

Titre: Construction rapide des g-fonctions pour l'optimisation financière
Title: sous incertitude des systèmes géothermiques

Auteur: Bernard Dusseault
Author:

Date: 2020

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Dusseault, B. (2020). Construction rapide des g-fonctions pour l'optimisation
Citation: financière sous incertitude des systèmes géothermiques [Thèse de doctorat,
Polytechnique Montréal]. PolyPublie. <https://publications.polymtl.ca/5421/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/5421/>
PolyPublie URL:

**Directeurs de
recherche:** Philippe Pasquier
Advisors:

Programme: Génie minéral
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Construction rapide des *g-functions* pour l'optimisation financière sous
incertitude des systèmes géothermiques**

BERNARD DUSSEAUT

Département des génies civil, géologique et des mines

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*
Génie minéral

Septembre 2020

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Cette thèse intitulée :

**Construction rapide des *g-functions* pour l'optimisation financière sous
incertitude des systèmes géothermiques**

présentée par **Bernard DUSSEAULT**

en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Michel BERNIER, président

Philippe PASQUIER, membre et directeur de recherche

Massimo CIMMINO, membre

Bertrand FRANÇOIS, membre externe

"Rien ne vaut la recherche lorsqu'on veut trouver quelque chose"

J.R.R. Tolkien

REMERCIEMENTS

Mes remerciements vont tout d’abord à mon directeur de recherche, le professeur Philippe Pasquier, pour sa grande patience au fil des années, les deux conférences auxquelles j’ai pu participer, son dévouement visible pour ses étudiants, sa rigueur et sa disponibilité malgré les embûches de la vie, un laboratoire mobile, une année sabbatique et une chaire de recherche. Un merci tout spécial pour m’avoir fait découvrir l’univers incroyablement enrichissant de la géothermie de basse température qui m’était totalement inconnu avant que je m’embarque, presque à l’aveuglette, dans les études supérieures.

Merci également au professeur Denis Marcotte qui, sans même être mon co-directeur officiellement, m’a plusieurs fois conseillé durant mon parcours et avec qui j’ai fini par écrire deux articles.

Je remercie aussi les membres du jury, les professeurs Michel Bernier, Massimo Cimmino et Bertrand François qui ont pris le temps de lire et d’évaluer ma thèse.

Je profite aussi de l’occasion pour saluer tous les professeurs, les étudiants et les techniciens du département que j’ai côtoyés à travers mes études du baccalauréat au doctorat. En ordre chronologique, je tiens particulièrement à saluer mes collègues que sont Alain, Fanny, Pascale, Émilie, Léo, Gabrielle, Vincent, Pier-Luc, Karine, Gabriel, Stéphanie et Jules pour les excellents moments que l’on a partagés ensemble à Polytechnique et à l’extérieur.

Je ne pourrais bien entendu pas terminer sans remercier du fond du cœur ma femme merveilleuse, Marie-Michèle, qui m’a écouté pendant des années parler de Matlab, de réseau neuronal et d’optimisation non-linéaire. Merci également à ma mère, dont la mémoire est toujours avec nous, et à mon père et qui m’ont supporté tout au long de ma vie, m’ont encouragé à poursuivre à l’université même si depuis longtemps ils ne pouvaient plus m’aider dans mes devoirs.

La réalisation de ce projet a été possible grâce à des bourses d’études supérieures et de l’aide financière octroyées par le Conseil de recherches en sciences naturelles et génie du Canada (CRSNG). Leur soutien m’a permis de me concentrer pleinement sur mon projet durant mon passage aux cycles supérieurs. Qu’ils en soient ici remerciés.

RÉSUMÉ

Le couplage de thermopompes avec un échangeur de chaleur souterrain permet de réaliser d'importantes économies d'énergie par comparaison aux appareils de chauffage et de climatisation traditionnellement installés. Au Québec, la majorité des projets utilisant cette géothermie de basse température ont recourt à des échangeurs de chaleur en boucles fermées verticales où un fluide caloporteur circule à travers un ensemble de forages en utilisant un circuit fermé de tuyaux. Ces systèmes sont efficaces, mais ont le désavantage important d'être coûteux à l'achat et à l'installation. Pour pallier ce problème, il est courant de concevoir des systèmes hybrides où les pompes à chaleur géothermiques ne permettent pas de couvrir la totalité de la pointe de demande thermique des bâtiments. Utilisés conjointement, des équipements moins dispendieux de chauffage et de climatisation d'appoint permettent de satisfaire les besoins thermiques résiduels lorsque nécessaire.

Faisant intervenir davantage d'équipements mécaniques, de paramètres et de variables, ces systèmes géothermiques hybrides sont donc plus difficiles à dimensionner. La taille de l'échangeur de chaleur et le nombre d'appareils auxiliaires se doivent d'être bien ajustés pour assurer la rentabilité du système. Ainsi, à cause de leur complexité, le dimensionnement de systèmes géothermiques hybrides est particulièrement sensible aux incertitudes qui planent sur ce genre de projet : évolution du coût de l'énergie et de la demande en chaleur des bâtiments, propriétés thermiques du sol et coûts initiaux de construction. L'objectif général de cette thèse est de proposer une méthodologie qui permet de mitiger ces risques lors de la conception de systèmes géothermiques hybrides en boucles fermées verticales.

Premièrement, une nouvelle méthode permettant le calcul de la *g-function* du champ de captage géothermique a été élaborée. Cette méthode mise sur une reformulation des problèmes de superposition spatiale et temporelle des contributions thermiques de chaque forage en utilisant une approche par matrice de bloc pour en accélérer grandement la résolution. De plus, l'utilisation de polynômes de Tchebychev pour l'approximation du modèle de la source linéique finie permet son intégration analytique plutôt que de façon numérique, accélérant encore davantage la construction de la fonction de transfert. Les contributions de ces deux approches sont investiguées séparément, puis une fois combinées au sein d'une méthode d'optimisation. Leur combinaison permet d'obtenir la fonction de transfert d'un champ régulier de 50 forages sur un horizon de 40 ans en seulement 0.46 seconde, soit largement plus rapidement que les autres approches ayant déjà été publiées. Ces gains de temps ont été nécessaires pour permettre l'accomplissement du deuxième sous-objectif de recherche de la thèse.

Ensuite, pour accélérer au maximum le temps de calcul de la fonction de transfert d'un champ de forages, un réseau neuronal servant à leur estimation a été développé puis entraîné. Pour qu'il possède de bonnes capacités d'approximation et de généralisation, une base de données contenant 500 000 exemples d'entraînement a été assemblée. Compte tenu du nombre de fonctions à calculer, la méthode basée sur une approche matricielle établie pour le premier sous-objectif de cette thèse a été utilisée. Après avoir été entraîné durant 30 jours, ce réseau était capable d'approximer et de construire en une moyenne de 17 centièmes de secondes la *g-function* de champs comptant entre 1 à 10 forages. La rapidité d'approximation de ce réseau a été mise à profit dans le cadre du troisième sous-objectif de cette thèse lors du design par optimisation stochastique de systèmes géothermiques hybrides : un processus requérant le calcul itératif de dizaines de milliers de *g-functions*. Au delà du réseau particulier développé pour cet objectif, les travaux présentés dans cet article démontrent qu'il est possible d'utiliser des réseaux neuronaux pour l'estimation de fonction de transferts à long terme et ce, en fonction des besoins précis de l'utilisateur.

Troisièmement, une étude de cas a été réalisée portant sur l'impact de l'utilisation d'un algorithme stochastique permettant la conception d'un système géothermique hybride par programmation mixte en nombres entiers sous incertitude. Cet algorithme utilise conjointement le gradient conjugué contraint, une méthode de séparation et évaluation (*branch-and-bound*) et un balayage pour s'assurer que les forages et les thermopompes évoluent en nombres entiers durant le processus d'optimisation. De plus, pour assurer l'intégration du risque au processus de dimensionnement, un indicateur financier stochastique jamais utilisé en géothermie de basse température a été employé. Empruntée aux domaines du génie civil et de la finance, la valeur actuelle nette à risque permet de converger vers un optimum en utilisant la valeur actuelle nette anticipée moyenne du projet lors des pires scénarios potentiels. La réalisation de cette étude de cas a permis d'entrevoir que ce nouvel indicateur financier permet un dimensionnement conservateur des systèmes hybrides, réduisant grandement les pertes lors des pires scénarios sans diminuer de façon significative les gains lors des scénarios plus favorables. Cette façon de procéder est une alternative souhaitable aux analyses de sensibilité dans un contexte incertain, car elle s'attaque à mitiger l'incertitude à la source au lieu de la quantifier une fois un design proposé. Pour aider les lecteurs à implémenter la valeur actuelle nette à risque dans leurs propres optimisation, une étude paramétrique a été réalisée pour permettre d'établir l'impact des trois paramètres régissant cet indicateur financier.

Pour finir, les contributions apportées dans cette thèse permettront aux professionnels d'effectuer le dimensionnement sécuritaire de systèmes géothermiques hybrides en boucles fermées verticales. Ainsi, en misant sur l'efficacité de cette technologie, il sera possible de réduire la quantité d'énergie qui est consacrée au chauffage et à la climatisation.

ABSTRACT

The coupling of heat pumps with a geothermal heat exchanger enables significant energy savings when compared to traditional heating and cooling options. In Quebec, the majority of projects that incorporate low temperature geothermal energy choose to install vertical closed loop systems where a heat carrying fluid circulates through a collection of boreholes using a piping network. These systems are efficient but suffer from high upfront costs. To bypass this problem, it is common practice to install hybrid systems where geothermal heat pumps are undersized and cannot cover the building's energy needs in peak periods. Used alongside the heat pumps are auxiliary heating and cooling equipment that cover the remaining energy needs when necessary.

Because they involved more mechanical components, parameters and variables, hybrid systems are more difficult to size properly. The size of the geothermal heat exchanger and the auxiliary equipment must be carefully balanced to insure the financial profitability of the whole system. Therefore, because of their overall complexity, hybrid systems can be particularly sensitive to uncertainties such as the time variation of energy needs and energy costs, the thermal properties of the ground and initial construction costs. The main objective of this thesis is to propose a methodology that allows for the mitigation of those uncertainties during the sizing of hybrid vertical closed-loop systems.

First, a new calculation method for the assessment of g-functions for geothermal heat exchangers was developed. This method relies on a reformulation of the temporal and spatial superposition problems of thermal perturbations induced by individual boreholes in a field using a block matrix formulation. This greatly improved the calculation speed of g-function. Furthermore, the use of Tchebychev polynomials to approximate the finite line source model allowed its analytical integration instead of a numerical integration, which proved to be faster. The contributions of these two approaches were investigated both independently and together. Their combination enables the calculation of a g-function for a regular field of 50 boreholes over a period of 40 years in only 0.46 seconds, which is much quicker than published methods. Gains in computation times were mandatory in order to reach the second sub-objective of this thesis.

Second, to accelerate even further the calculation of g-functions, an artificial neural network dedicated to their estimation was developed and trained. To insure it possesses good approximation and generalization capabilities, a training database of 500000 g-functions was assembled. Because of the size of this database, the block matrix formulation method that

was developed for the first sub-objective of this thesis was used for its creation. After a training phase that lasted 30 days, this artificial neural network was able to approximate and reconstruct the g-function of a field that varies between one and ten boreholes in a matter of 17 hundredths of a second. This approximation speed was later put to use in the third sub-objective of this thesis during the design by stochastic optimization of hybrid ground-coupled heat pump systems: a process that requires the calculation of tens of thousands of g-functions. Apart from the sole development of another g-function construction tool, this work showed that it is possible to use artificial neural networks for the estimation of long-term g-function according to the specific needs of the user.

Third, a case study was conducted on the impacts of using a stochastic algorithm that allows for the design by optimization of hybrid ground-coupled heat pump systems with mixed-integer programming under uncertainty. This algorithm uses together the constraint conjugate gradient method, branch-and-bound and sweeping to ensures that boreholes and heat pumps stay on integer numbers all throughout the optimization process. Additionally, to integrate risk management into the design process, a stochastic financial indicator never used before in the field of geothermal energy was implemented. Borrowed from civil engineering and the finance sector, the net present value-at-risk reaches an optimum by considering multiple net present values of a project under the worst foreseeable circumstances. This case study showed that this new financial indicator sized conservative hybrid systems, reducing considerably financial losses during worst cases without reducing significantly the financial gains for the best cases. This design strategy is a better alternative to sensitivity analyses when faced with uncertain conditions, because it focuses on mitigating uncertainties instead of trying to quantify the resulting risk afterwards. To guide the readers into implementing the net present value-at-risk in their own optimizations, a parametrical study was also performed to assess the impacts on proposed designs of the three parameters associated with this financial indicator.

Finally, the original contributions of this thesis will give designers the possibility to safely size hybrid vertical closed-loop ground-coupled heat pump systems when faced with uncertainties. Thus, by focusing on the known efficiency of this technology, the overall amount of energy dedicated to heating and cooling can be reduced.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vii
TABLE DES MATIÈRES	ix
LISTE DES TABLEAUX	xii
LISTE DES FIGURES	xiv
LISTE DES SIGLES ET ABRÉVIATIONS	xviii
CHAPITRE 1 INTRODUCTION	1
CHAPITRE 2 REVUE DE LITTÉRATURE	5
2.1 Méthodes de simulation	5
2.1.1 Transferts de chaleur dans les sols	6
2.1.2 Couplage entre sol et bâtiment	11
2.2 Dimensionnement de systèmes	17
2.2.1 Algorithmes d'optimisation	19
2.2.2 Variables optimisées	21
2.3 Risques et incertitudes	24
2.3.1 Quantification	24
2.3.2 Mitigation	27
CHAPITRE 3 DÉMARCHE DE L'ENSEMBLE DU TRAVAIL	30
CHAPITRE 4 ARTICLE 1 : A BLOCK MATRIX FORMULATION FOR EFFICIENT G-FUNCTION CONSTRUCTION	32
4.1 Abstract	32
4.2 Introduction	32
4.3 A block matrix formulation for g-functions	34

4.3.1	Block matrix formulation	36
4.3.2	Chebyshev polynomials approximation	41
4.4	Verification with Two Independent Algorithms	45
4.4.1	Sequential Algorithm	45
4.4.2	Iterative Algorithm	46
4.4.3	Time stepping schemes	47
4.5	Validation Scenarios	48
4.6	Results and discussion	50
4.6.1	Computing efficiency	51
4.6.2	Gain of using Chebyshev polynomials	52
4.6.3	Accuracy and time stepping	53
4.7	Conclusion	55
4.8	Acknowledgments	56
4.9	Appendix A. Numerical example	56

CHAPITRE 5 ARTICLE 2 : EFFICIENT G-FUNCTION APPROXIMATION WITH ARTIFICIAL NEURAL NETWORKS FOR A VARYING NUMBER OF BORE-HOLES ON A REGULAR OR IRREGULAR LAYOUT

5.1	Abstract	62
5.2	Introduction	62
5.3	Construction of a g-function with a block matrix formulation	65
5.4	Methodology	68
5.4.1	Architecture	69
5.4.2	Inputs of the ANN	70
5.4.3	Outputs of the ANN	72
5.4.4	Randomization and coverage of the training data	72
5.4.5	Size of the training database and training algorithm	76
5.5	Results and discussion	76
5.5.1	Validation of the ANN for regular GHEs	78
5.5.2	Validation of the ANN for irregular GHEs	81
5.5.3	Computation times	83
5.6	Conclusion	84

CHAPITRE 6 ARTICLE 3 : USAGE OF THE NET PRESENT VALUE-AT-RISK TO DESIGN GROUND-COUPLED HEAT PUMP SYSTEMS UNDER UNCERTAIN SCENARIOS

6.1	abstract	87
-----	--------------------	----

6.2	Introduction	87
6.3	Design-based Optimization Under Uncertainty	93
6.3.1	Direct Simulation of Hybrid GCHP Systems	93
6.3.2	Efficient Computation of g-functions	95
6.3.3	Net Present Value-at-Risk	97
6.3.4	Optimization of a Mixed-Integer Programming Problem	99
6.3.5	Subsampling of the NPV Distribution	100
6.4	Case Study Description	102
6.4.1	Fixed design parameters used to compute the NPVaR	103
6.4.2	Uncertain Design Parameters	105
6.4.3	Design Variables Optimized	107
6.5	Algorithm Validation	108
6.5.1	Deterministic use of the NPVaR	109
6.5.2	Stochastic use of the NPVaR	110
6.6	Gains at Using the NPVaR	111
6.6.1	Actual NPV and Risk Management	112
6.6.2	Impact on the design and operation of GCHP systems	115
6.7	Conclusion	121
6.8	Appendix - Fine-tuning of the NPVaR related parameters	123
6.8.1	Impact of the fraction of the worst outcomes α	123
6.8.2	Impact of the buffer β	124
6.8.3	Impact of the sample size ω	124
CHAPITRE 7	DISCUSSION GÉNÉRALE	127
7.1	Article 1 : A Block Matrix Formulation for Efficient g-Function Construction	127
7.2	Article 2 : Efficient g-Function Approximation with Artificial Neural Networks for a Varying Number of Boreholes on a Regular or Irregular Layout	129
7.3	Article 3 : Usage of the Net-Present-Value-at-Risk to Design Ground Source Heat Pump Systems Under Uncertain Scenarios	132
CHAPITRE 8	CONCLUSION ET RECOMMANDATION	134
8.1	Synthèse des travaux	134
8.2	Améliorations futures	135

LISTE DES TABLEAUX

2.1	Nombres et natures des variables utilisées par différents auteurs dans le cadre de dimensionnement par optimisation.	22
4.1	Thermal and geometrical parameters used for validation Scenario 1, 2 and 3.	48
4.2	Summary of the validation scenarios.	50
4.3	Computation times when using the sequential method with Chebyshev polynomials and the block matrix formulation without Chebyshev polynomials.	53
4.4	Computation times and maximum relative errors for Scenario 1, 2 and 3. The sequential method is used without Chebyshev polynomials while the block matrix formulation is used with Chebyshev polynomials. . .	55
4.5	Parameters used in the numerical example of Annex A	56
4.6	g-function presented according to the time and characteristic time of the simulation for the numerical example of Annex A.	60
5.1	Ranges of input parameters of the ANN.	73
5.2	Example of Halton coordinates.	75
5.3	Reduced ranges of input parameters during the validation phase. . . .	77
5.4	Computation time for a 100-year simulation	84
6.1	Costs considered for the calculation of the NPV and NPVaR of the case study analyzed in Sections 4 and 5.	104
6.2	Boundaries and initial guesses for the 5 design parameters optimized in this work.	107
6.3	Ranges and increments used during the sweeping of the five design variables.	108
6.4	Results and computational time for the sweep, the deterministic and stochastic optimization that use the NPV and NPVaR as objective function.	111
6.5	Energy and power coverage of the designs optimized with the NPV and NPVaR ($\alpha = 0.1$).	115
6.6	Annual and monthly operating costs for the months of January and July and payback time for the systems design by NPVaR ($\alpha = 0.1$), by NPV and the reference system.	117
6.7	Impact of the NPVaR related parameters on the optimum location. .	123

7.1	Taille de la base de données d'entraînement requise en fonction du nombre de forages maximal accepté par le réseau neuronal.	130
-----	---	-----

LISTE DES FIGURES

1.1	Exemple d'un système géothermique en boucles fermées verticales hybride incluant quatre paramètres incertains (en gras).	2
1.2	Exemple de la fonction de transfert à court terme et à long terme d'un champ de captage géothermique en fonction du temps t et du temps caractéristique $t_s = H^2/9\alpha$ (avec H la longueur des puits et α la diffusivité thermique).	4
2.1	Comparaison entre le maillage de la différence finie (gauche) et des éléments finis (droite).	7
2.2	Représentation schématique de la source linéique infinie comme une superposition infinie de sources ponctuelles.	8
2.3	Représentation schématique d'une thermopompe en mode chauffage. .	11
2.4	Évolution du COP et de la CAP d'une thermopompe, en chauffage et en climatisation, en fonction de la température d'entrée d'eau (EWT pour <i>Entering Water Temperature</i>).	12
2.5	Configurations possibles de distribution de fluide caloporteur de la sortie d'une thermopompe à son retour : a) parallèle, b) série ou c) mixte.	16
2.6	Démonstration de la relation entre la VaR et la CVaR. La CVaR constitue la moyenne des pires cas situés entre la VaR et la perte maximale pouvant être subie.	28
4.1	Illustration of the thermal interactions between the boreholes of a GHE.	35
4.2	The six first Second type Chebyshev polynomials in their original unshifted domain.	42
4.3	Sampling points used by Chebyshev polynomials superimposed on the approximation the time derivative of the FLS model.	43
4.4	Sampling points used by Chebyshev polynomials superimposed on the FLS model.	43
4.5	Overview of the block matrix algorithm.	44
4.6	Configuration of the GHE in all scenarios (from left to right, 1 to 3).	49
4.7	Transfer functions obtained for all three scenarios (from top to bottom, 1 to 3).	51
4.8	Relative errors committed during all three scenarios (from top to bottom, 1 to 3).	54

5.1	Short- and long-term transfer functions according to the characteristic time t_s designed to provide directly the EWT of a heat pump. The short-term transfer function is obtained by the ANN presented by Pasquier <i>et al.</i> (2018) while the long-term transfer function is obtained after correction of a g-function constructed by the block matrix formulation of Dusseault <i>et al.</i> (2017). Adapted from Pasquier <i>et al.</i> (2018).	64
5.2	The mathematical model used for the creation of single artificial neuron.	68
5.3	Proposed artificial neural network with 2 fully connected hidden layers used to approximate a g-function.	70
5.4	Illustration of how the distances within a GHE are extracted and formatted to construct input vector x	71
5.5	Comparison of 250 points distributions obtained with Halton sequence (left) and uniform number generator (right).	75
5.6	Synthetic hourly heat load signal presented for one year and used to assess the error made by the ANN on the fluid temperature. This signal is replicated five times to simulate over 5 years the heat carrying fluid temperatures of the Case Study. . .	77
5.7	Visual comparison of the g-functions provided by the ANN and the BMF for the best, worst and median case amongst the 2500 regular.	79
5.8	Geometry of the GHE for the best (blue), median (green) and worst (red) case. .	79
5.9	Distribution of the maximal absolute relative errors and maximal absolute errors between the g-functions provided by the ANN and the block matrix formulation for regular GHEs.	80
5.10	Distribution of the maximal relative errors between the fluid temperatures provided usign the ANN approximated g-functions and the block matrix formulation provided g-functions over 100 years for regular (blue) and irregular (red) GHEs. .	80
5.11	Visual comparison of the g-functions provided by the ANN and BMF for the best, worst and median cases amongst the 2500 irregular GHEs.	82
5.12	Configurations of the GHE on the best (blue), median (green) and worst (red) cases for irregularly placed boreholes.	82
5.13	Distribution of the maximal absolute relative errors and maximal absolute errors between the g-functions provided by the ANN and the block matrix formulation for irregular GHEs.	83
6.1	Main components of a simplified hybrid ground-coupled heat pump system where the five design parameters (H , n_b , n_{HP} , L_x , L_y) are in bold and the four uncertain parameters considered in this work are underlined.	88

6.2	The net present value-at-risk (NPVaR) corresponds to the average of the worst fraction (α) of the outcomes of a NPV distribution composed of ω samples. The smaller the α value, the more cases that feature extreme losses are used to compute the NPVaR.	92
6.3	Hourly simulation output obtained with the method described in Section 2.1 for a case where the number of installed GCHPs is 6, the minimum temperature limit on the source side is $-2^{\circ}C$ and where the GHE is composed of 9 boreholes of 150 m. From top to bottom, the figures show the number of active heat pumps in operation, the EWT, the COP and CAP, the hourly cumulative consumption and peak monthly electric consumption and finally the cumulative monthly operating costs of the whole system.	95
6.4	Simplified flow-chart of the algorithm used to optimize the NPVaR with a detailed breakdown of the subsampling scheme retained to speedup its evaluation.	101
6.5	Building loads of a 5-storey long-term care facility located in Montreal, Canada. The first value corresponds to January 1 st , while the negative and positive values correspond to heating and cooling loads respectively.	102
6.6	Percentage of the energy needs shown in Fig. 6.5 that can be covered according to the relative capacity of the GCHP system. The arrows points towards the three systems analyzed in Section 6.6.	103
6.7	Coefficients of performance (<i>COP</i>) and capacity (<i>CAP</i>) of the GCHPs used in the case study.	105
6.8	Probability density functions used to describe the four uncertain design parameters used in this work. The PDFs express, with the exception of the monthly electricity cost, the probability that the values retained in a base case scenario be multiplied by a given positive value.	106
6.9	NPV (upper figure) and NPVaR (lower figure) obtained by sweep of the design variables n_b , n_{HP} , H , L_x and L_y to assess the global optimum of each cost function. The black horizontal dashed line represents the best NPV or NPVaR reached by optimization while the red circle shows the location of the sweeping optimum.	109

6.10	Close up view of the upper part of Fig. 6.9 for $n_b = 9$. Each color corresponds to a different borehole length (black for $H = 50$ m, blue for $H = 100$ m and red for $H = 150$ m as highlighted by the ellipses), each line style is for a different value of n_{HP} while the sweepings of L_x and L_y is illustrated by the boxes.	110
6.11	Actual NPVs after 10 years of operating the hybrid GCHP systems calculated using the design proposed by the stochastic NPVaR using $\alpha = 0.1$, $\beta = 0.05$ and $\omega = 100$ (top), using $\alpha = 0.2$, $\beta = 0.05$ and $\omega = 100$ (middle) and using the deterministic algorithm (bottom) based on the NPV as objective function. The blue histograms are associated with the statistical distributions of Fig. 6.8, and the red lines to the bounded uniform distributions.	113
6.12	Cumulative distribution functions of the NPVs shown in the histograms of Fig. 6.11. The three above correspond to the original distribution of uncertain parameters and the three below are for the flat distributions.	114
6.13	Semi-logarithmic presentation of the power versus energy coverage curve associated with the long term care center of the case study. In red, the derivative is presented alongside the original curve in blue without its own axis simply to show its behavior.	116
6.14	Yearly energy billing breakdowns for the operating costs of the NPV (System type 1) and NPVaR-based designs (System type 2) and a reference system without GCHPs (System type 3).	118
6.15	Cashflow differences between the NPV and NPVaR-based designs. The dotted black lines illustrates the impact of the discount rate and electricity tariffs changes on the value of cashflow.	119

LISTE DES SIGLES ET ABRÉVIATIONS

AAT	All-at-a-time
ANN	Artificial neural networks
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BACCO	Bayesian analysis of computer code outputs
B & B	Branch-and-bound
BHE	Borehole heat exchanger
BMF	Block matrix formulation
CAP	Capacity
CF	Cash flow
COP	Coefficient of performance
CPU	Central processing unit
CRSNG	Conseil de recherches en sciences naturelles et génie du Canada
CVaR	Conditional value-at-risk
erf	Error function
EWT	Entering water temperature
FAST	Fourier amplitude sensitivity test
FFT	Fast Fourier transform
FLS	Finite-line source model
GA	Genetic algorithm
GCHP	Ground-coupled heat pump
GHE	Ground / geothermal heat-exchanger
HP	Heat pump
Hy-GCHP	Hybrid ground-coupled heat pump
LWT	Leaving water temperature
MIP	Mixed-integer programming
MINLP	Mixed-integer non-linear programming
NSERC	National sciences and engineering research council of Canada
NPV	Net present value
NPVaR	Net present value-at-risk
NSGA	Nondominated sorting genetic algorithm
OAT	One-at-a-time
PDF	Probability density function

PSO	Particule swarm optimization
SCI	Source cylindrique infinie
SCIM	Source cylindrique infinie mouvante
SLI	SOurce linéique infinie
SLF	Source linéique finie
SLFM	Source linéique finie mouvante
SPEDE	Système de plafonnement et d'échange de droits d'émission
TRC	Thermal resistance capacity
VAN	Valeur actuelle nette
VANaR	Valeur actuelle nette-à-risque
VaR	Value-at-risk / Valeur-à-risque

CHAPITRE 1 INTRODUCTION

En puisant et en rejetant de l'énergie dans le sol à l'aide d'un échangeur de chaleur souterrain, il est possible de chauffer et de climatiser un bâtiment, et ce aussi bien au plus froid de l'hiver que durant les canicules de l'été. Il existe différents types d'échangeurs de chaleur souterrains. Au Québec, les systèmes en boucles fermées verticales représentent la majorité des systèmes installés, soit 82.9% en 2010 (CCEG, 2012). Dans ces systèmes, les transferts de chaleur sont réalisés par conduction en faisant circuler un fluide caloporteur, généralement un mélange d'eau glycolée pour prévenir le gel, dans des tuyaux installés en boucles fermées dans des forages.

Ces systèmes permettent de réaliser des économies d'énergie (Omer, 2008) grâce au couplage d'une pompe à chaleur à l'échangeur de chaleur géothermique. Une thermopompe géothermique, à l'instar d'un réfrigérateur, permet d'extraire de la chaleur d'un milieu froid pour l'acheminer vers un milieu chaud en activant son compresseur (Self *et al.*, 2013). Ainsi, en consommant principalement l'énergie électrique nécessaire à l'activation des pompes à chaleur, un bâtiment peut aller chercher dans le sol environnant des kilowattheures de chauffage ou de climatisation gratuits à l'année, d'où les économies d'énergies par rapport aux méthodes de chauffage et de climatisation où chaque kilowattheure utilisé doit être payé.

Comblent l'entière des besoins d'un bâtiment par la géothermie est techniquement possible, mais requiert d'importants investissements initiaux. À cet effet, Hénault *et al.* (2016) ont démontré qu'un système sous-dimensionné pour un bâtiment commercial dans la région de Montréal peut coûter plus de 250000\$. Combiné aux prix planchers de l'électricité et du gaz naturel au Québec (Bernard *et al.*, 2011), il est facile de concevoir pourquoi les acteurs du milieu optent souvent pour l'installation de systèmes hybrides. Ces systèmes, représentés schématiquement à la Figure 1.1, permettent de réduire substantiellement la taille de l'échangeur de chaleur souterrain en ayant recours à des systèmes de chauffage et de climatisation d'appoint qui s'activent en support des thermopompes lors des périodes de pointe (Klein *et al.*, 2014). Dimensionner un système hybride est une tâche complexe qui requiert de tenir compte d'un grand nombre de paramètres afin d'estimer la rentabilité du système. Parmi ces paramètres, certains sont inférés sur la base de jeux de données partiels, mesurés, estimés en se basant sur des modèles numériques, tirés de la littérature, connus à force d'expériences ou tout simplement posés.

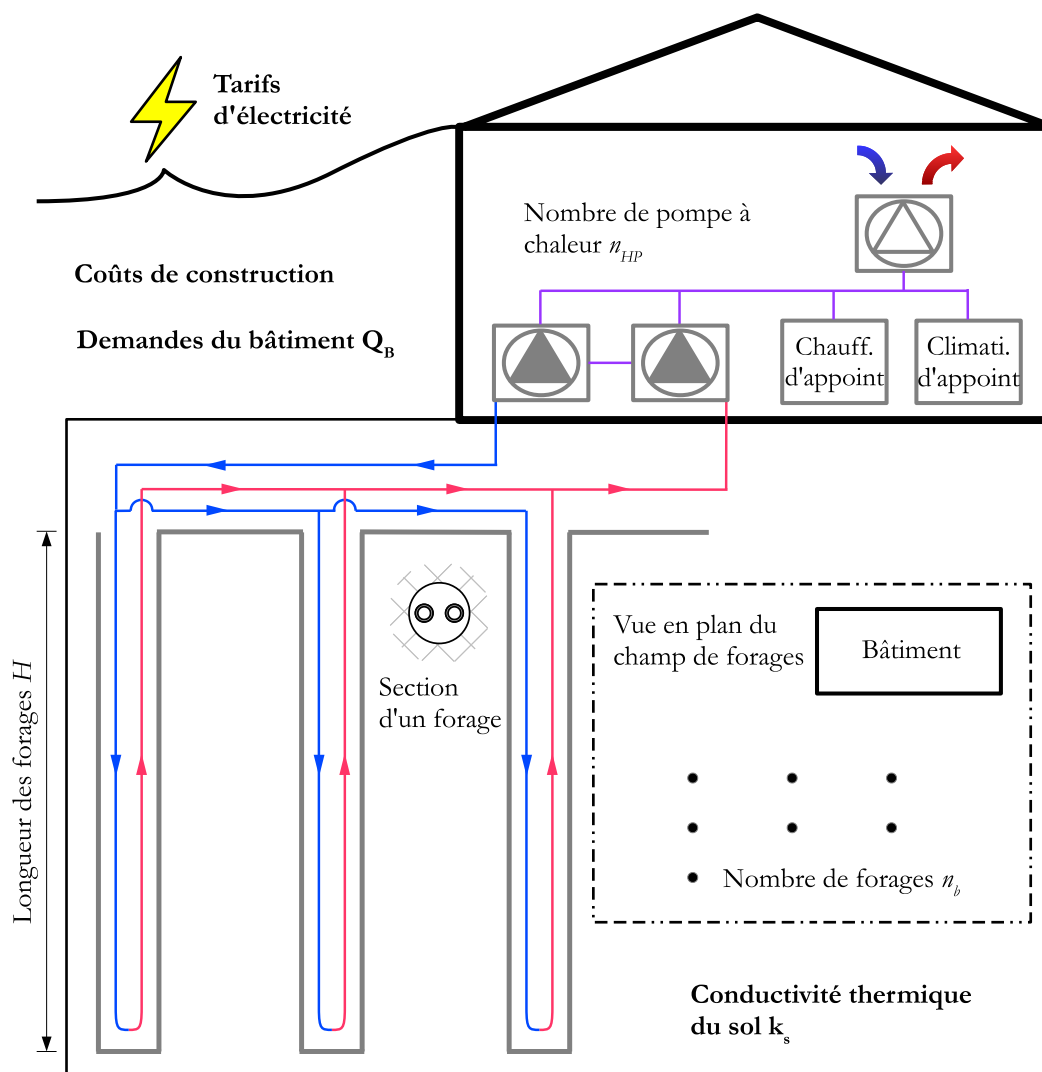


Figure 1.1 Exemple d'un système géothermique en boucles fermées verticales hybride incluant quatre paramètres incertains (en gras).

Une des façons de dimensionner de tels systèmes est d'avoir recours aux indicateurs financiers dans le cadre d'un processus d'optimisation (qu'il soit automatisé à l'aide d'un algorithme ou effectué manuellement). Utiliser un algorithme d'optimisation conjointement avec un module qui permet la simulation horaire d'un système hybride permet d'exprimer les besoins énergétiques horaires des systèmes auxiliaires.

Une fois les besoins énergétiques du système simulés, son coût d'opération horaire peut être calculé et actualisé en une valeur unique de référence. Additionnée aux coûts de construction initiaux du système, cela permet de calculer la valeur actuelle nette du système (VAN) et de la comparer à celle d'un système de référence (par exemple, des résistances électriques couplées

à une thermopompe air-air ou encore un système de chauffage central au gaz naturel). Utiliser la VAN comme fonction objectif est un choix intéressant, car seule une analyse rigoureuse prenant en compte des indicateurs financiers permet d'établir la rentabilité d'un système énergétique (Adnot *et al.*, 2012).

Or, des paramètres incertains apparaissant en gras à la Figure 1.1 (variation des tarifs d'électricité dans le temps, propriétés thermiques du sol, coûts de construction et demandes thermiques du bâtiment) introduisent justement une grande part d'incertitude dans le calcul de la VAN d'un système. Bien entendu, si la VAN est sujette à une incertitude, il en découle naturellement qu'un design se basant sur son calcul sera lui aussi imprécis et donc risqué. Le recours à des approches stochastiques semble de ce fait une approche intéressante pour tenter de mitiger à la source les risques inhérents à ces incertitudes. Cette hypothèse de recherche est au cœur de la présente thèse qui vise à développer une approche pour tenir compte du risque lors du dimensionnement de systèmes géothermiques hybrides en boucles fermées verticales.

Bien que cette approche soit prometteuse, elle ne fait qu'exacerber un problème qui existe déjà au stade du design par optimisation, soit les temps de calcul prohibitifs. Pour contourner cette limitation et rendre viable un algorithme d'optimisation stochastique, il incombe donc de travailler en parallèle à réduire significativement les temps de calculs associés au dimensionnement de systèmes dans un contexte déterministe. À cet égard, une piste de solution intéressante a été soulevée par Hénault (2015) en conclusion de son mémoire de maîtrise. Il a remarqué que le calcul de la fonction de transfert des échangeurs de chaleur souterrains consomme jusqu'à 75% du temps total associé au dimensionnement. Cela est dû au fait qu'à travers le processus d'optimisation, le calcul de milliers voire de dizaines de milliers de fonctions de transfert est requis. Ainsi, réussir à réduire ne serait-ce qu'un peu le temps nécessaire pour construire une fonction de transfert aurait des répercussions majeures sur le temps requis pour la complétion du processus d'optimisation.

Une fonction de transfert, baptisée *g-function* par Eskilson (1987) lorsqu'appliquée aux échangeurs de chaleur souterrains, est un modèle mathématique qui permet de tracer une relation entre l'entrée et la sortie d'un système linéaire. Un exemple de *g-function* est présenté à la Figure 1.2. En géothermie, on la définit comme étant la corrélation entre les taux d'extraction de chaleur prescrits dans un échangeur de chaleur souterrain et la température moyenne aux parois des forages de ce même échangeur. L'élaboration de nouvelles méthodes de calcul de fonctions de transfert étant un domaine actif en recherche depuis les années 80, cette thèse apportera une contribution originale sur le sujet en puisant dans les dernières avancées en mathématiques et en programmation. Plus précisément, les problèmes de superpositions spatiale et temporelle seront reformulés et des polynômes seront utilisés pour accélérer la

résolution du modèle de la source linéique finie de Claesson et Javed (2011).

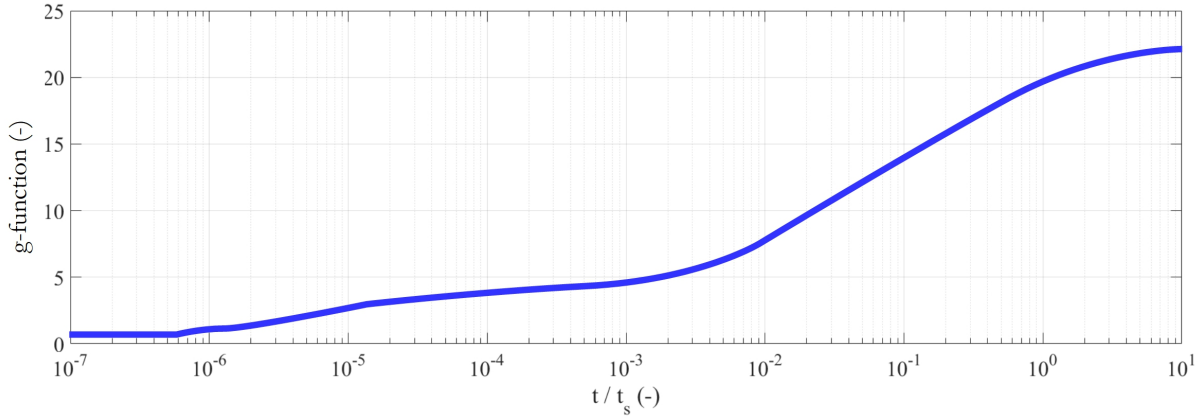


Figure 1.2 Exemple de la fonction de transfert à court terme et à long terme d'un champ de captage géothermique en fonction du temps t et du temps caractéristique $t_s = H^2/9\alpha$ (avec H la longueur des puits et α la diffusivité thermique).

Comme deuxième objectif, un réseau neuronal sera créé puis entraîné afin d'accélérer davantage le calcul de fonctions de transfert en utilisant la capacité et la puissance d'apprentissage des réseaux de neurones artificielles. La base de données de fonctions de transfert requis pour son entraînement sera assemblée en utilisant la méthode développée pour le premier objectif. Le but sera de rendre ce réseau neuronal aussi polyvalent que possible afin qu'il couvre un maximum de cas de figure pouvant survenir lors d'un processus de dimensionnement par optimisation (nombre variable de forages et de thermopompes, disposition des puits, propriétés thermiques du sol, taille et rayon des puits, etc.).

Comme dernier objectif, un algorithme d'optimisation stochastique sera développé pour procéder au dimensionnement de systèmes géothermiques hybrides en boucles fermées verticales. Cet algorithme utilisera au maximum le réseau neuronal développé durant le deuxième objectif pour l'évaluation de fonctions de transfert, recourant au premier lorsque cela sera impossible. Finalement, cet algorithme servira à réaliser une étude de cas se déroulant à Montréal pour démontrer l'impact de recourir à une méthode stochastique sur le temps de retour sur investissement, les risques encourus, la facilité d'implémentation et le temps de calcul.

CHAPITRE 2 REVUE DE LITTÉRATURE

Tout modèle numérique est limité dans la qualité de ses prédictions par la qualité des données qui y entrent. La géothermie de basse température, située à la frontière entre la mécanique du bâtiment et le génie géologique, n'échappe pas à ce principe. En effet, le dimensionnement d'un échangeur de chaleur souterrain en boucle fermée verticale et de tous les équipements mécaniques requis pour son opération nécessite l'utilisation d'une grande quantité de données concernant l'immeuble et le milieu géologique dont toutes ne seront pas connues avec la même certitude. À l'échelle d'un projet, toutes ces incertitudes finissent par se cumuler et peuvent mener à des conclusions erronées si aucune mesure n'est prise pour en tenir compte. Interviennent alors les outils stochastiques.

Les méthodes stochastiques, par exemple les simulations de Monte-Carlo, permettent d'intégrer les incertitudes aux simulations dans le but de quantifier l'incertitude d'une valeur d'intérêt. Malheureusement, cela est effectué au détriment du temps de calcul. En effet, ces méthodes nécessitent une quantité importante de simulations individuelles pour proposer une solution. Ainsi, il va sans dire que l'implémentation de méthodes stochastiques lors d'un problème aussi complexe que le dimensionnement de systèmes géothermiques hybrides va de pair avec l'utilisation des meilleures méthodes de simulations et des meilleurs algorithmes d'optimisation. Ces deux aspects seront couverts durant cette revue de littérature pour voir les limitations actuelles du domaine et mettre en lumière de potentielles lacunes. Pour finir, un tour d'horizon sur les façons dont l'incertitude est traitée en géothermie et dans d'autres domaines est effectué afin de situer les travaux de cette thèse sur l'échiquier de la recherche.

2.1 Méthodes de simulation

Modéliser l'interaction entre un échangeur de chaleur souterrain et un bâtiment est un processus itératif. Dans un premier temps, une quantité de chaleur est introduite ou extraite dans l'échangeur de chaleur à l'aide d'une thermopompe. Cette chaleur va être prélevée ou diffusée dans le sol environnant à l'aide d'un fluide qui circule dans un réseau de tuyaux parcourant les forages. La quantité de chaleur introduite dans le sol, la configuration et le nombre de forages, l'enfouissement des têtes de forage, leur diamètre, le nombre de boucles de tuyaux par forages, la conductivité thermique du sol, des tuyaux et du coulis et la nature du fluide sont autant de facteurs qui vont dicter de quelle façon et à quelle vitesse le sol va se réchauffer ou se refroidir.

Dans un deuxième temps, après avoir réagi avec le milieu géologique, le fluide caloporteur va retourner aux thermopompes. La température à laquelle il va entrer dans l'équipement mécanique va dicter les performances de ce dernier en jouant sur son coefficient de performance et sa capacité à fournir du chaud et du froid (ces deux paramètres sont abordés plus en détail à la Section 2.2.1). Ces deux paramètres vont à leur tour influencer la quantité de chaleur qui pourra être envoyée au sol par le fluide, recommençant le cycle. Ainsi, on comprend pourquoi la simulation de l'opération d'une thermopompe géothermique requiert à la fois le calcul des transferts de chaleur entre le fluide et le sol et entre le fluide et le bâtiment à travers la pompe à chaleur. Les deux prochaines sous-sections vont offrir un tour d'horizon des méthodes qui permettent de réaliser ces deux opérations.

2.1.1 Transferts de chaleur dans les sols

Il existe deux grandes familles de méthodes permettant de calculer les effets de l'utilisation d'un échangeur de chaleur souterrain sur le sol environnant et sur lui-même. Law et Dworkin (2016) démontrent en validant leur modèle par éléments finis à l'aide de données expérimentales que les simulations numériques sont la méthode la plus précise pour y arriver. Leur précision provient en grande partie de la souplesse inhérente des méthodes numériques leur permettant d'incorporer n'importe quelle stratigraphie à leur modèle et de représenter plus fidèlement la géométrie et la composition des forages. L'autre méthode est le recours aux modèles analytiques.

D'autres auteurs arrivent à la même conclusion quant à la précision des méthodes numériques par rapport aux méthodes analytiques en ayant recours à des techniques de simulation différentes. Monzó *et al.* (2013, 2014, 2018) utilisent également les éléments finis, quoique sans valider leur modèle à l'aide de données expérimentales. Eskilson (1987) et Hecht-Méndez *et al.* (2010) procèdent plutôt par différences finies. Rees et He (2013) recourent quant à eux aux volumes finis.

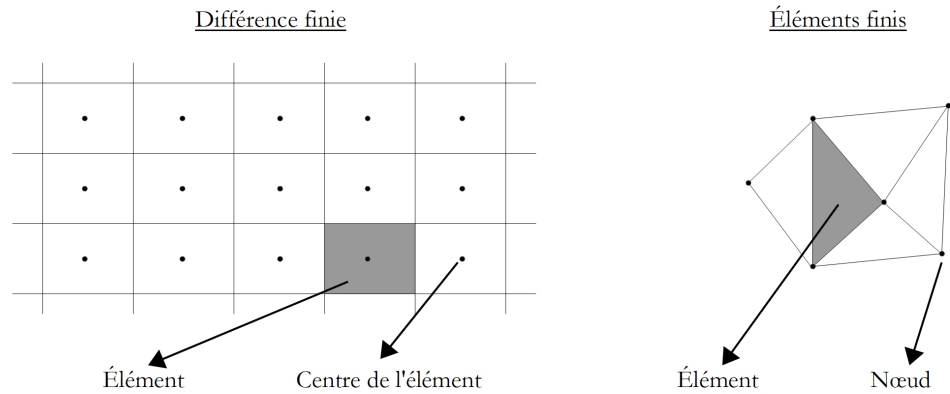


Figure 2.1 Comparaison entre le maillage de la différence finie (gauche) et des éléments finis (droite).

La principale différence entre la différence finie et les éléments finis réside dans le fait que la différence finie se cantonne à un maillage basé sur des formes plus régulières, comme on peut le voir à la Figure 2.1. Cela réduit la précision de ce motif de maillage par rapport aux deux autres, mais à l'avantage de limiter également les temps de calcul. Cependant, Chapuis (2009) indique que même la différence finie ne peut rivaliser en temps de calcul avec les méthodes analytiques pour le dimensionnement de systèmes dû aux grands nombres de modélisations qui sont requises.

C'est Ingersoll et Plass (1948) ainsi qu'Ingersoll *et al.* (1950, 1954) qui ont développé le premier modèle analytique permettant de calculer l'impact thermique d'un forage à distance. Leur modèle, celui de la source linéique infinie (SLI) représenté à la Figure 2.2, considère un empilement infini de sources de puissance q par unité de longueur (en W/m) pour calculer la perturbation thermique induite à une distance r après un temps t . Leur modèle est intéressant pour sa simplicité d'implémentation, mais commet des erreurs non négligeables pour de petites valeurs de t parce qu'il ne permet pas de prendre en compte la capacité thermique des puits géothermiques. Cela provient du fait que l'eau, le coulis, les tuyaux et le fluide caloporteur lui-même ont une inertie thermique qui doit être franchie lorsqu'on applique une charge en chauffage ou en climatisation dans un puits avant que cette perturbation ne soit ressentie par le sol environnant.

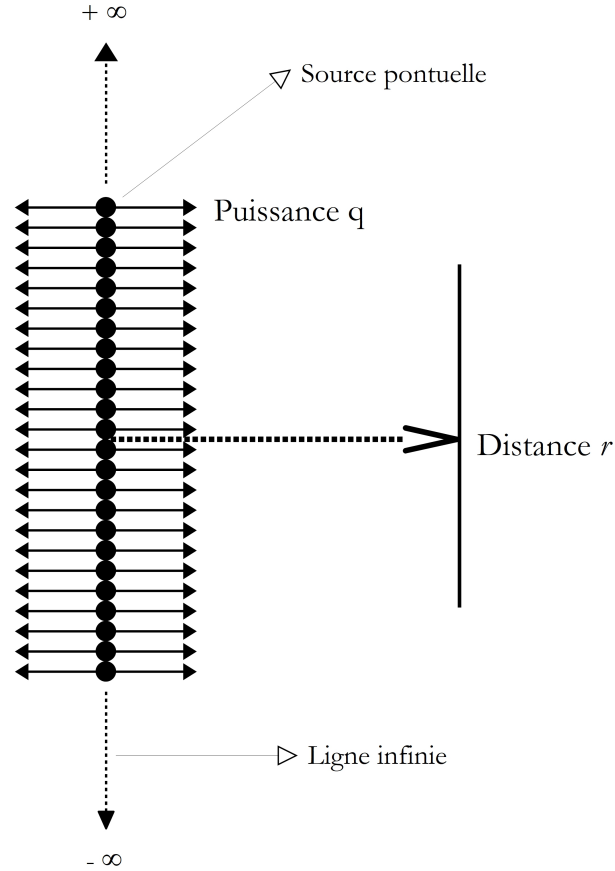


Figure 2.2 Représentation schématique de la source linéique infinie comme une superposition infinie de sources ponctuelles.

Utilisant la solution développée par Carslaw et Jaeger (1946), Ingersoll *et al.* (1950) propose un modèle ressemblant davantage à un puits, car pourvu d'un rayon non nul, appelé la source cylindrique infinie (SCI). Son évaluation fait intervenir l'intégration numérique de fonctions de Bessel de premier et de deuxième ordre et s'avère être plus coûteuse en temps de calcul. Plusieurs auteurs revisitent par la suite la SCI pour en accélérer le calcul en ayant recours à des transformations d'Euler (Cooper, 1976) ou en établissant des corrélations (Bernier, 2000, 2001; Bernier et Salim-Shirazi, 2007) à partir de valeurs tabulées par Ingersoll *et al.* (1954).

Marcotte *et al.* (2010) mettent toutefois en garde contre l'utilisation de méthodes analytiques d'extension infinie à cause des importantes imprécisions qu'elles peuvent causer lors de l'évaluation des perturbations thermiques dans les sols. En date d'aujourd'hui, il existe de nombreux modèles de sources linéiques finies (SLF), tous développés depuis le début des années 2000. Ces modèles font intervenir la profondeur d'enfouissement des têtes de forages D et permettent de cibler un point précis dans le sol pour y évaluer les changements de températures en fonction de ses coordonnées en r, z .

Zeng *et al.* (2002) sont parmi les premiers à proposer un modèle de SLF. Leur modèle suppose une profondeur d'enfouissement nulle. Lamarche et Beauchamp (2007) simplifient par la suite le modèle de la SLF pour qu'il ne nécessite la résolution que d'une seule intégrale, accélérant par le fait même sa résolution. Chapuis (2009) généralise ensuite la formulation de Lamarche et Beauchamp pour inclure les cas où $D \geq 0$. Finalement, Claesson et Javed (2011) présentent la formulation suivante (Eq. 2.1 à 2.3) grandement simplifiée de la SLF qui permet de connaître la perturbation thermique moyenne induite par un forage émetteur i auprès d'un forage récepteur j , sur sa hauteur, à une distance r et considérant la conductivité thermique du milieu géologique k_s et sa chaleur spécifique volumétrique C_s :

$$\Delta T(r_{j \rightarrow i}, t_k) = \frac{q}{4\pi k_s} \int_{\frac{1}{\sqrt{(4k_s t_k)/C_s}}}^{\infty} e^{-r_{j \rightarrow i}^2 s^2} \frac{Y(Hs, Ds)}{Hs^2} ds \quad (2.1)$$

où

$$Y(h, d) = 2 \operatorname{ierf}(h) + 2 \operatorname{ierf}(h + 2d) - \operatorname{ierf}(2h + 2d) - \operatorname{ierf}(2d) \quad (2.2)$$

et où

$$\operatorname{ierf}(X) = X \operatorname{erf}(X) - \frac{1}{\sqrt{\pi}} (1 - e^{-X^2}) \quad (2.3)$$

et où $r_{j \rightarrow i}$ est la distance entre le forage j et le forage i , t_k est le pas de temps considéré, q est le flux de chaleur émis par le forage j , k_s est la conductivité thermique du sol et C_s est la chaleur spécifique volumétrique du sol.

Lorsque le comportement thermique d'un échangeur de chaleur souterrain doit être déterminé à court terme, à l'échelle horaire ou plus finement, la capacité thermique des forages se doit d'être prise en compte et les modèles analytiques présentés jusqu'à maintenant n'offrent pas cette possibilité. Lamarche et Beauchamp (2007) proposent une nouvelle formulation analytique permettant de tenir compte de la réponse thermique à court terme. Pour connaître avec précision le comportement thermique des puits, un modèle TRC (pour *Thermal Resistance Capacity*) peut aussi être utilisé. De tels modèles sont utilisés entre autres par les auteurs suivants : Najib *et al.* (2019), Pasquier et Marcotte (2014) et Shulte *et al.* (2016).

Pour donner davantage de vraisemblance à leur modèle analytique, certains auteurs introduisent des modèles plus complexes offrant de nouvelles possibilités s'approchant davantage des méthodes numériques. Riviera *et al.* (2016) proposent une nouvelle formulation de la SLF qui permet de tenir compte de la chaleur provenant de la surface. Ghelin (2002) mentionne que pour un bon dimensionnement, l'effet de l'écoulement souterrain se doit d'être pris en compte. Cela dit, des études plus récentes menées par Samson *et al.* (2018) démontrent que

l'effet de l'eau semble n'avoir un impact significatif que pour des nombres de Péclet supérieurs à 10^{-2} . Tye-Gingras et Gosselin (2014) considèrent malgré tout cet effet et mettent au point une formulation analytique mixte qui utilise la SCI et une SLF se déplaçant dans l'espace (SCIM et SLFM pour SCI et SLF mouvante) pour émuler l'écoulement souterrain. Plus récemment, Hu (2017) ainsi que Erol et François (2017) développent un nouveau modèle analytique permettant de tenir compte des effets d'un écoulement d'eau irrégulier à travers une géologie stratifiée.

Pour simuler plus fidèlement les interactions entre forages, Cimmino et Bernier (2014) mettent au point une méthode semi-analytique morcelant chaque puits de l'échangeur de chaleur en plusieurs segments de SLF. Cimmino et Bernier (2015) valident expérimentalement ce modèle qui s'avère plus précis que l'utilisation d'une seule SLF pour représenter chaque puits. Avec la création de cette méthode semi-analytique, qui constitue en quelque sorte un compromis entre la précision des méthodes numériques et la facilité d'implémentation des méthodes analytiques, une partie de la vitesse de calcul de ces dernières est sacrifiée. À titre d'exemple, Cimmino et Bernier (2014) mentionnent pouvoir calculer l'ensemble des interactions thermiques entre les 100 forages d'un échangeur de chaleur souterrain durant une période de 8 ans au coût de 433 secondes de calculs environ. La précision de leur méthode est de 1.4% en comparant avec une méthode de référence numérique, et ce, une fois le régime permanent atteint. Cette précision est rendue possible en subdivisant chacun des forages de leur simulation en 12 segments distincts. Lamarche (2017) revisite la méthode semi-analytique pour tenter de mitiger ce problème de temps de calcul qui découle du grand nombre de segments devant être utilisés. Pour ce faire, il utilise des profils de flux de chaleur polynomiaux pour réduire leur nombre afin d'en accélérer le calcul sans compromettre la précision.

Cui *et al.* (2006) jettent les bases d'un modèle analytique d'extension finie qui permet de calculer les perturbations thermiques induites par un forage oblique qui part de la surface du sol. Marcotte et Pasquier (2009) généralisent leur modèle pour permettre de traiter des forages enfouis d'une distance quelconque sous la surface. Puis, Lamarche (2011) propose une nouvelle formulation analytique simplifiée plus facile à implémenter pour le modèle de Marcotte et Pasquier (2009). En 2016, Lazzarotto propose une série d'articles (2016b; 2016c; 2016a) où la méthodologie semi-analytique de Cimmino et Bernier (2014) est utilisée puis validée numériquement pour calculer les transferts de chaleur dans des échangeurs de chaleur souterrains contenant des forages inclinés. Plus récemment, Lamarche et Pasquier (2019) ont même reformulé les principaux modèles analytiques afin de pouvoir tenir compte d'une variation linéaire de la charge dans le temps.

En résumé, il existe de nombreuses méthodes numériques et analytiques permettant d'estimer,

avec différents degrés de détails et de vraisemblances, les transferts de chaleur dans le sol induits par un échangeur de chaleur souterrain. Il est admis que les méthodes numériques sont plus précises que les méthodes analytiques, mais également plus lentes, ce qui restreint leur déploiement pour des méthodes itératives de dimensionnement. Les méthodes analytiques, malgré leur rapidité d'implémentation et de résolution, font toutes intervenir l'intégration d'au moins une fonction complexe telle des fonctions de Bessel ou des fonctions d'erreur, ce qui nécessite malgré tout leur résolution par des méthodes numériques.

2.1.2 Couplage entre sol et bâtiment

Sans aborder les aspects thermodynamiques de la question, il est important de savoir comment fonctionne une thermopompe, cet appareil essentiel à la géothermie de basse température, pour comprendre comment le sol et un bâtiment peuvent interagir thermiquement entre-eux. La Fig. 2.3 illustre schématiquement les différentes composantes d'une thermopompe.

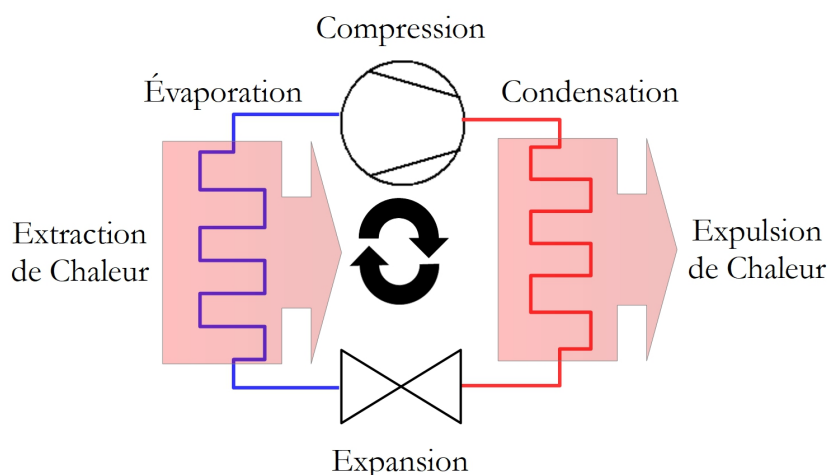


Figure 2.3 Représentation schématique d'une thermopompe en mode chauffage.

Les thermopompes réussissent à extraire de la chaleur d'un milieu froid pour la déplacer vers un milieu chaud en procédant de la façon suivante. Dans un circuit fermé, un réfrigérant est circulé entre deux échangeurs séparés par une valve d'expansion. Le fluide est détendu dans un des échangeurs, ce qui lui permet d'extraire de la chaleur du milieu avec lequel il est en contact en abaissant localement sa température et provoque son évaporation. Le fluide est ensuite comprimé et déplacé jusqu'à l'autre échangeur où il va se condenser et relâcher du même coup de la chaleur vers un milieu ayant des besoins de chauffage. Ce cycle peut être

inversé à tout moment, c'est ce qui permet aux thermopompes d'opérer à la fois en chauffage et en climatisation en fonction des besoins du bâtiment.

Bien que chaque thermopompe soit conçue pour être opérée sur une gamme différente de températures, leur capacité à transporter de la chaleur d'un milieu à l'autre varie grandement au fil de leur opération. Dans le cas de thermopompes géothermiques qui sont conçues pour travailler avec un fluide caloporteur arrivant d'un échangeur de chaleur souterrain, c'est la température de ce fluide même qui va dicter leur coefficient de performance (COP) et leur capacité (CAP). Le COP est un ratio entre énergie fournie par une thermopompe et énergie requise pour la faire fonctionner tandis que la CAP est la puissance qui peut être développée par la thermopompe. Ces paramètres n'ont pas à être déterminés empiriquement par les installateurs et les ingénieurs. Ils sont plutôt fournis directement par les fabricants sous forme de valeurs tabulées ou encore de schémas illustrant la variation du COP et de la CAP en fonction de la température d'entrée du fluide aux thermopompes, par exemple les deux graphiques de la Fig. 2.4.

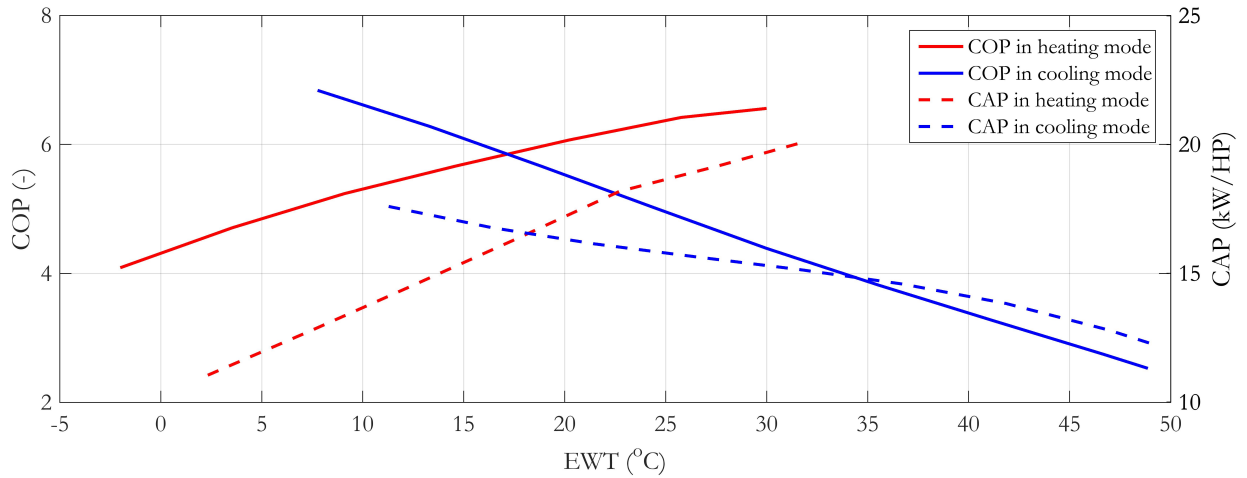


Figure 2.4 Évolution du COP et de la CAP d'une thermopompe, en chauffage et en climatisation, en fonction de la température d'entrée d'eau (EWT pour *Entering Water Temperature*).

Simuler simultanément l'opération de thermopompes et d'un échangeur de chaleur souterrain à l'intérieur d'un seul module de simulation cohérent est possible, et Rui *et al.* (2018) rapportent une bonne précision en procédant ainsi après comparaison avec des données expérimentales. Cependant, leur article ne comporte aucun temps de calcul et il est raisonnable de penser que le logiciel qu'ils ont développé n'est utilisable qu'à des fins de recherche dans son itération actuelle. Il est plus courant d'appliquer des taux d'injection ou d'extraction de chaleur à des forages, calculer leurs interactions thermiques résultantes dans le sol et dé-

duire la température d'entrée du fluide caloporteur avant de soumettre cette température aux thermopompes pour connaître leur COP et leur CAP. De façon générale, l'Eq. 2.4 indique comment cette température d'entrée du fluide aux thermopompes est obtenue :

$$T_f(t) = T_g + q(t) \cdot R_b + \Delta T(t) \quad (2.4)$$

Dans cette formule, T_f correspond à la température moyenne du fluide caloporteur entrant aux thermopompes pendant un temps t donné, T_g correspond à la température initiale du sol, q est la charge thermique appliquée par mètre de forage dans l'échangeur de chaleur, R_b est la résistance thermique des forages et $\Delta T(t)$ est l'incrément de température moyen aux parois des forages.

Le terme R_b est un terme qui rend compte de la difficulté qu'a la chaleur à être transférée par conduction et par convection thermique dans un puits entre le fluide et la paroi. Ce terme est estimé en combinant l'action isolante du coulis géothermique, des parois des tuyaux servant à la circulation du fluide caloporteur et de la convection qui s'effectue dans le fluide même. Parmi les méthodes les plus populaires permettant de calculer R_b , on compte celles d'Hellström (1991) basées sur la ligne source et sur une approximation de la méthode multipole de Claesson et Bennet (1987) et celle de Bennet *et al.* (1987). Cette dernière est d'ailleurs revisitée et améliorée par Claesson et Hellstrom (2011). Ce terme, R_b , n'a besoin d'être calculé qu'une seule fois lors de la simulation d'un système, car il ne change pas une fois le système mis en marche lorsque le débit de circulation ne change pas. Ainsi, il n'occupe donc qu'un très faible pourcentage du temps total de calcul.

Cimmino et Bernier (2014) classent les types de conditions frontières utilisées dans la littérature en trois catégories : un taux d'extraction de chaleur uniforme à travers l'échangeur de chaleur souterrain et des températures moyennes variables aux parois des forages comme dans les travaux de Zeng *et al.* (2002) et Claesson et Javed (2011), un taux d'extraction de chaleur différent pour chaque forage avec une température moyenne unique aux parois et une température uniforme imposée aux parois des forages comme dans la thèse d'Eskilson (1987). Malayappan et Spitler (2013) ainsi que Naldi et Zanchini (2018) mettent en garde contre les méthodes qui utilisent le premier type de condition frontière, révélant avec leurs exemples des erreurs pouvant aller jusqu'à 6.6 % et 8.7 % selon les champs de captage géothermiques qu'ils ont modélisés. Marcotte et Pasquier (2014), ayant comparé avec un modèle par éléments finis, démontrent que les erreurs commises sont plutôt de 3.5%. Les auteurs jugent cette différence acceptable et l'attribuent davantage à l'utilisation d'une formule analytique qu'aux conditions frontières employées.

Marcotte et Pasquier (2008) démontrent que le terme $\Delta T(t)$ peut également être obtenu en convoluant la fonction de transfert de l'échangeur de chaleur souterrain, établi par Eskilson (1987), avec un signal d'incrément de charges thermiques au sol. Tel qu'expliqué au Chapitre 1, une fonction de transfert, ou *g-function*, est un outil mathématique qui permet de lier des taux d'extraction de chaleur imposés aux parois des forages d'un champ de captage géothermique avec la température de ses parois. Chaque champ de captage géothermique possède sa propre fonction de transfert puisque celle-ci dépend à la fois de la géométrie du champ ainsi que des propriétés thermiques du sol environnant.

Une fois déterminée, la fonction de transfert d'un champ permet de traiter rapidement et de façon holistique l'ensemble des interactions thermiques entre les forages d'un champ de captage. Cela a comme conséquence de simplifier et d'accélérer les allers-retours entre déterminations de T_f et opération des thermopompes lors de la simulation d'un système.

La recherche de nouvelles façons d'obtenir rapidement et précisément la fonction de transfert d'un champ de captage géothermique est un domaine où il s'effectue beaucoup de recherche en géothermie de basse température. Eskilson (1987) est le premier à obtenir la fonction de transfert d'un champ en utilisant des simulations numériques basées sur la différence finie. Pour simplifier le problème à résoudre lors de l'élaboration de ses *g-functions*, il pose 5 hypothèses et opère certaines simplifications en jouant sur les conditions frontières de son modèle. La température de l'air est constante durant l'année, le sol est homogène, les têtes des forages sont isolées, les puits sont alimentés par le fluide caloporteur en parallèle et la température du sol est uniforme verticalement sur la hauteur des forages. Les conditions frontières qu'il utilise imposent une température uniforme aux parois de tous les forages du champ de captage géothermique. En discrétisant ses forages sur leur hauteur en cellules, Eskilson détermine numériquement les flux de chaleur qui résultent de l'uniformité qu'il impose aux températures sur les profils verticaux des forages. En normalisant la température aux parois des forages par les taux d'extractions de chaleur, Eskilson obtient ainsi une fonction de réponse adimensionnelle.

Pour déterminer plus rapidement la fonction de transfert d'un champ, de nombreux auteurs ont eu recours aux modèles analytiques présentés précédemment (SLI, SCI, et SLF) afin d'éviter les temps de calcul associés aux méthodes numériques. Quantifier les interactions thermiques entre les forages pour obtenir une fonction de transfert requiert que des formules analytiques soient utilisées à chaque pas de temps et pour chaque forage, sous l'influence d'une puissance de chauffage unitaire puis superposées temporellement et spatialement.

La façon de superposer temporellement les perturbations thermiques des différents puits d'un échangeur de chaleur souterrain exerce une grande influence sur le temps de calcul

d'une fonction de transfert. En effet, la formule classique de superposition temporelle :

$$\Delta T(r, q, t) = \sum_{i=1}^{n_t} \Delta T(r, q_i - q_{i-1}, t - t_{i-1}) \quad (2.5)$$

ne contient aucune récurrence. La somme qu'elle représente doit donc être recalculée au complet à chaque pas de temps. Yavuzturk et Spitler (1999) ainsi que Bernier *et al.* (2004) présentent des schémas d'agrégation des charges afin de réduire le nombre de calculs nécessaires dans cette opération. Moyennant une perte de précision, les charges horaires les plus lointaines du pas de temps actuel sont fusionnées en une charge moyenne qui couvre un intervalle de temps plus grand.

En parallèle, Marcotte et Pasquier (2008) proposent d'exprimer la superposition temporelle comme un produit de convolution dans le domaine spectral entre un modèle analytique et un signal d'incrément de charges thermiques. Grâce aux gains de temps réalisés, cela rend inutile l'agrégation des charges thermiques. Remarquant que les modèles SLI, SCI et SLF sont des fonctions au comportement lisse et régulier, Marcotte et Pasquier proposent également de réduire le nombre d'évaluations à effectuer et d'interpoler à l'aide de splines cubiques pour obtenir une évolution horaire des températures aux parois des forages. Les auteurs rapportent une accélération des temps de calcul par cette procédure de plusieurs ordres de grandeur, une conclusion corroborée par Cimmino *et al.* (2012).

Pasquier et Marcotte (2013), en utilisant l'approche par produits de convolution qu'ils ont développée, jettent les bases d'une méthode itérative pour trouver la fonction de transfert d'un champ de captage géothermique. Cette approche modifie itérativement et simultanément les flux de chaleur de tous les puits à tous les pas de temps pour tenter de reproduire un signal de température aux parois des forages. Cette façon de procéder est polyvalente et précise, mais devient exponentiellement plus lente pour un grand nombre de forages ou pour de longues simulations.

Lamarche (2009), Cimmino *et al.* (2013) ainsi que Lazzarotto (2014), afin de simplifier le calcul de la contribution historique des perturbations thermiques à chaque pas de temps, ont l'idée de traiter séquentiellement les transferts de chaleur ainsi que les températures moyennes aux puits en utilisant des transformées de Laplace. Lazzarotto (2014) mentionne, mais sans fournir les équations pertinentes, que son algorithme permet de calculer la fonction de transfert d'un champ où les puits ne sont pas nécessairement alimentés en parallèle. Cette possibilité constitue une première pour déterminer les *g-functions* de façon analytique d'un champ où les puits ne sont pas alimentés en parallèle. Cette hypothèse sur l'approvisionnement des puits en fluide est héritée, rappelons-le, des travaux pionniers d'Eskilson

(1987).

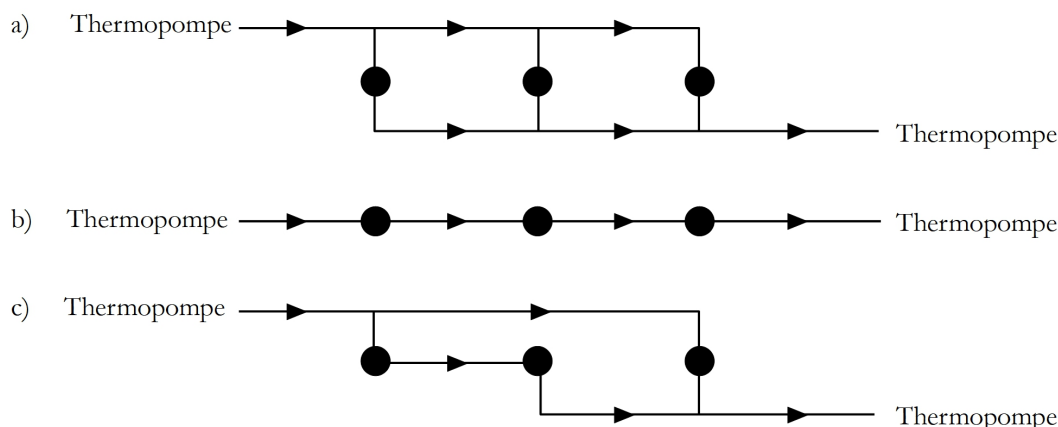


Figure 2.5 Configurations possibles de distribution de fluide caloporteur de la sortie d'une thermopompe à son retour : a) parallèle, b) série ou c) mixte.

Marcotte et Pasquier (2014), reprenant l'idée d'un champ où les puits peuvent être alimentés par les différentes configurations présentées à la Figure 2.5, développent une approche par transformées de Fourier basée sur une méthode qu'ils ont eux-mêmes publiée plus tôt (Marcotte et Pasquier, 2008). Ils proposent ainsi un nouveau type de fonction de transfert basée sur T_f au lieu de la température aux parois des forages. Leurs résultats sont confrontés à des simulations numériques par éléments finis 3D et concordent avec une bonne précision considérant qu'ils sont obtenus à l'aide d'une approche analytique. Leur nouveau type de fonction de transfert leur permet de découvrir que, à réseaux d'alimentation identiques, l'ordre dans lequel les forages sont parcourus par le fluide caloporteur n'a pas d'incidence sur la façon dont le sol va se chauffer aux alentours sur le long terme. Cimmino (2019) développe également une approche pour traiter plusieurs connexions possibles entre forages en intégrant le concept de modèle semi-analytique qu'il a développé pour la SLF (Cimmino et Bernier, 2014). Lamarche (2017), en combinant les approches de Marcotte et Pasquier (2014) et celle de Lazzarotto (2014), propose une approche permettant de traiter des champs connectés de n'importe quelle façon et qui peuvent être parcourus par plusieurs réseaux indépendants de fluide caloporteur.

Lorsque l'on calcule la fonction de transfert d'un champ contenant un grand nombre de puits et sur une longue période, le modèle analytique utilisé devra être évalué des dizaines ou des centaines de milliers de fois. Nécessairement, certains éléments seront similaires entre toutes ces évaluations. Cimmino (2018) met à profit cette redondance pour accélérer l'évaluation des fonctions de transfert dans son ensemble. Il rapporte un temps de calcul de 27 minutes

pour un champ de 512 forages distribués aléatoirement dans un champ. À titre comparatif, Chen *et al.* (2019) rapportent des temps de calcul records de 1 min. 30 pour l'établissement de la fonction de transfert d'un champ de 40 puits sur 30 ans en utilisant une version dérivée de la SLF par rapport au temps de Fourier en conjonction avec des transformées de Fourier.

En résumé, le couplage entre sol et bâtiment est un processus complexe à modéliser et l'utilisation de fonctions de transfert pour représenter le champ de captage géothermique aide à simplifier ce processus. Les dernières années ont vu l'abandon des méthodes numériques pour leur détermination, suivi d'une complexification des méthodes analytiques servant à les déterminer. Malgré tout, leur détermination reste un processus itératif et peu rapide où des modèles analytiques sont utilisés en série avant d'être superposés spatialement et temporellement.

2.2 Dimensionnement de systèmes

Effectuer la conception d'un système géothermique hybride demande d'ajuster un ensemble de variables à travers leur plage de validité respective dans le but de maximiser ou minimiser une caractéristique du système. L'évolution de cette caractéristique à travers le processus d'optimisation est suivie par le calcul d'une fonction objectif. Puisque la perspective de réaliser des économies est une des raisons les plus importantes poussant les gens à investir dans l'installation de la géothermie, il est courant de voir un indicateur économique comme la valeur actuelle nette (VAN, présentée à l'Eq. 2.6) d'un projet servir de fonction objectif. Utiliser la VAN est un choix judicieux, car selon Adnot *et al.* (2012), seule une analyse rigoureuse prenant en compte différents indicateurs financiers permet d'établir la rentabilité d'un système énergétique. Erdeweghe *et al.* (2018) corroborent cette affirmation en démontrant, dans l'étude de cas d'une centrale géothermique de basse température située en Europe, que le contexte financier à lui seul peut faire la différence entre des gains de 11.63 millions d'euros et des pertes de 9.91 millions d'euros sur un horizon de 30 ans.

$$VAN = \sum_{t=0}^T \frac{CF_t}{(1+i)^t} \quad (2.6)$$

où t désigne la période en cours, généralement une année, T est l'horizon de temps considéré, CF_t est le flux de trésorerie à la période t et i est le taux d'actualisation considéré.

Bien que cela ne garantisse pas de converger vers une solution optimale unique (Ehrgott, 2005), il est possible de suivre simultanément l'évolution de plusieurs fonctions objectif. Cela est utile lorsque l'on cherche à balancer des enjeux économiques avec des enjeux environne-

mentaux ou technologiques. Sayyadi *et al.* (2009), Sanaye et Hajabdollahi (2010), Sayyadi et Nejatollahi (2011), Ghanei *et al.* (2014), Wen *et al.* (2016) de même que Rao et Saroj (2018) travaillent tous sur des problèmes d’optimisation utilisant un indicateur financier, généralement la VAN, avec un indicateur thermodynamique qui sert à caractériser l’efficacité des transferts dans leurs échangeurs de chaleur. Pu *et al.* (2017) préconisent une approche de minimisation de l’entropie et d’un terme intégrateur, emprunté à la mécanique des fluides, faisant intervenir les pertes de charge dans les tuyaux du système. Sivasakthivel *et al.* (2014) utilisent directement la résistance thermique des forages, leur longueur ainsi que le COP des thermopompes comme fonctions objectif. Zeng *et al.* (2015) utilisent également trois fonctions objectifs pour maximiser les économies annuelles d’énergie et d’argent et minimiser les émissions de gaz à effet de serre. Hadidi (2015) cherchent plutôt à éviter les baisses de pression dans leur système, réduire les coûts annuels totaux ainsi que la surface totale de leur échangeur de chaleur.

Il existe d’autres approches plus simples et plus rapides pour aider au dimensionnement d’échangeurs de chaleur qui ne requièrent pas l’utilisation de fonctions objectif. L’ASHRAE (ASHRAE Handbook Ch.34, 2011; Kavanaugh et Rafferty, 1997) propose une équation de dimensionnement standard qui requière quelques paramètres (température initiale du sol, certaines demandes thermiques du bâtiment et les résistances thermiques effectives associées). Ces demandes sont typiquement les pires consommations moyennes sur des horizons de 10 ans, un mois et 6 heures.) pour fournir une longueur totale de mètres de forage à installer. Cette équation a par la suite été corrigée par Bernier *et al.* (2008) avec l’ajout d’un paramètre représentant une pénalité de température et à nouveau par Bernier (2006) et par Philippe *et al.* (2010) pour permettre le dimensionnement d’un système hybride directement. Cullin *et al.* (2015) mettent cependant en garde contre la précision d’une telle procédure, car elles pourraient mener à l’occasion à d’importants sur et sous-dimensionnements. Cela n’empêche pas Retkowski et Thöming (2014) d’utiliser ces équations de dimensionnement pour valider la précision de leur propre méthodologie d’optimisation. Au-delà de leur précision, le recours à une formule de dimensionnement s’accompagne d’un problème sous-jacent important. Bien qu’elle puisse fournir une longueur totale optimale de forages pour répondre à la demande d’un bâtiment, il lui est impossible de fournir l’évolution horaire de T_f . Ce paramètre est requis pour simuler l’opération heure par heure des thermopompes, comme le montrent Shonder *et al.* (1999) dans leur étude comparative de six méthodes de dimensionnement différentes. Dans un contexte où le coût d’exploitation des thermopompes intervient pour établir la rentabilité d’un projet, à l’aide de la VAN par exemple, il est nécessaire d’obtenir ces températures horaires.

Des logiciels de simulation existent qui permettent d’obtenir ces températures, entre autres

valeurs d'importance, pour aider au dimensionnement d'un système hybride. TRNSYS (2000) est un logiciel de simulation thermique dynamique appliqué au bâtiment qui intègre les variables d'emplacement, de matériaux, d'architecture et de concepts énergétiques choisis, y compris les systèmes solaires et géothermiques. Sa base de données peut être enrichie de nouveaux composants, systèmes et concepts énergétiques et est partagée entre utilisateurs. Cui *et al.* (2015) y ont partagé un module permettant la simulation d'un champ de captage géothermique où les forages peuvent être connectés en parallèle ou en série. De Rosa *et al.* (2015) et Ruiz-Calvo *et al.* (2015) y ont travaillé sur un module permettant d'obtenir la réponse à court terme d'un forage plus rapidement qu'avec l'implémentation initiale du logiciel. Nombreux sont les auteurs, comme Courchesne-Tardif (2011) et Pu *et al.* (2017) qui utilisent TRNSYS pour élaborer de nouvelles méthodologies d'optimisation ou valider leurs derniers résultats. Hackel *et al.* (2009) utilisent même TRNSYS pour conduire l'étude paramétrique de quatre types de bâtiments sur six sites différents afin d'en tirer des équations de premiers degrés devant servir au dimensionnement de champ de captage géothermique pour des contextes dominés en climatisation.

Pour éviter de faire manuellement des allers-retours entre module de simulation, équation de dimensionnement et calcul d'une fonction objectif, il est possible de confier les rênes du processus d'optimisation à un algorithme. De façon générale, on remarque que les méthodes permettant une optimisation d'un plus grand nombre de variables ont tendance à recourir à des heuristiques plutôt qu'à des méthodes d'optimisation mathématique basées sur le gradient. Das et Suganthan (2011) et Ayala *et al.* (2016) corroborent cette observation et notent que les algorithmes évolutionnaires sont plus adaptés pour effectuer des recherches d'optimum dans des espaces de haute dimensionnalité. Ce genre d'algorithme, comme les algorithmes génétiques et génétiques quantiques utilisés par Li *et al.* (2019) par exemple, ne garantit pas l'atteinte d'un optimum global d'autant plus qu'il nécessite beaucoup de réglages fins de la part de l'utilisateur pour assurer leur bon fonctionnement. Pour réduire la dimensionnalité et la complexité des problèmes d'optimisation à résoudre, certains auteurs tels que Holl et Pelz (2016), Mizutani *et al.* (2003) et Onishi *et al.* (2013) décident de fragmenter le problème en sous-problèmes ayant chacun leur propre fonction objectif à résoudre. Selon les différents auteurs, les sous-problèmes sont traités en série ou en parallèle.

2.2.1 Algorithmes d'optimisation

La méthode du gradient conjugué inventée par Hestenes et Stiefel (1952) est une méthode itérative d'optimisation. Elle se sert du gradient de la fonction objectif pour déterminer la direction et le pas de sa prochaine itération. Retkowski et Thöming (2014) l'utilisent en

supposant que les interactions thermiques entre les puits sont inexistantes. Robert et Gosselin (2014) l’emploient également en posant l’hypothèse que toutes les variables sont continues dans leur problème.

Hénault *et al.* (2016) développent un algorithme d’optimisation en nombres continus utilisant conjointement les gradients conjugués contraints ainsi que le balayage. Cet article a constitué le point de départ et la source d’inspiration de la thèse actuelle. Leur algorithme optimise 6 variables (nombre de forages, espacement entre les lignes et les colonnes de forages dans le champ, température de consigne minimale et maximale des thermopompes et nombre de thermopompes) et tient compte de nombreux paramètres financiers pour converger vers le système hybride qui maximise la VAN. De plus, il prend en compte les stratégies de contrôle prévues durant l’opération d’un bâtiment, ce qui permet une utilisation maximale du système proposé (Jinggang *et al.*, 2009). Leur algorithme d’optimisation est utilisé conjointement avec un module qui permet la simulation horaire d’un système géothermique hybride (Pasquier *et al.*, 2013). Ce module permet d’exprimer les besoins énergétiques horaires des systèmes auxiliaires.

La méthode du gradient conjugué n’est qu’une des méthodes d’optimisation itératives qui utilisent le gradient ou la hessienne d’une fonction objectif pour converger. Parmi les plus utilisées en géothermie, on retrouve les méthodes de Newton et de quasi-Newton (Broyden, 1967), la descente la plus rapide (Curry, 1944) et la région de confiance intérieure (*interior trust region*) pour des problèmes de minimisation non-linéaires (Coleman et Li, 1996). Toutes ces méthodes sont conçues pour des problèmes où les variables à optimiser sont continues, mais peuvent être adaptées pour des variables entières en utilisant des techniques de programmation pour nombres entiers appliquées à des problèmes non linéaires que Belotti *et al.* (2013) ont couvertes en détail dans une excellente revue de littérature.

Une des plus simples méthodes qui existe pour contourner la contrainte sur les variables continues est le balayage. Cette contrainte touche le nombre de thermopompes et de forages, deux variables entières qui sont souvent optimisées par les auteurs. Le balayage est une façon de procéder qui requiert un nombre important de calculs de la fonction objectif, ce nombre augmentant exponentiellement selon le nombre de variables balayées. Cela n’empêche pas Alavy *et al.* (2013) de même que Bayer *et al.* (2014) d’utiliser exclusivement le balayage pour une optimisation à deux variables ainsi que Ni *et al.* (2011) pour une optimisation à un seul paramètre.

Pour remédier à cette contrainte sur les variables continues, d’autres approches ont été essayées. Hackel *et al.* (2008) maximisent la VAN d’un système hybride reposant sur le modèle DST d’Hellström (1989) en utilisant une métaheuristique par essaim particulière (PSO pour

particle swarm optimization). Huang *et al.* (2015) fonctionnent par optimum local de Pareto pour résoudre leur problème contenant deux fonctions objectifs. Retkowski et Thöming (2014) implémentent une méthode de séparation et évaluation (*branch and bound*) pour contraindre le nombre de puits de leur modèle à évoluer en nombres entiers. Sivasakthivel *et al.* (2014) utilisent la méthode de Tagushi pour introduire des variables entières dans leur modèle contenant huit variables et trois fonctions objectifs. Sanaye et Hajabdllahi (2010) et Yousefi *et al.* (2013) utilisent eux aussi des métaheuristiques, respectivement l'algorithme NSGA-II (*nondominated sorting genetic algorithm*) et la recherche harmonique, dans lesquelles certaines de leurs variables sont entières. Mizutani *et al.* (2003) utilise un logiciel appelé DICOPT++, développé par Grossman *et al.* (2003), qui propose plusieurs méthodes d'optimisation mathématiques classiques reposant sur les nombres entiers. Plus récemment, Guo et Meggers (2017) utilisent des méthodes de programmation en nombres entiers mixte hybride non linéaire (MINLP), sans toutefois faire référence à un algorithme en particulier, pour dimensionner un système de chauffage urbain à partir d'un ancien puits de pétrole en Pennsylvanie.

2.2.2 Variables optimisées

Dans un système géothermique hybride, tous les paramètres n'ont pas le même impact sur la ou les fonctions objectifs considérées. Il est facile, par exemple, de concevoir intuitivement que le nombre de puits n'a pas la même incidence sur le coût et la rentabilité d'un système que le diamètre des tuyaux de circulation du fluide caloporteur. Pour cette raison, différents auteurs choisissent de travailler avec différentes variables d'optimisation. Leurs raisons sont motivées par leur expérience personnelle, l'algorithme d'optimisation qu'ils utilisent, les fonctions objectifs qu'ils emploient ou encore sur les résultats d'analyses de sensibilité qu'ils ont effectuées pour évaluer l'impact de leurs variables sur l'optimum.

Alavy *et al.* (2013) et Bayer *et al.* (2014), limité par leur algorithme d'optimisation par balayage, ne travaillent qu'avec deux variables. Zhang *et al.* (2016) avec la méthode de Hooke-Jeeves, tout comme Caputo *et al.* (2006) avec un algorithme génétique pour le design d'échangeurs de chaleur dans des applications industrielles, utilisent trois variables d'optimisation. Robert et Gosselin (2014) avec leur approche par gradient conjugué et Pu *et al.* (2017) avec un mélange de krigeage et d'algorithmes génétiques ont recours à quatre variables d'optimisation. Huang *et al.* (2015) avec leur méthode par optimum local de Pareto tiennent compte de 5 variables différentes. Hénault *et al.* (2016) utilisent les 6 variables mentionnées précédemment avec une combinaison de gradients conjugués contraints et de balayage. Sanaye et Hajabdllahi (2010) avec la méthode NSGA-II pour le design d'échangeurs de chaleur

à calandre et Turgut (2016) avec une version modifiée de PSO pour le design d'échangeurs à plaques et ailettes utilisent 7 variables différentes dans leur modèle. Sayyadi *et al.* (2009), à l'aide d'un algorithme génétique, et Sivasakthivel *et al.* (2014) et Pandey *et al.* (2017), tous deux en utilisant la méthode Tagushi, ainsi que Mizutani *et al.* (2003), en utilisant l'un des algorithmes offerts dans le logiciel DICOPT++, procèdent à des optimisations selon 8 variables d'intérêts. Yousefi *et al.* (2015), en utilisant un nouveau type de PSO intelligent capable d'apprentissages, intègrent 9 variables différentes à leurs optimisations. Pour finir, en implémentant une méthode multiobjective à algorithme évolutionnaire, Sayyadi et Nejatolahi (2011) incorporent 12 variables à leur algorithme d'optimisation. Les travaux et les variables utilisées sont résumés dans le tableau suivant (Tableau 2.1).

Nombre (-)	Article	Variables
1	Ni <i>et al.</i> (2011)	η_H
2	Alavy <i>et al.</i> (2013)	η_C, η_H
	Bayer <i>et al.</i> (2014)	Q_H, Q_C
3	Zhang <i>et al.</i> (2016)	n_b, H, S
	Caputo <i>et al.</i> (2006)	B, d_{out}, C_{in}
4	Robert et Gosselin (2014)	n_b, H, S, p
	Pu <i>et al.</i> (2017)	d_p^*, l^*, T_{in}^*, Re
5	Huang <i>et al.</i> (2015)	n_b, H, S, r_b, r_{out}
6	Hénault <i>et al.</i> (2016)	$EWT_{min}, EWT_{max}, S_x, S_y, n_b, n_{HP}$
7	Sanaye et Hajabdollahi (2010)	$t_a, d_{out}, t_p, H, n_b, B, B_{cut}$
	Turgut (2016)	$T_H, T_C, F_H, F_F, F_T, L_L, n_F$
8	Sayyadi <i>et al.</i> (2009)	$\Delta T_{cond}, \Delta T_{evap}, \Delta T_{GHEX}, \Delta T_{br}, \Delta T_{air},$ $\Delta T_{sup}, \Delta T_{sub}, r_{out}$
	Sivasakthivel <i>et al.</i> (2014)	$r_{out}, r_b, Q, k_g, EWT, p_s, k_p, \dot{m}$
	Pandey <i>et al.</i> (2017)	$C_f, k_f, \mu_f, \rho_f, k_p, r_p, p_s, \dot{m}$
	Mizutani <i>et al.</i> (2003)	$n_F, T_1, d_{in}, d_{out}, T_2, n_B, T_3, T_4$
9	Yousefi <i>et al.</i> (2015)	L_a, L_b, F_H, F_F et $F_T, l_F, N_a, mb_1, mb_2$
12	Sayyadi et Nejatolahi (2011)	$T_{cond}, T_{evap}, LWT, EWT, LWT_{CT},$ $EWT_{CT}, LD_{cond,I}, LD_{cond,II}, LD_{evap}, \dot{m},$ $\Delta T_{sub}, \Delta T_{sup}$

Tableau 2.1 Nombres et natures des variables utilisées par différents auteurs dans le cadre de dimensionnement par optimisation.

Dans le Tableau 2.1, η_C et η_H sont des pourcentages des charges totales en climatisation et en chauffage, Q_H et Q_C sont les charges en chauffage et en climatisation, n_b est le nombre de forage, H est la longueur des forages, S est l'espacement entre forages, B est l'espacement entre tubes, d_{out} est le diamètre extérieur du tube, C_{in} est le diamètre intérieur de la coque, d_p^* , l^* et T_{in}^* sont des paramètres adimensionnels décrivant la géométrie d'un puits et la

température d'entrée d'eau aux thermopompes, Re est le nombre de Reynolds, r_b est le rayon du puits, r_{out} est le rayon extérieur des tuyaux, EWT est la température d'entrée d'eau aux thermopompes, n_{HP} est le nombre de thermopompes, t_a est l'arrangement des tubes, t_p est la longueur des tubes, B_{cut} est la ratio de coupe des cloisonnements, T_H et T_C sont les températures chaudes et froides des ailettes, F_H , F_F et F_T sont les dimensions et fréquence des ailettes, L_L est la longueur des lances, n_F est le nombre d'ailettes, ΔT_{cond} est la différence entre la température d'entrée de la saumure et celle du réfrigérant dans le condenseur, ΔT_{evap} est la différence entre la température de l'air de sortie et celle du réfrigérant surchauffé dans l'évaporateur, ΔT_{GHEX} est la différence entre la température de sortie de la saumure et celle du sol dans l'échangeur de chaleur souterrain, ΔT_{br} est la différence de température entrée / sortie de la saumure dans l'échangeur de chaleur souterrain, ΔT_{air} est la différence de température entrée / sortie de l'air dans l'évaporateur, ΔT_{sup} est l'ampleur du surchauffage dans l'évaporateur, ΔT_{sub} est l'ampleur du refroidissement dans le condenseur, k_g est la conductivité thermique du coulis, p_s est l'espacement entre les tuyaux dans le forage, k_p est la conductivité thermique des tuyaux et $\dot{m}$ est le débit massique du fluide caloporteur, C_f , k_f , μ_f et ρ_f sont la chaleur massique, la conductivité thermique, la viscosité et la masse volumique du fluide caloporteur, r_p est le rayon des tuyaux, T_1 est le nombre de passes de tuyaux, T_2 est le patron d'arrangement des tubes, n_B est le nombre de cloisons, T_3 est le type de tête de tuyau, T_4 renseigne sur l'allocation du fluide, L_a et L_b sont les longueurs des circuits chaud et froid, l_F est le décalage entre deux rangées d'ailettes, N_a est le nombre de couches chaudes, mb_1 et mb_2 sont les écoulements massiques sur les circuits froid et chaud, T_{cond} et T_{evap} sont les températures de condensation et d'évaporation à saturation, LWT est la température de l'eau à la sortie des thermopompes, LWT_{CT} et EWT_{CT} sont les températures de l'eau à la sortie et à l'entrée de la tour de refroidissement, $LD_{cond,I}$ et $LD_{cond,II}$ sont deux ratios entre la longueur et le diamètre de l'enveloppe du condenseur et LD_{evap} est ratio entre la longueur et le diamètre de l'enveloppe de l'évaporateur.

En résumé, les auteurs ayant publié sur le dimensionnement d'échangeurs de chaleur souterrains et hors-sol dans les dernières décennies ont développé et utilisé un large éventail de méthodes. Après un tour d'horizon des méthodes les plus utilisées, on distingue trois principales catégories dans lesquelles elles peuvent être regroupées :

1. Les équations de dimensionnement. Elles sont rapides à utiliser, nécessitent peu de paramètres, mais ne permettent pas d'estimer la rentabilité d'un système.
2. Les méthodes mathématiques basées sur les dérivées premières et secondes de la fonction objectif. Elles convergent en un nombre fini, mais potentiellement élevé d'itérations et garantissent l'optimalité de la solution pour des problèmes à une fonction

objectif.

3. Les méthodes basées sur des heuristiques et des méta-heuristiques. Elles sont intéressantes pour les problèmes comptant de nombreuses variables d'optimisation, mais ne peuvent garantir l'attente de l'optimum global du domaine à l'étude. De plus, elles requièrent plusieurs réglages fins de la part de L'utilisateur au moment de leur implémentation.

À part les équations de dimensionnement, toutes les méthodes peuvent utiliser un nombre variable de paramètres d'optimisation conjointement à une ou plusieurs fonctions objectifs. Lorsque des variables entières se trouvent à être incluses au processus d'optimisation, comme un nombre de forages ou de thermopompes par exemple, les auteurs ont le choix d'ignorer ces contraintes et de les traiter comme des variables continues ou encore d'incorporer des méthodes d'optimisation en nombres entiers à leurs algorithmes.

2.3 Risques et incertitudes

La façon dont incertitude et risque sont mesurés et gérés varie grandement d'un domaine à l'autre. En géothermie de basse température, cette problématique est traitée presque uniquement par des analyses de sensibilité et jamais de mesures de mitigation ne découlent de ces analyses. Une analyse de sensibilité permet d'évaluer comment une variation dans la sortie d'un modèle peut être attribuée à une variation des intrants autour de l'optimum. Ce type d'analyse est effectué *a posteriori* après qu'un design ait été proposé.

Cette façon de procéder n'est pas la norme en finance par exemple où la gestion du risque fait partie intégrante de la façon dont les portefeuilles d'investissements sont constitués. De ce fait, pour étoffer cette section de la revue de littérature, plusieurs articles à l'extérieur de la géothermie sont explorés pour donner une vue d'ensemble de ce qui est envisageable dans la gestion du risque.

2.3.1 Quantification

Il existe trois types d'analyse de sensibilité utilisé dans la modélisation de phénomènes environnementaux qui sont bien documentés dans un état de l'art publié par Pianosi *et al.* (2016). Le *ranking*, ou priorisation des paramètres, vise à quantifier puis à classer les différents paramètres que l'on introduit dans un modèle en fonction de leur impact sur les valeurs de sortie et potentiellement sur une fonction objectif qui en dépend. Le *screening*, ou dépistage, vise à identifier les paramètres d'entrée du modèle dont la contribution est négligeable, généralement pour pouvoir les exclure d'un processus d'optimisation par la suite dans le

but de réduire la dimensionnalité du problème. Le *mapping*, ou cartographie, vise à explorer et caractériser les zones de l'espace dans lequel les paramètres d'entrée peuvent évoluer et qui produisent des résultats remarquables parmi les paramètres de sortie, par exemple des résultats extrêmes ou incohérents.

Il existe de nombreuses méthodes qui s'inscrivent dans ces trois types d'analyses de sensibilité. Certaines servent à établir une analyse de sensibilité localement, autour d'un optimum fourni par un algorithme d'optimisation par exemple, tandis que d'autres donnent une idée plus générale de la sensibilité d'un modèle face à ses paramètres d'entrée (Pianosi *et al.*, 2016). Certaines de ces méthodes, de type *one-at-a-time* (OAT), perturbent itérativement un paramètre d'entrée à la fois tandis que d'autres, de type *all-at-a-time* (AAT), en modifient plusieurs simultanément.

Dans le cadre de cette thèse, comme l'objectif est d'intégrer l'impact de l'incertitude des charges horaires thermiques d'un bâtiment, des coûts de construction, de l'évolution du coût de l'énergie ainsi que de la conductivité thermique des sols simultanément durant le dimensionnement d'un système géothermique hybride (voir les paramètres qui apparaissent à la Fig. 1.1), seules les analyses de type AAT sont considérées. En effet, comme mentionné dans Charitopoulos et Dua (2017), l'optimisation de systèmes énergétiques présente souvent des variables qui, sans être directement corrélées, exercent une influence les unes sur les autres à cause de la complexité des modèles utilisés. Il est donc logique que, à cause de la complexité de modéliser les transferts de chaleur entre bâtiment et sol environnant, que les analyses reposent sur des méthodes de type AAT. Or, c'est tout le contraire qui est observé dans la littérature.

Parmi les nombreux auteurs publiant sur l'optimisation présentés à la section précédente, la majorité de ceux qui s'adonnent à proposer des analyses de sensibilité de leurs résultats le fait par perturbation itérative des variables d'optimisation pour en quantifier l'impact. Un exemple serait Hadidi (2015) qui, à cause de la complexité de son modèle d'optimisation à trois fonctions objectif, recourt à une méthode de perturbation itérative (OAT) des variables pour constater visuellement leur impact sur les fronts de Pareto qu'il a calculés. Zurmühl *et al.* (2019) procèdent également à une analyse de sensibilité OAT pour les trois paramètres incertains de leur design : la conductivité thermique du sol, le prix de l'électricité et les coûts des forages. Ils concluent que, à variation égale, les coûts de forages entraînaient des changements de 1.35 à 8 fois plus importantes sur l'optimum de la valeur totale du projet que les deux autres paramètres. Bien que leur analyse soit rigoureuse selon les standards de la méthode qu'ils ont employée, leurs résultats ne permettent pas de prédire l'impact sur l'optimum si plusieurs paramètres étaient modifiés simultanément.

Les rares exceptions à cette façon de procéder parmi les auteurs présentés précédemment sont Huang *et al.* (2015), qui intègrent leur analyse de sensibilité directement dans leur algorithme d’optimisation pour identifier les variables ayant un impact fort sur leur design avant de débiter, et Wen *et al.* (2016) qui recourent au krigeage simple pour produire une analyse de sensibilité de leurs paramètres de design. Les travaux de Huang *et al.* (2015), bien que simplifiant le design en réduisant le nombre de paramètres d’optimisation, ne permettent pas pour autant de quantifier l’impact des paramètres d’entrée sur l’optimum une fois l’optimisation terminée. Quant aux travaux de Wen *et al.* (2016), utiliser le krigeage pour estimer la sensibilité d’un modèle constitue un recours à un modèle dit de substitut, ou *surrogate model* en anglais. L’emploi d’un tel modèle est toujours sujet lui-même à un certain degré d’incertitude difficilement quantifiable et est sensible aux paramètres que l’utilisateur lui fournit et aux choix que fait le concepteur lors de son élaboration selon Pianosi *et al.* (2016).

Fernandez *et al.* (2017), sans toutefois avoir pour finalité la quantification de l’incertitude dans un processus d’optimisation, utilisent une combinaison de plusieurs méthodes (Morris OAT, Sobol-Jansen et la technique BACCO, pour *Bayesian analysis of computer code outputs*) afin de circonscrire l’incertitude de leur modèle et établir les variables ayant le plus d’impact sur ce dernier. Leur but dans ce processus était de démontrer qu’une analyse de sensibilité permet de calibrer un modèle de simulation de puits verticaux sur TRNSYS plus finement, plus rapidement et en évitant les biais du modélisateur.

Cannavó (2012), plus loin de la géothermie de basse température, utilise une approche basée sur la variance des paramètres en combinant la méthode FAST (pour *Fourier Amplitude Sensitivity Test*) (Cuckier *et al.*, 1973) et l’approche de Sobol (1993) afin de quantifier la capacité d’ajustement d’un modèle de source volcanique analytique à des données de déformations synthétiques. Les méthodes basées sur la variance, comme la méthode FAST, obtiennent une approximation de la sensibilité d’un ou plusieurs paramètres en propageant leur variance à travers un modèle pour voir quelle sera la variance sur l’optimum calculé. Cette méthode a l’avantage d’être facile à implémenter. En revanche, elle nécessite plusieurs milliers d’évaluations du modèle par paramètre considéré pour correctement propager leur variance, a tendance à sous-estimer la corrélation entre les paramètres lorsqu’elle existe et ne traite pas bien les paramètres dont la variance est multimodale ou asymétrique. Dans de tels cas, la variance de l’optimum ne capturera pas bien sa sensibilité aux paramètres d’entrée, comme noté par Pianosi *et al.* (2016).

2.3.2 Mitigation

Déterminer la nature, l'origine ou encore quantifier les incertitudes affectant un ou plusieurs paramètres d'un design est secondaire dans cette thèse. L'accent est mis sur l'intégration de ces incertitudes dans le but de les mitiger à la source et ce, qu'elles proviennent d'effets d'échelle, d'hypothèses de modélisation, d'erreurs expérimentales, de l'incertitude entourant l'interprétation d'un essai de réponse thermique, un cas traité par Pasquier et Marcotte (2020), ou encore de l'utilisation de valeurs types issues de la littérature. Les méthodes présentées dans cette section, empruntées au domaine de la finance et utilisées habituellement pour la gestion de portfolio, intègrent toutes les incertitudes sans les quantifier.

Rockafellar et Uryasev (1999) ont développé une mesure appelée la *conditional value-at-risk* (CVaR). Cet outil, intimement lié à la *value-at-risk* (VaR) dont il s'inspire, permet de juger un portefeuille d'investissements en fonction de la moyenne des pertes qu'il est susceptible d'essuyer lors des pires scénarios envisageables. Au contraire, la VaR renseigne sur la perte maximale qu'il est susceptible d'essuyer à la limite d'un intervalle de confiance donné. Cette distinction est fondamentale pour comprendre la différence entre ces deux approches : la VaR renseigne sur une perte réalistement envisageable tandis que la CVaR prend uniquement en compte les pires retours sur investissements simulés pour proposer un choix d'investissement davantage à l'abri des pertes.

Obtenir la VaR ou la CVaR d'un portefeuille est un processus stochastique s'apparentant à une simulation de Monte-Carlo. Cette procédure est représentée schématiquement à la Fig. 2.6. Pour la VaR, une fois le nombre de portefeuilles simulés souhaités atteint, le plus bas retour sur investissement situé à la limite des α cas considérés par le gestionnaire devient la valeur recherchée. Au contraire, pour un deuxième gestionnaire souhaitant connaître la CVaR de ces mêmes scénarios, c'est la moyenne des cas non considérés par le premier (les $1 - \alpha$ cas) qui constituent la valeur qu'il recherche. Utiliser la CVaR comme fonction objectif pour un problème d'optimisation focalise la recherche d'investissements sur la queue de la distribution des pertes, comme noté par Chikhaoui (2015), et s'avère un indicateur sécuritaire lorsque l'historique des investissements disponibles est instable, inconnu ou n'obéit pas à une distribution précise. Cette méthode a l'avantage de traiter simultanément toutes les incertitudes en essayant directement de les mitiger au lieu de les quantifier.

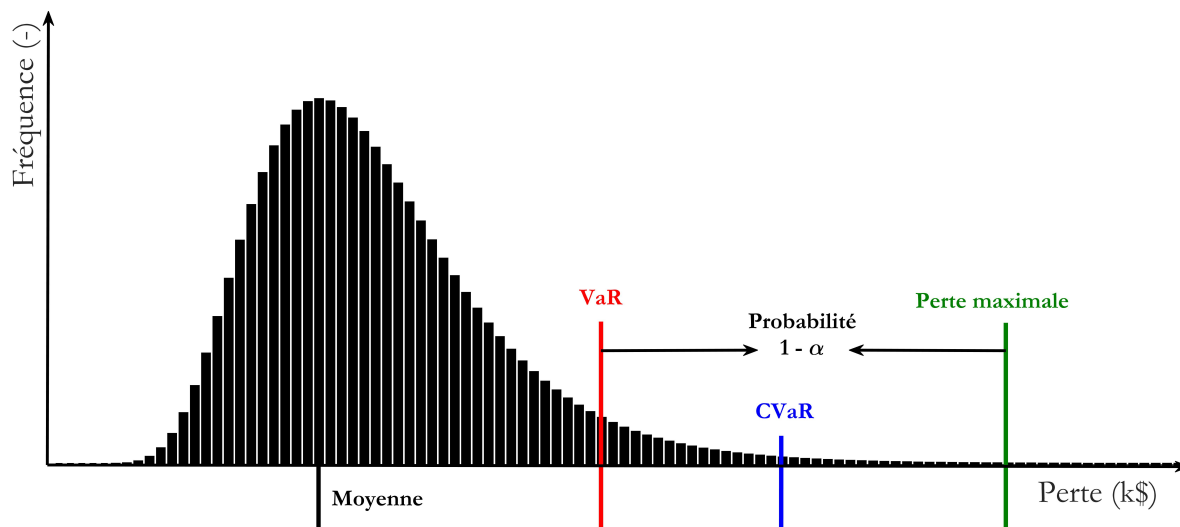


Figure 2.6 Démonstration de la relation entre la VaR et la CVaR. La CVaR constitue la moyenne des pires cas situés entre la VaR et la perte maximale pouvant être subie.

Depuis l'élaboration de la CVaR, cet indicateur financier a été intégré avec succès dans plusieurs problèmes d'optimisation en dehors du domaine de la finance, sans toutefois atteindre la géothermie de basse température. Yamout (2005) utilise les CVaR dans un problème complexe d'allocation de ressources en eau à l'intérieur d'une communauté dans laquelle les besoins des particuliers et des entreprises sont tous sujets à incertitude. Khodabakhsh et Sirouspour (2016) intègrent les CVaR à leur modèle de contrôle d'un micro réseau de distribution d'électricité avec piles servant à l'emménagement pour mitiger les risques. Ghaffari *et al.* (2016) et Bourry *et al.* (2008) utilisent les CVaR pour réduire les risques de « déversement de vent » (du vent inutilisé) dans un réseau électrique où l'électricité est générée en partie par des champs d'éoliennes. Ces derniers comparent d'ailleurs leur modèle avec de vraies données d'opération d'une ferme d'éoliennes dans le nord-ouest du Danemark. À cause des nombreuses incertitudes entourant les bulletins météorologiques dont les opérateurs se servent pour faire fonctionner leur ferme, l'utilisation des CVaR permet selon les auteurs de réaliser des économies substantielles en mitigeant le risque à la source.

Une des particularités intrinsèques de la méthode CVaR rend cependant son application directe pour la géothermie de basse température impossible. En effet, elle nécessite que l'éventail des solutions possibles, soit les actions de différentes compagnies pouvant être achetées pour constituer un portefeuille d'investissement par exemple, puisse être combiné librement de façon linéaire. Autrement dit, les différentes actions, chacune affublée d'une incertitude sur son cours futur, doivent pouvoir être acquises indépendamment puis combinées pour former un portefeuille représentant la tolérance au risque du gestionnaire. Cela est possible dans le

domaine de la finance ou lorsque l'on considère l'achat individuel d'éoliennes, mais à cause des interactions thermiques entre puits dans le sol et du couplage non linéaire entre sol et bâtiment, chaque puits et chaque thermopompe ne peuvent pas être considérés comme des investissements distincts puis combinés au sein d'un système hybride selon la tolérance au risque de l'utilisateur.

Des contraintes similaires empêchant la combinaison linéaire de solutions distinctes s'appliquent au dimensionnement d'infrastructures dont différents designs ne peuvent être combinés pour formuler un design optimal. Ye et Tiong (2000), pour contourner cette limitation, ont développé la méthode de la *Net Present Value-at-risk* (ou valeur actuelle nette à risque, VANaR). Cette méthode, de façon similaire au CVaR, propose l'utilisation d'un indicateur financier comme fonction objectif qui représente l'espérance des VANs lors des pires scénarios envisageables (les $1 - \alpha$ cas) pour les variables incertaines. Le pourcentage des pires scénarios à considérer est laissé encore une fois à la discrétion de l'utilisateur en fonction de la tolérance au risque du client. Dans leur article, les auteurs démontrent que deux projets de centrales peuvent apparaître similaires sous la loupe d'indicateurs financiers traditionnels (période de retour sur investissement, taux de rendement interne, VAN), mais différer grandement lorsque la VANaR est considérée. Par exemple, on estime une VAN de 569 *k*\$ pour les deux centrales aux termes du projet, mais lorsque les 5% pires scénarios sont considérés, le projet de centrale A vaut malgré tout 69 *k*\$ en moyenne tandis que le projet de centrale B engendre des pertes moyennes de 302 *k*\$.

En résumé, au stade actuel, la géothermie de basse température ne met en place que des mesures de détermination du risque. Ces mesures, des analyses de sensibilité pour la plupart, permettent de renseigner sur la stabilité d'un design, mais ne mettent pas à l'abri des pertes lorsque les conditions réelles d'opération diffèrent des conditions supposées lors du dimensionnement. Pour contourner cette limitation, d'autres domaines comme la finance, les réseaux de distribution d'énergie contenant des sources d'électricité intermittentes, l'hydrologie et le génie civil utilisent déjà des indicateurs financiers performants qu'il serait intéressant d'exporter et d'adapter pour le design de systèmes géothermiques hybrides.

CHAPITRE 3 DÉMARCHE DE L'ENSEMBLE DU TRAVAIL

Tel qu'introduit au Chapitre 1, le but global visé par cette thèse est de tenir compte du risque lors du dimensionnement de systèmes géothermiques hybrides en boucles fermées verticales. Ce but s'articule en trois objectifs qui découlent des lacunes ayant été mises en évidence dans la revue de littérature du Chapitre 2.

Premier objectif : Développer une nouvelle méthode de calcul rapides de *g-function* utilisant une formulation linéaire exprimée par des matrices de blocs et des polynômes de Tchebychev.

Au chapitre 4, le premier enjeu entourant les défis du couplage entre sol et bâtiment est abordé avec l'élaboration d'une nouvelle méthode pour le calcul de fonctions de transfert. Elle est présentée sous la forme d'un premier article de revue (Dusseault *et al.*, 2018). La méthode développée, appelée formulation par matrices de bloc (BMF), s'articule autour de l'utilisation de grandes matrices creuses permettant de résoudre simultanément un système d'équations linéaires. La résolution de ce système permet le calcul de l'ensemble des interactions thermiques entre les puits du champ géothermique et donc, permet de calculer la fonction de transfert d'un champ en une seule étape. Cette méthode se veut non seulement rapide à exécuter, mais également simple à implémenter dans un langage de programmation.

La construction des grandes matrices de bloc utilisées par la BMF requiert plusieurs dizaines de milliers de résolution du modèle de la source linéique finie (SLF) (Claesson et Javed, 2011). À cause du nombre de SLF à calculer et de la complexité de l'intégrale qu'elle contient, cette opération consomme une grande partie du temps total de calcul de la BMF. Pour y remédier, une formulation de la SLF dérivée par rapport au temps a été élaborée. Cette nouvelle version de la SLF ne contient pas d'intégrale, peut être approximée par un polynôme puis intégrer analytiquement au lieu de numériquement après coup. Cette opération produit le même résultat que l'évaluation du modèle initial, mais puisque cette intégration est résolue plus rapidement, cela permet de réduire davantage le temps de détermination de la fonction de transfert.

Deuxième objectif : Utiliser un réseau neuronal afin de calculer plus rapidement des *g-functions* dans des configurations régulières et irrégulières d'échangeurs de chaleur souterrains.

Au Chapitre 5, un second article de revue (Dusseault et Pasquier, 2019) présente une méthodologie qui permet l'élaboration et l'entraînement de réseaux neuronaux (ANN) pour

l'estimation de fonctions de transferts. Il explique en détail comment une base de données de 500,000 *g-functions* conçue pour l'entraînement d'un ANN a été assemblée à l'aide de la BMF issue du premier article de revue. Ensuite, l'architecture utilisée dans la construction du réseau est présentée. Cette architecture, inspirée d'un ANN présenté en conférence et ayant fait l'objet d'une publication qui n'est pas présentée dans cette thèse (Dusseault et Pasquier, 2018), permet de construire des *g-functions* pour des champs dont les forages sont placés selon des configurations régulières et irrégulières dans un champ. De plus, il permet de traiter des champs allant de un à dix forages, des rayons et des profondeurs d'enfouissement de puits variables ainsi que des propriétés thermiques du sol changeantes. Le but visé par l'élaboration de cet ANN est d'accélérer suffisamment le calcul de fonctions de transfert pour impacter significativement le temps total de calcul lors du design par optimisation de systèmes géothermiques hybrides.

Troisième objectif : Réaliser une étude de cas sur l'impact de l'utilisation d'un algorithme permettant la mitigation du risque lors de la conception d'un système géothermique hybride par programmation mixte en nombres entiers sous incertitude.

Au chapitre 6, les deux derniers enjeux abordés au Chapitre 2, soit l'intégration de l'incertitude dans le design et l'optimisation en nombres entiers, sont abordés sous la forme d'un troisième article de revue écrit pour cette thèse (Dusseault et Pasquier, 2020). Dans cet article, on présente l'algorithme développé permettant la conception d'un système géothermique hybride par programmation en nombres entiers mixte. Ce dernier utilise une combinaison de gradient conjugué contraint et de séparation et évaluation pour contraindre le nombre de forages et de thermopompes à évoluer en nombre entier. De plus, la valeur actuelle nette à risque (VANaR) y est utilisée comme fonction objectif pour le problème d'optimisation. Cet indicateur financier emprunté au génie civil permet d'intégrer l'incertitude à un design pour fournir un optimum en fonction d'un seuil de tolérance au risque. Pour finir, une étude de cas réalisée au moyen de ce nouvel algorithme y est présentée puis discutée pour voir quelles sont les implications positives et négatives de l'utilisation de la VANaR.

Le chapitre 7 présente une discussion générale portant sur les principales contributions de ces recherches et le chapitre 8 clôt la thèse en présentant les conclusions et suggère des recommandations pour des travaux futurs.

CHAPITRE 4 ARTICLE 1 : A BLOCK MATRIX FORMULATION FOR EFFICIENT G-FUNCTION CONSTRUCTION

Bernard Dusseault^{1*}, Philippe Pasquier¹, Denis Marcotte¹

Renewable Energy, Volume 121, June 2018, Pages 249-260

¹ Civil, Geological and Mining Department, Polytechnique Montreal, C.P. 6079 Succ. Centre-ville, Montreal H3C 3A7, Canada.

* Corresponding author : bernard.dusseault@polymtl.ca.

4.1 Abstract

A compact block matrix formulation allowing fast construction of the g-function of a ground heat exchanger is presented. This new formulation is nor sequential nor iterative, doesn't require use of Laplace or Fourier transform and provides a g-function by solving a single system of linear equations assembled in a block matrix system. A method to accelerate the integration of the finite-line source model based on Chebyshev polynomials is also introduced. Although it suffers from a fixed cost in terms of computational time, this approach allows to speed up the g-function assessment even further when used jointly with the block matrix formulation on large fields. By using both strategies, constructing the g-function of a ground heat exchanger composed of 50 regularly spaced boreholes can be achieved in less than half-second while committing only a small relative error. The speed and compactness of the block matrix formulation could be useful to design ground heat exchangers with optimization-based algorithms, which can require the assessment of several thousand g-functions.

4.2 Introduction

For the simulation and design of ground heat exchangers (GHE), the concept of g-function introduced by Eskilson (1987) a few decades ago is still widely used to simplify the computation of the complex thermal interactions occurring between the various boreholes of a GHE. This useful numerical tool acts as a transfer function linking the total ground heat exchange rate and the average borehole wall temperature (Eskilson, 1987). By convolving a time varying heat load signal with the g-function, one can obtain easily the evolution of the fluid temperature over time and predict the behavior of a ground-coupled heat pump system (Marcotte et Pasquier, 2008).

The construction of a g-function is however site specific as it is dependent on the field geometry and ground thermal properties. Numerous methods have been proposed in recent years to construct, under different assumptions on the boundary conditions (Cimmino, 2015), the g-function of a GHE. It is nowadays common practice to use the finite-line source (FLS) model (Claesson et Javed, 2011; Lamarche et Beauchamp, 2007; Zeng *et al.*, 2002) to represent the heat emanating from a borehole and to assess the thermal perturbation caused by the source at the location of other boreholes. Thus, the borehole wall temperature can be computed for each borehole of the GHE. Using the FLS model, Cimmino et Bernier (2014) implemented a semi-analytical method to generate g-functions for GHE. A similar procedure, combining numerical and analytical models, was also used in (Lazzarotto, 2016b,c,a) to account for arbitrarily inclined boreholes. The problem of inclined boreholes was also tackled by other authors (Cui *et al.*, 2006; Lamarche, 2011; Marcotte et Pasquier, 2009).

It was observed (Cimmino *et al.*, 2012) that implementing a heat-load aggregation scheme to speed up the computation (Bernier *et al.*, 2004) can lead to a loss in accuracy. To avoid that outcome, the computation of heat transfers and average borehole wall temperatures can be addressed sequentially with Laplace transforms to deal with convolution products (Cimmino *et al.*, 2013; Lamarche, 2009; Lazzarotto, 2014). Marcotte and Pasquier used fast Fourier transforms (Marcotte et Pasquier, 2008; Pasquier et Marcotte, 2013) and modified the nature of the transfer function to describe GHEs where the boreholes are connected in parallel, in series or in a mixed arrangement (Marcotte et Pasquier, 2014). To simulate GHEs where heat is simultaneously extracted from and injected to a GHE, a network-based methodology enabling the operation of a GHE under different piping arrangements was proposed (Lazzarotto, 2014). A step further has been made recently (Lamarche, 2017) with a methodology allowing multi-years simulations for complex and irregular borehole fields having different inlet conditions in a matter of seconds.

Most algorithms now allow the construction of a g-function in a few minutes, which is quite reasonable for simulation or sizing purposes. However, for optimization-based design of GHE (Hénault *et al.*, 2016; Huang *et al.*, 2015; Retkowski et Thöming, 2014), the optimization process of the cost or utility function requires the construction of several hundreds of g-functions to evaluate the objective function value and its derivative. In fact, Hénault (2015) observed for a financial optimization problem that the sole assessment of the g-functions can take up 75 % of the total computation time, which was about 3 hours.

At the moment, the objective function gradient is usually assessed by finite-differentiation. However, recently the limitations of using a finite-difference scheme when assessing the gradient were shown (Dusseault *et al.*, 2017). It was also observed that computing the gradient

analytically could provide more accurate results while speeding up the whole optimization process. Furthermore, it was demonstrated that for a simple objective function, analytical derivation provides more accurate gradients, which leads to a smaller number of iterations to reach convergence. This exploratory work (Dusseault *et al.*, 2017) was based on a sequential algorithm (Marcotte et Pasquier, 2014) for the g-function construction. Although the development was limited to a single optimization parameter, the analytical derivation of the gradient proved cumbersome in a sequential context.

The goal of this article is to present a new compact block matrix formulation to construct the g-function of a GHE. This new formulation is nor sequential nor iterative, doesn't require use of Laplace or Fourier transform, provides a dimensionless g-function by solving a single system of linear equations and has the potential to ease gradient derivation. In this paper, the compact block matrix formulation will be first presented. Secondly, an additional method used to speed up the evaluation of the FLS model and relying on Chebyshev polynomials will be illustrated. This approach allows to reproduce, with significant accuracy, the FLS solution. It will be shown that by using both a block matrix formulation and Chebyshev polynomials, computation times of less than half a second can be achieved to assess the g-function of a 50 boreholes field over a period of 40 years.

4.3 A block matrix formulation for g-functions

The approach brought forth in this section is twofold. First of all, a block matrix notation will be presented step-by-step to express simultaneously all thermal perturbations that occur between the boreholes at all time steps, without requiring Fourier or Laplace transforms. Then, we will show how the integrand of the FLS model proposed by Claesson et Javed (2011) can be approximated by a Chebyshev polynomials and integrated analytically instead of being evaluated numerically. These two approaches can be used jointly to further accelerate the calculation of g-functions.

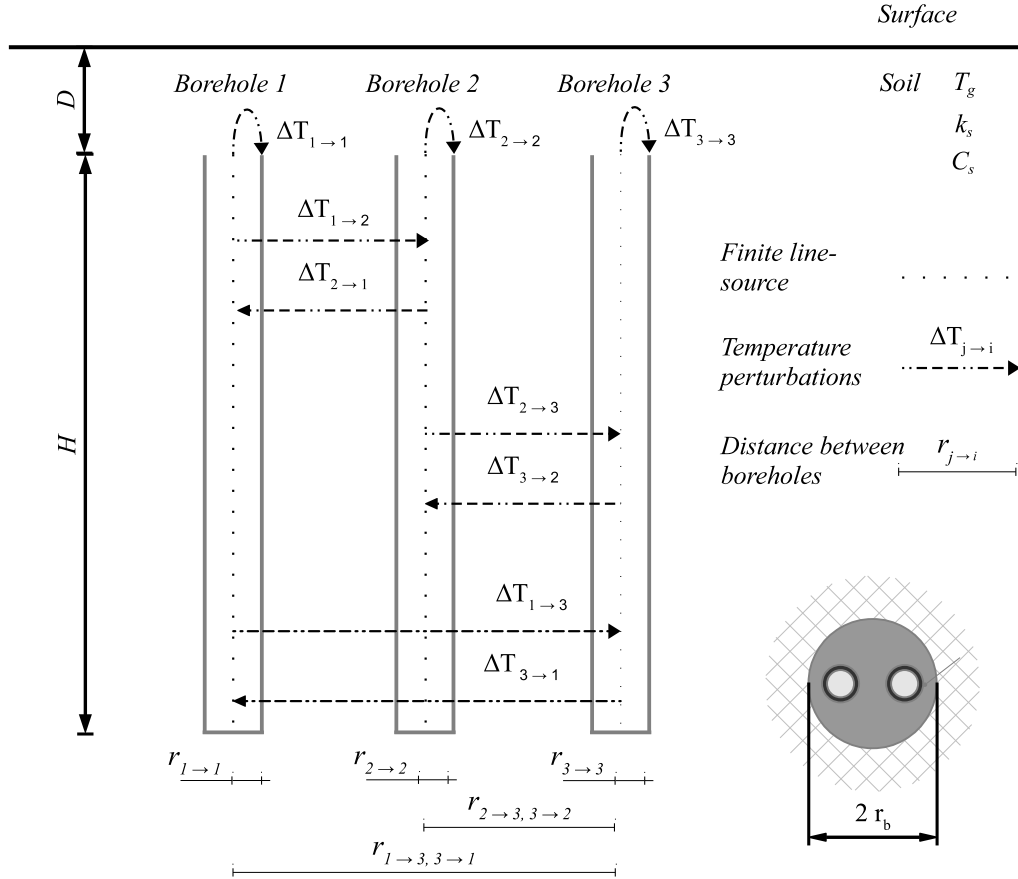


Figure 4.1 Illustration of the thermal interactions between the boreholes of a GHE.

In this article, the FLS model (Eq. 4.1 to 4.3) is used, whether as-is or approximated with Chebyshev polynomials, to compute the thermal influence of borehole j has on borehole i at time t_k . For a borehole of length H , buried at a depth D in a medium of thermal conductivity k_s , the thermal perturbation caused by borehole j on borehole i located at a distance $r_{j \rightarrow i}$ (see Figure 4.1) is given by :

$$\Delta T(r_{j \rightarrow i}, t_k) = \frac{\tilde{q}}{4\pi k_s} \int_{1/\sqrt{(4k_s t_k)/C_s}}^{\infty} \exp(-r_{j \rightarrow i}^2 s^2) \frac{Y(Hs, Ds)}{Hs^2} ds \quad (4.1)$$

where

$$Y(h, d) = 2 \operatorname{ierf}(h) + 2 \operatorname{ierf}(h + 2d) - \operatorname{ierf}(2h + 2d) - \operatorname{ierf}(2d) \quad (4.2)$$

and where

$$\mathbf{ierf}(X) = X \mathbf{erf}(X) - \frac{1}{\sqrt{\pi}} (1 - \exp(-X^2)) \quad (4.3)$$

Throughout this work, the notation $g_{j \rightarrow i}$ will refer to the solution of Eq. 4.1 obtained with a constant heat flux $\tilde{q}_j = 1 \text{ W/m}$ and at a distance $r_{j \rightarrow i}$ from source j . Therefore, $g_{j \rightarrow i} = \Delta T(r_{j \rightarrow i}, t_k) / \tilde{q}$ and the units of g are in $m^\circ C/W$.

4.3.1 Block matrix formulation

To illustrate easily the block matrix formulation, the equations behind it are approached step-by-step in the following section by first considering three simple situations. The first one consists in a single borehole simulated at several time steps (Section 4.3.1) while the second one is a GHE made of several boreholes but simulated for a single time step (Section 4.3.1). By proceeding this way, the mechanics of the method is grasped more easily before tackling the third situation which represents a GHE made of many boreholes and simulated at several time steps in subsection 4.3.1.

Single borehole and multiple time steps

Let's start by considering a GHE formed of a single borehole that emits a unit heat load $\tilde{q}$. To obtain the g -function of this GHE, one must first calculate the evolution of its wall temperature throughout the simulation. In such field, no spatial superposition is required to account for the interacting boreholes. However, temporal superposition is required to account for the thermal influence the borehole has on itself as time goes by. At any given moment in time (t_k), the temperature at the borehole wall (T_b) can be represented by the following equation (Eq 4.4) :

$$T_b(t_k) = T_g + \Delta T(r, q, t_k) \quad (4.4)$$

where T_g is the initial ground temperature and $\Delta T(t_k)$ is the temperature perturbation resulting from past and present thermal interactions between the borehole and itself.

As noted by Marcotte et Pasquier (2008), the temperature computation is a convolution product that can be solved in the time or frequency domain using the incremental heat load function f given by :

$$f(t_k) = q(t_k) - q(t_{k-1}) \quad (4.5)$$

This function serves to track the temporal variations in heat extraction rates in each borehole of the GHE throughout a simulation. However, since a single borehole is considered in this

section, the heat extraction rate over the field is $q(t_k) = \tilde{q} = 1$ and f takes for now the value of 1 during the first time step and 0 for all subsequent time steps.

With this definition for f , the temperature perturbations can be expressed along the borehole wall ($r_{1 \rightarrow 1} = r_b$) by

$$\Delta T(r_{1 \rightarrow 1}, q, t) = (\mathbf{f} * \mathbf{g})(t) = \sum_{k=1}^n f(t_k) g(t - t_{k-1}) \quad (4.6)$$

where n is the number of time steps and where $t_0 = q_0 = 0$.

Equation 4.6 can now be expressed easily in a matrix form by :

$$\begin{bmatrix} \Delta T(t_1, r_{1 \rightarrow 1}) \\ \Delta T(t_2, r_{1 \rightarrow 1}) \\ \Delta T(t_3, r_{1 \rightarrow 1}) \\ \vdots \\ \Delta T(t_n, r_{1 \rightarrow 1}) \end{bmatrix} = \begin{bmatrix} g(t_1 - t_0) & 0 & 0 & \dots & 0 \\ g(t_2 - t_0) & g(t_2 - t_1) & 0 & \dots & 0 \\ g(t_3 - t_0) & g(t_3 - t_1) & g(t_3 - t_2) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g(t_n - t_0) & g(t_n - t_1) & \dots & g(t_n - t_{n-2}) & g(t_n - t_{n-1}) \end{bmatrix} \begin{bmatrix} f(t_1) \\ f(t_2) \\ f(t_3) \\ \vdots \\ f(t_n) \end{bmatrix} \quad (4.7)$$

or in a more compact matrix notation by :

$$\Delta \mathbf{T} = \mathbf{G} \mathbf{f} \quad (4.8)$$

Solution of the previous equation (Eq. 4.8) provides, simultaneously, the temperature perturbations at a distance r for the n time steps.

Eq. 4.4 can now be used to assess the borehole wall temperature T_b providing the initial ground temperature is known. To calculate the non dimensional g-function for the case of a single borehole, one only has to construct the convolution matrix $\mathbf{G}$ using the FLS model computed with a unit heat load $q(t_k) = \tilde{q} = 1$ W/m and the incremental heat load vector $\mathbf{f}$ corresponding to $q(t_k) = 1$ W/m and solve Eq. 4.8. The adimensional g-function is then obtained by $\tilde{g}(t_k) = 2\pi k_s \Delta T(t_k) / \tilde{q}$.

However, let's suppose for the moment that nor $f(t_k)$ nor $\Delta T(t_k, r)$ are known. This system of linear equations has now $2n$ unknowns and only n equations. To ensure the uniqueness of the solution, a constraint on the heat extraction rates q needs to be added. For the case of a single borehole, this constraint states that $q(t_k) = 1 \forall k$, which leads to the following incremental constraint vector :

$$\mathbf{\Lambda} = [1 \quad 0 \quad 0 \quad \dots \quad 0]' \quad (4.9)$$

a n by 1 vector that consists of a single 1 followed by $n - 1$ zeros. In matrix notation, the constraints on q become

$$\mathbf{I} \mathbf{f} = \mathbf{\Lambda} \quad (4.10)$$

where $\mathbf{I}$ is a n by n identity matrix. Using a matrix notation, Eq. 4.4 can be rewritten as follows :

$$\mathbf{G} \mathbf{f} - (\mathbf{T}_b - T_g) = \mathbf{0} \quad (4.11)$$

Combining the constraints described by Eq. 4.10 and the system of linear equations of Eq. 4.11 produces the following matrix system :

$$\begin{bmatrix} \mathbf{G} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ -(\mathbf{T}_b - T_g) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{\Lambda} \end{bmatrix} \quad (4.12)$$

where $\mathbf{0}$ is either a $n \times n$ square matrix of zeros (left-side of the equation) or a $n \times 1$ vector of zeros (right-side of the equation).

Solving the system of linear equations described by Eq. 4.12 gives vectors $\mathbf{f}$ and $-(\mathbf{T}_b - T_g)$. Following the work of Marcotte et Pasquier (2014), the dimensionless g-function for a GHE made of a one borehole ($m = 1$) is given by the following equation :

$$\tilde{g}(t_k) = 2\pi k_s m (T_b(t_k) - T_g) / \tilde{q} \quad (4.13)$$

Multiple boreholes, one time step

Let's start the second part of this demonstration by considering a GHE composed of m boreholes each emitting a different heat flux q_j at a single time step ($n = 1$). In such field, let's put all the distances between the boreholes in a matrix $\mathbf{r}$. This matrix contains the distances $r_{j \rightarrow i}$ between the source j and the destination i as follows :

$$\mathbf{r} = \begin{bmatrix} r_{1 \rightarrow 1} & r_{1 \rightarrow 2} & \dots & r_{1 \rightarrow m} \\ r_{2 \rightarrow 1} & r_{2 \rightarrow 2} & \dots & r_{2 \rightarrow m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m \rightarrow 1} & r_{m \rightarrow 2} & \dots & r_{m \rightarrow m} \end{bmatrix} \quad (4.14)$$

where

$$r_{j \rightarrow i} = r_b \quad \forall \quad i = j \quad (4.15)$$

To obtain the g-function of this GHE, one must find the m values of q_j ensuring that the wall

temperature of each borehole is the same at the end of the single time step of this simulation. Here, contrarily to the first case, the spatial superposition is required to account for the thermal interactions between the boreholes. For any destination borehole i in the field, the temperature at the borehole wall (T_{b_i}) can be represented by the following equation :

$$T_{b_i} = T_g + \sum_{j=1}^m \Delta T_{j \rightarrow i} \quad (4.16)$$

where $\Delta T_{j \rightarrow i}$ is the temperature perturbation induced at the wall of borehole i resulting from the heat emitted by one of the m surrounding boreholes. Since the simulation of this GHE contains only one time step, Eq. 4.16 can be rewritten as follows :

$$T_{b_i} = T_g + \sum_{j=1}^m g(r_{j \rightarrow i}) f_j \quad (4.17)$$

To make sure that the borehole wall temperature at each borehole is the same, a constraint needs to be added to ensure that $T_b = T_{b_i} \forall i$. Also, since there are several boreholes, the sum of the m heat extraction rates during the first and only time step of this case needs to be equal to 1. This is expressed by :

$$\sum_{j=1}^m f_j = \mathbf{1}' \mathbf{f} = 1 \quad (4.18)$$

Combining these two sets of constraints leads to the following matrix system :

$$\begin{bmatrix} g_{1 \rightarrow 1} & g_{2 \rightarrow 1} & g_{3 \rightarrow 1} & \cdots & g_{m \rightarrow 1} & 1 \\ g_{1 \rightarrow 2} & g_{2 \rightarrow 2} & g_{3 \rightarrow 2} & \cdots & g_{m \rightarrow 2} & 1 \\ g_{1 \rightarrow 3} & g_{2 \rightarrow 3} & g_{3 \rightarrow 3} & \cdots & g_{m \rightarrow 3} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{1 \rightarrow m} & g_{2 \rightarrow m} & g_{3 \rightarrow m} & \cdots & g_{m \rightarrow m} & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_m \\ -(T_b - T_g) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (4.19)$$

In Eq. 4.19, the different terms of $g_{j \rightarrow i}$ are lay out following the same disposition as the terms in the matrix $\mathbf{r}$. By using a compact matrix notation with $\mathbf{0}$ and $\mathbf{1}$ to describe $n \times 1$ vectors of zeros or ones, Eq. 4.19 becomes finally :

$$\begin{bmatrix} \hat{\mathbf{g}} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix} \begin{bmatrix} \mathbf{f} \\ -(T_b - T_g) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \quad (4.20)$$

where the hat notation $\hat{}$ refers to the interactions between the boreholes.

Since there is only one time step, solution of Eq. 4.20 gives the heat extraction rates ($f_j = q_j$) at each borehole ensuring that $T_b = T_{b_i} \forall i$. Then, the $-(T_b - T_g)$ value can be used with Eq. 4.13 to assess the g-function $\tilde{g}$ of this GHE for a single evaluation time.

Multiple boreholes, multiple time steps

This section aims to show how the previous equations can be combined to account simultaneously for the temporal and the spatial superposition of the thermal interactions between boreholes in a GHE. In order to do this, a block matrix $\hat{\mathbf{G}}$ formed of triangular matrix $\mathbf{G}_{j \rightarrow i}$ have to be constructed. By constructing the block matrix $\hat{\mathbf{G}}$ with the same pattern as matrix $\mathbf{r}$, we make sure each sub-matrix $\mathbf{G}_{j \rightarrow i}$ gets multiplied with its corresponding incremental heat extraction rate $\mathbf{f}_j$:

$$\Delta \mathbf{T}_{i \rightarrow j} = \mathbf{G}_{j \rightarrow i} \mathbf{f}_j \quad (4.21)$$

Once combined with the general version of the two previously mentioned conditions on heat extraction rates, a general block matrix version of the systems of linear equations of Eq. 4.12 and Eq. 4.20 can be formulated as :

$$\begin{bmatrix} \mathbf{G}_{1 \rightarrow 1} & \mathbf{G}_{2 \rightarrow 1} & \mathbf{G}_{3 \rightarrow 1} & \dots & \mathbf{G}_{m \rightarrow 1} & \mathbf{I} \\ \mathbf{G}_{1 \rightarrow 2} & \mathbf{G}_{2 \rightarrow 2} & \mathbf{G}_{3 \rightarrow 2} & \dots & \mathbf{G}_{m \rightarrow 2} & \mathbf{I} \\ \mathbf{G}_{1 \rightarrow 3} & \mathbf{G}_{2 \rightarrow 3} & \mathbf{G}_{3 \rightarrow 3} & \dots & \mathbf{G}_{m \rightarrow 3} & \mathbf{I} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{G}_{1 \rightarrow m} & \mathbf{G}_{2 \rightarrow m} & \mathbf{G}_{3 \rightarrow m} & \dots & \mathbf{G}_{m \rightarrow m} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} & \mathbf{I} & \dots & \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \\ \vdots \\ \mathbf{f}_m \\ -(\mathbf{T}_b - T_g) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \Lambda \end{bmatrix} \quad (4.22)$$

Using the $\hat{}$ notation to indicate that a matrix or a vector has been replicated m times, an even more compact notation is given by :

$$\begin{bmatrix} \hat{\mathbf{G}} & \hat{\mathbf{I}} \\ \hat{\mathbf{I}}' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ -(\mathbf{T}_b - T_g) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{0}} \\ \Lambda \end{bmatrix} \quad (4.23)$$

The terms in the vector $-(\mathbf{T}_b - T_g)$ can then be used in Eq. 4.13 to assess the g-function $\tilde{g}$ of a multiple boreholes GHE for all of the time steps of the simulation simultaneously. To help the reader understand how Eq. 4.22 can be constructed for a given GHE, a small numerical example is presented in Annex A.

4.3.2 Chebyshev polynomials approximation

This section presents a methodology that allows for faster evaluation of the FLS model (Eq. 4.1 to Eq. 4.3), which can be used jointly to the block matrix method, to speed up the evaluation time of this FLS model. Smaller computation times are achieved by estimating the integration within the FLS model by using polynomials instead of resolving it numerically (Trefethen, 2007). In the Results and Discussion section, the case study developed for the present article will be solved using both the FLS model and Chebyshev polynomials to show the efficiency of the proposed method.

We will show in Section 4.6 that this analytical approach proves to be faster, especially when the FLS model is used several dozens or hundreds of times to account for all the time steps and all the thermal interactions in large BHE.

Analytical derivation

The simplest existing formulation of the FLS model is provided by Claesson et Javed (2011). Their formulation involves a single time integral. Here, we seek to eliminate this integration. In order to do so, the analytical formula (Eq. 4.1) is derived according to time. The FLS model is once again computed with a heat flux $\tilde{q} = 1$ (W/m). According to the notation introduced in Eq. 4.1, the derivative of the FLS model is equal to :

$$\Delta\dot{T}(r, t) = \frac{-\tilde{q}}{4\pi\sqrt{k_s C_s t}} \exp(-C_s r^2 / 4k_s t) Y \left(\frac{H}{\sqrt{4k_s t / C_s}}, \frac{D}{\sqrt{4k_s t / C_s}} \right) \quad (4.24)$$

Polynomial approximation

In a second step, $\Delta\dot{T}(r, t)$ is approximated using Chebyshev polynomials of second kind. Thus, the polynomial approximation of Eq. 4.24, $\Delta\dot{T}_c$, is provided by :

$$\Delta\dot{T}(r, t) \cong \Delta\dot{T}_c(r, t) = \sum_{k=0}^n c_k U_k(t), t \in [-1, 1] \quad (4.25)$$

where c_k are the unknown polynomial coefficients and $U_k(t)$ are the Chebyshev polynomials of the second kind and of order k evaluated at t . The Chebyshev polynomials $U_k(t)$ are defined by recurrence using the following starting values :

$$U_0(t) = 1 \quad (4.26)$$

$$U_1(t) = 2t \quad (4.27)$$

$$U_{k+1}(t) = 2tU_k(t) - U_{k-1}(t) \quad (4.28)$$

As an example, the first six polynomials that are constructed using the recurrence defined by Eqs. 4.26 to 4.28 are shown in Fig 4.2.

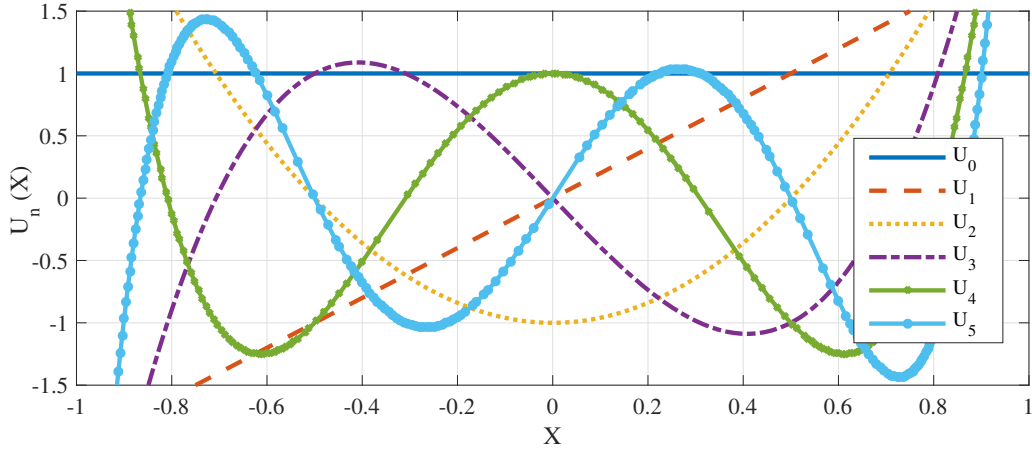


Figure 4.2 The six first Second type Chebyshev polynomials in their original unshifted domain.

Since the default range of Chebyshev polynomial is $[-1, 1]$, shifted polynomials of the form :

$$U_k^*(t) = U_k(at - b) \quad (4.29)$$

are used in order to expand it. The coefficients a and b are selected so that the new range encompasses all the duration of the simulated period.

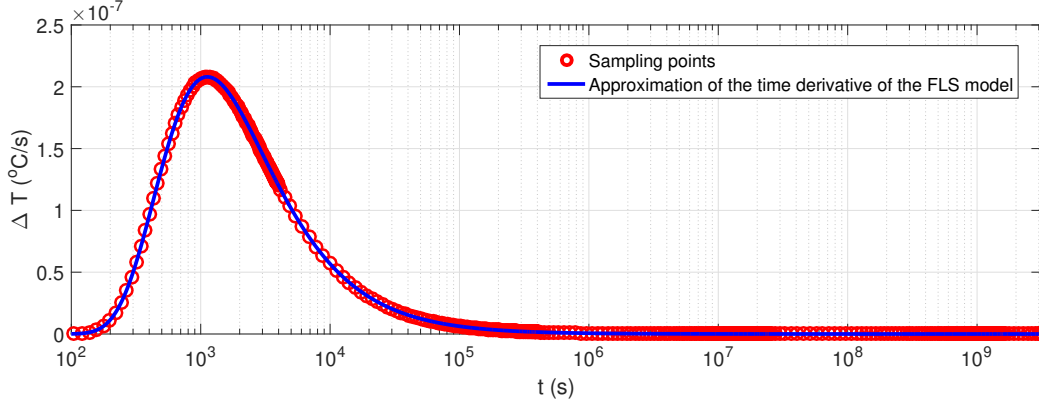


Figure 4.3 Sampling points used by Chebyshev polynomials superimposed on the approximation the time derivative of the FLS model.

To implement the Chebyshev polynomials approximation of Eq. 4.24 and find the unknown coefficient a , b and c , the simple and efficient open-source package Chebfun (Driscoll *et al.*, 2014) was used. To reduce the computation burden, Eq. 4.24 was however sampled at the sampling points illustrated in Fig 4.3 and 4.4. The Chebfun package was then able to find the c coefficients and the polynomial degree minimizing the error between $\Delta\dot{T}(r, t)$ and $\Delta\dot{T}_c(r, t)$ down to machine precision.

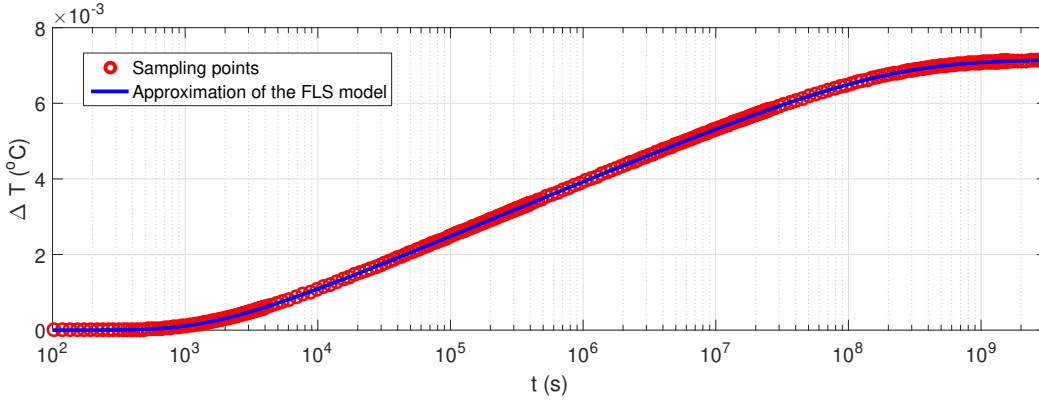


Figure 4.4 Sampling points used by Chebyshev polynomials superimposed on the FLS model.

Polynomial Integration

Finally, with $\Delta\dot{T}_c(r, t)$ now expressed as a linear combination of shifted Chebyshev polynomials, the temperature derivative can now be rapidly integrated back according to time

through :

$$\Delta T(r, t) \cong \int_0^t \Delta \dot{T}_c(r, t) dt = \int_0^t \sum_{k=0}^n b_k U_k^*(t) dt \quad (4.30)$$

A useful relationship between Chebyshev polynomials of the first ($\mathbb{T}_k$) and second type (U_k) links the integral of second type polynomials through :

$$\int U_k dt = \mathbb{T}_{k+1}/(k+1) \quad (4.31)$$

Using Eq. 4.31, the integration is performed analytically and the numerical integration on the right side of Eq. 4.30 is avoided, thus saving computation time for the solution of the integral.

Figure 4.4 compares the results of Eq. 4.30 to the temperature obtained by numerical integration of Eq. 4.1. This new polynomial formulation of the FLS model is accurate up to machine precision while being evaluated much faster. As we will see in the result section, although this procedure does suffer from a fixed computation cost, using Chebyshev polynomials ends up being cost-effective most of the time since analytical integration is more rapidly computed than numerical integration by a computer. To summarize the computation steps, Figure 4.5 presents a list of the inputs and outputs of the proposed method as well as a brief summary of the overall algorithm.

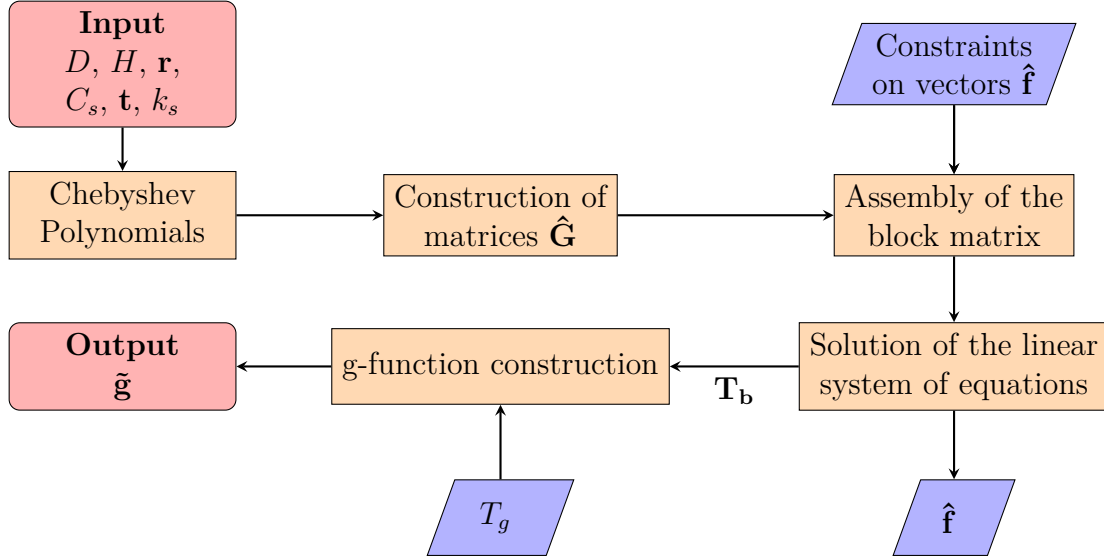


Figure 4.5 Overview of the block matrix algorithm.

4.4 Verification with Two Independent Algorithms

To verify the implementation of the block matrix formulation, the g-function given by the block matrix formulation was compared with the one obtained with two independent algorithms conveniently named the sequential algorithm (Marcotte et Pasquier, 2014) and the iterative algorithm (Pasquier et Marcotte, 2013; Pasquier *et al.*, 2013). In addition to this double verification process, the speed of the block matrix formulation was compared with the one of the sequential method for two reasons. First, the sequential algorithm is the direct predecessor of the block matrix approach and it was a legitimate question to see how reformulating the sequential system of equations into a single system led to computation gains. Second, the sequential method was, to the best of our knowledge, one of the fastest method. Finally, though being fairly slow and inefficient for large GHEs, the iterative algorithm was deemed more accurate because it modifies simultaneously the heat-fluxes of every boreholes at all time steps until it finds good agreement amongst the borehole wall temperatures across the field. The following section presents briefly the two independent algorithms used for the verification process.

4.4.1 Sequential Algorithm

The block matrix formulation developed in the present article is greatly inspired by the sequential algorithm of Marcotte et Pasquier (2014) and therefore share many similarities in its formulation and nomenclature. This algorithm allows the calculation of g-functions for boreholes that can be connected either in parallel, series or in a mixed-arrangement. However, for simplicity and since the method developed in the present article has been developed for pure parallel configurations, the equations for series and mixed-arrangement won't be covered here.

The method works by solving sequentially, time step per time step, Eq. 4.32 to obtain the temperature perturbation at each borehole of the field given an historical temperature perturbation vector $\mathbf{h}$:

$$\Delta \mathbf{T} = \hat{\mathbf{g}} \hat{\mathbf{q}} + \mathbf{h} \quad (4.32)$$

To insure uniqueness of the solution, constraints on the heat extraction rates similar to Eq. 4.18 are added to Eq.4.32, which leads to the matrix formulation of Eq. 4.33 :

$$\begin{bmatrix} \hat{\mathbf{g}} & \mathbf{1} \\ \mathbf{1}' & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{q}} \\ -(T_b - T_g) \end{bmatrix} = \begin{bmatrix} \mathbf{h} \\ 1 \end{bmatrix} \quad (4.33)$$

This system, although similar to the block matrix formulation introduced in the present paper, is designed to split the thermal perturbations in the field into two distinct parts : a current and an historical part. In fact, the product $\hat{\mathbf{g}}\hat{\mathbf{q}}$ accounts for the current thermal perturbations while vector $\mathbf{h}$ accounts for all past perturbations. The system of linear equations given by Eq. 4.33 must be solved sequentially, time step per time step, for $\hat{\mathbf{q}}$ and $-(T_b(t_k) - T_g)$. The solution vector $\hat{\mathbf{q}}$ is then used to compute the historical vector $\mathbf{h}_k$ for time step t_k through :

$$\mathbf{h}_k = \hat{\mathbf{g}}\hat{\mathbf{q}} \quad (4.34)$$

Once solved for all time steps, the time evolution of T_b allows for the calculation of the g-function through Eq. 4.35. It is important to mention that this method uses a value of $\tilde{q} = 1/H(W/m)$ to compute the $g_{j \rightarrow i}$. This has for consequence a small difference in the formulation of the g-function calculation method once the values of $\mathbf{T}_b$ are obtained. For the sequential method, the g-function is obtained by :

$$\tilde{g}(t_k) = 2\pi k_s m (T_b(t_k) - T_g)/q \quad (4.35)$$

4.4.2 Iterative Algorithm

The foundations of the iterative algorithm used in this work have first been presented in Pasquier et Marcotte (2013) and used later to simulate a GHE composed of boreholes of varying length, dip and strike angles (Pasquier *et al.*, 2013) and to simulate a GHE system in the scope of an optimization task (Hénault *et al.*, 2016). The algorithm is quite efficient at reconstructing $\tilde{g}$ for small GHEs, but its performances decrease rapidly with an increase of the number of boreholes and time steps. The fact that the algorithm uses a constant time step over the whole simulation duration allows a better sampling of the thermal perturbations caused by the surrounding heat sources present in the GHE. Consequently, the iterative algorithm is deemed more accurate and can be used to assess the accuracy of the block matrix formulation proposed in this work.

Using p to express the iteration index, the algorithm modifies the heat extraction rate of heat source j at time step k according to the following update formula :

$$q_{jk}^{p+1} = q_{jk}^p + \Delta q_{jk}^p \quad (4.36)$$

The m updated vectors q_{jk}^{p+1} are then used to construct vector $\hat{\mathbf{f}}^{p+1}$, which in turn allows computation of the borehole wall temperature $\mathbf{T}_{bj}^{p+1}$ of each source i . Using the matrix notation

of Eq. 4.23, the borehole wall temperature at iteration p is easily provided by :

$$\mathbf{T}_b^p = \hat{\mathbf{G}}\hat{\mathbf{f}}^p \quad (4.37)$$

The heat flux correction Δq is based on the difference between the temperature at the borehole wall of heat source j and the average borehole temperature of the m boreholes $\bar{T}_b$ present in the GHE :

$$\Delta q_{jk}^p = (\bar{T}_b^p - T_{bjk}^p) \rho_{jk} \quad (4.38)$$

To increase the convergence rate of the algorithm, the term within the parentheses are multiplied by ρ_{jk} , a time dependent and source-specific coefficient (see Pasquier et Marcotte (2013) for details). Once a convergence is reached, the algorithm ensures that the borehole wall temperature T_b of each borehole is equal to the average borehole temperature $\bar{T}_b$. The g-function is then obtained simply by :

$$\tilde{g}_k = 2\pi k_s \sum_{j=1}^m T_{bjk} / (q_{jk}) \quad (4.39)$$

4.4.3 Time stepping schemes

Three different algorithms are used in this work to construct g-functions. Each method having its own time stepping scheme, the following section presents succinctly these schemes to allow the interested reader to reproduce the results presented hereinafter.

The block matrix method uses a geometrical progression with a common ratio of 2 in its time steps which allows to cover all the simulation period in only a few time steps. Concretely, the first time step of the geometrical progression starts at 1 hour and doubles every two time steps until the end of the simulation. In a 40-year simulation, this leads to 38 unequal time steps. As a result of this implementation, the built-in convolution matrix allows to assess the temporal superposition accurately with unequal time steps.

A geometric progression scheme with a common ratio of 2 was selected for the block matrix method mainly owing to its ease of implementation and because our tests indicated that doubling the simulation time step every two time steps was a good trade-off between speed and accuracy for a variety of GHE layout. It is worth noting that the block matrix formulation is not limited to a given time stepping scheme and one can implement a linear scheme that accounts for every hour of a simulation if desired, or use a higher common ratio. However, using a linear scheme to construct the g-function can rapidly lead to memory management problems since this temporal discretization leads to an assembly of massive matrices.

The sequential method needs constant time steps. This is a direct consequence of using direct and inverse Fourier transforms to perform the convolution product. Because of this limitation, to cover the simulation duration and reduce the computation burden, the simulation length is split into sub-simulations according to the procedure recommended by Cimmino *et al.* (2013). The whole simulation is split into 23 time steps of 1 hour, followed by 29 time steps of 1 day, than 11 steps of 1 month, 9 steps of 1 year and finally 1 time step per decade until the end of the simulation. In a 40-year simulation, this leads to 46 time steps, which is comparable to the number of time steps obtained for the block matrix method.

The iterative algorithm modifies simultaneously the heat extraction rate of each borehole at every hour of the simulation before reaching convergence and therefore uses constant time steps of 1 hour. In a 40-year long simulation, the iterative algorithm uses 350,400 time steps. It would be possible to use longer time steps to improve performance of the iterative method. However, using short time steps allows to sample accurately the thermal interactions between the boreholes.

4.5 Validation Scenarios

In this section, the three scenarios developed to benchmark the block matrix formulation method are presented. For each scenario, a g-function will be calculated by the proposed method, as well as by the sequential and iterative methods. The time taken for the calculations of the g-functions, with and without the use of Chebyshev polynomials, will be measured.

Tableau 4.1 Thermal and geometrical parameters used for validation Scenario 1, 2 and 3.

Parameter	Symbol	Value	Unit
Buried depth	D	2	m
Soil thermal conductivity	k_s	2.5	$W/m^{\circ}C$
Borehole radius	r_b	0.08	m
Soil volumetric heat capacity	C_s	2	$MJ/m^3^{\circ}C$
Borehole length	H	150	m
Initial ground temperature	T_g	0	$^{\circ}C$

Each scenario presents a GHE which contains boreholes arranged in a given configuration. Apart from the borehole locations and the simulated period, the three scenarios share all the same thermal parameters, borehole length, radius and buried depth. Table 4.1 lists the

common parameters used throughout Validation Scenario 1, 2 and 3.

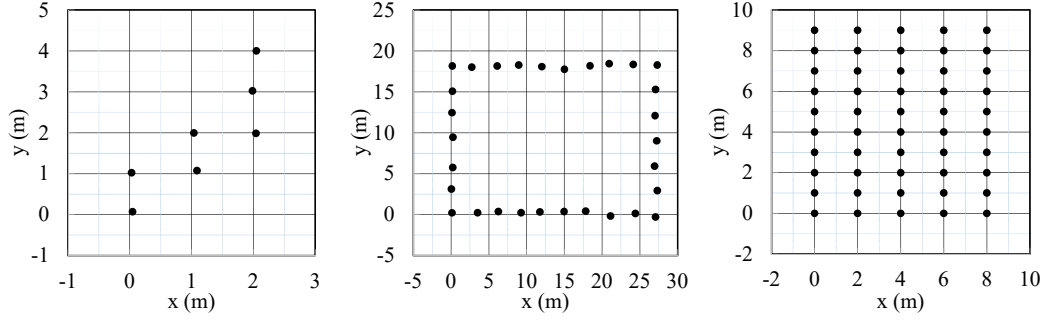


Figure 4.6 Configuration of the GHE in all scenarios (from left to right, 1 to 3).

The GHE of Scenario 1 contains seven 30 meters' boreholes (Fig. 4.6, left) distributed on an irregular grid across an area of 2 by 4 meters. The x and y coordinates of each boreholes are modified by a small random value in order to make each pair of boreholes spaced by a unique distance and to limit the benefits associated with a symmetry. The goal of Scenario 1 is to compare the late time g-function, which has reached its final plateau. This is meant to assess the differences between each method after a steady-state is almost reached in the field. To clearly assess the late behavior of the g-function, the simulation of Scenario 1 lasts 200 years.

In Scenario 2, the GHE contains 30 boreholes of 150 meters (Fig. 4.6, center) distributed randomly around a rectangle of 17.5 by 27.5 meters. The boreholes are arranged as if they were surrounding a building. Again, the x and y coordinates of each boreholes are modified by a small random value to reduce the symmetry in the GHE. The mean distance between two adjacent boreholes is 2.5 meters. This GHE being notably bigger than the one in Scenario 1, the simulation time had to be reduced to 40 years to allow the iterative method to converge in a reasonable time.

In the third scenario, a GHE of 50 boreholes of 150 meters located on a regular grid having a clear symmetry (Fig. 4.6, right) is simulated. The GHE spans across an area of 9 by 8 meters. The rows in this field are separated by 1 meter intervals while its columns are 2 meters apart from each other. This very small spacing between the boreholes was chosen to generate a steep g-function and a demanding test case for the tested algorithm. The simulation time is again maintained to 40 years.

Scenarios 1 and 2 are intended to compare the three computation methods in situations where no symmetry is present. This is necessary for an objective comparison since the single time step implementation of the block matrix method inherently takes advantage more efficiently

of the symmetry in the GHE than its sequential counterpart. The first two scenarios were then designed to not unduly favor the block matrix formulation. Note that all simulations were performed on a standard laptop computer (Intel i7, 2.6 GHz, 16 Go of RAM). Table 4.2 summarizes the three validation scenarios.

Tableau 4.2 Summary of the validation scenarios.

	Duration	Number of boreholes	Grid
Scenario 1	200 years	7	Irregular
Scenario 2	40 years	30	Irregular
Scenario 3	40 years	50	Regular

4.6 Results and discussion

The g-functions computed with the methods described in Section 4.3 and 4.4 for the validation scenarios presented in Section 4.5 are illustrated in Figure 4.7. In addition, Table 4.4 lists the computation times to obtain each g-function as well as the maximum relative error of the block matrix formulation and sequential algorithm with respect to the iterative method. Table 4.4 also presents the computation time obtained by constructing the g-functions using the sequential method with the help of Chebyshev polynomials to solve the finite-line source model as well as using the block matrix formulation without polynomials. This is done to show the direct contribution of the Chebyshev polynomials to the efficiency of the proposed method.

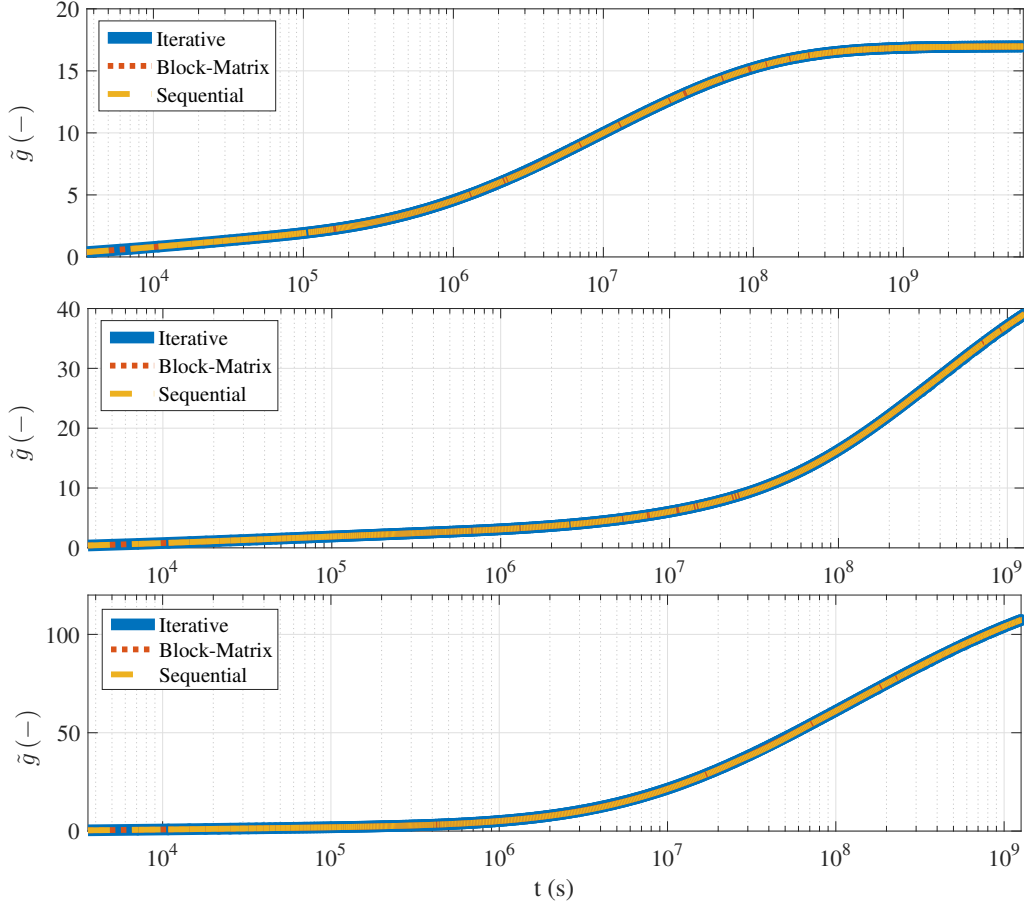


Figure 4.7 Transfer functions obtained for all three scenarios (from top to bottom, 1 to 3).

4.6.1 Computing efficiency

As shown in Table 4.4, the iterative method systematically presents higher computation times than the other two methods. When using the fastest computation times, the sequential method is between 12 to 31 times faster than the iterative approach while the block matrix formulation is between 13 and 145 times faster. This behavior was expected and is mainly due to the fact that the iterative method modifies the heat extraction rates on all times steps and boreholes simultaneously, which leads to a heavy memory and CPU use and lower efficiency.

For the small and irregular GHE of Scenario 1, the block matrix method (that includes the use of Chebyshev polynomials) is approximately 33% faster than the sequential method with a computation time of about 0.37 sec. For the larger but irregular GHE of Scenario 2, the sequential method calculated the g-function two times faster than the block matrix formulation with a running time of 1.18 sec. Finally, for the large and regular GHE of Scenario 3, the

block matrix formulation is 5 times faster than the sequential method with a computation time of 0.46 s (see the last row of Table 4.4). To put these results in perspective and highlight the efficiency of the block matrix formulation and sequential approach, computation times in the order of minutes were reported for similar GHEs (Cimmino *et al.*, 2013).

A breakdown of the computational effort has been done for validation of scenario 3 following the main computation steps detailed in Figure 4.5. For a total computation time of 0.46 s, 74.4 % of the total time is dedicated to the construction of Chebyshev polynomials, 5.6 % is used to construct the various matrices $\mathbf{G}$, 2.2 % is dedicated to the assembly of the linear system of equations, 9.9 % is used to solve the linear system and finally 5.8 % of the computation time is dedicated to the calculation of the g-function, which includes an interpolation using cubic splines to get a hourly g-function. The remaining 2.1 % of the total computation task is spread between a lot of small and fast subroutines such as preemptively creating empty vectors and matrices, importing and allocating the input parameters into the algorithm and finding the unique distances within $\mathbf{r}$. From this breakdown of the computational effort, it is clear that most of the computation is dedicated to the construction and integration of the Chebyshev polynomials. Reducing the computational burden associated to this specific step could reduce significantly the overall computation time.

These first results suggest using the matrix formulation for large regular GHE fields, the sequential method for large irregular fields, or either one of these techniques for small fields containing no symmetry. The number of redundancies in the distances of matrix $\mathbf{r}$ explains why the block matrix formulation is faster for Scenario 3. Indeed, when there is a redundancy in the components of matrix $\mathbf{r}$, some sub-matrices $\mathbf{G}_{j \rightarrow i}$ are computed only once and reused multiple times to fill the block matrix $\hat{\mathbf{G}}$ of Eq. 4.22 and 4.37 or the interaction matrix $\hat{\mathbf{g}}$ of Eq. 4.33. This reduces the number of times the solution of the FLS model (Eq. 4.1) is required and leads to a significant reduction in computation time. This redundancy benefits to all three methods, but the gain is much larger for the block matrix formulation, hence a faster computation time for Scenario 3.

Although the block matrix method has poorer performances than the sequential method in Scenario 2, the block matrix formulation is easier to implement and understand as it does not rely on spectral approaches.

4.6.2 Gain of using Chebyshev polynomials

The fastest computation times attributed to the proposed method in Scenario 1 and 3 can't be attributed only to the use of the block matrix formulation. Indeed, the use of Chebyshev polynomials to speed up the integration in Eq. 4.1 results in additional gains to construct

the g-functions. The computation times presented in Table 4.3 indicate that combined use of block matrix and Chebyshev polynomials results in a beneficial synergy in Scenario 1 and 3. Indeed, for these scenarios g-function computed with the block matrix formulation are obtained 4 to 5 times faster with than without the use of Chebyshev polynomials (Tables 4.3 and 4.4). For Scenario 2, the benefit of using Chebyshev polynomials is almost negligible and could explain the poorest performance of the proposed method. Surprisingly, the sequential method also benefits from the Chebyshev polynomials, but only in Scenario 3 with a degree of symmetry. This could be explained by the fact that the polynomials have a base cost in computation times that prevents them from being competitive on their own for simpler and irregular GHEs due to the loops it has to undertake.

Tableau 4.3 Computation times when using the sequential method with Chebyshev polynomials and the block matrix formulation without Chebyshev polynomials.

Scenario	Calculation method		Computation time
			(s)
1	Sequential	With polynomials	1.13
	Block matrix	Without polynomials	1.52
2	Sequential	With polynomials	10.79
	Block matrix	Without polynomials	2.75
3	Sequential	With polynomials	1.32
	Block matrix	Without polynomials	1.82

4.6.3 Accuracy and time stepping

Figure 4.8 shows the relative errors that exist between the sequential method, the block matrix formulation and the iterative algorithm during the simulation. The maximum relative errors of each scenario are summarized in Table 4.4. When considering the underlying assumptions of g-functions and their practical use, the accuracies reported in this work become negligible, even irrelevant. It is important to note that this accuracy analysis was done to understand the close link between the time stepping scheme used and the error obtained since using long time steps reduces simultaneously computation time and accuracy.

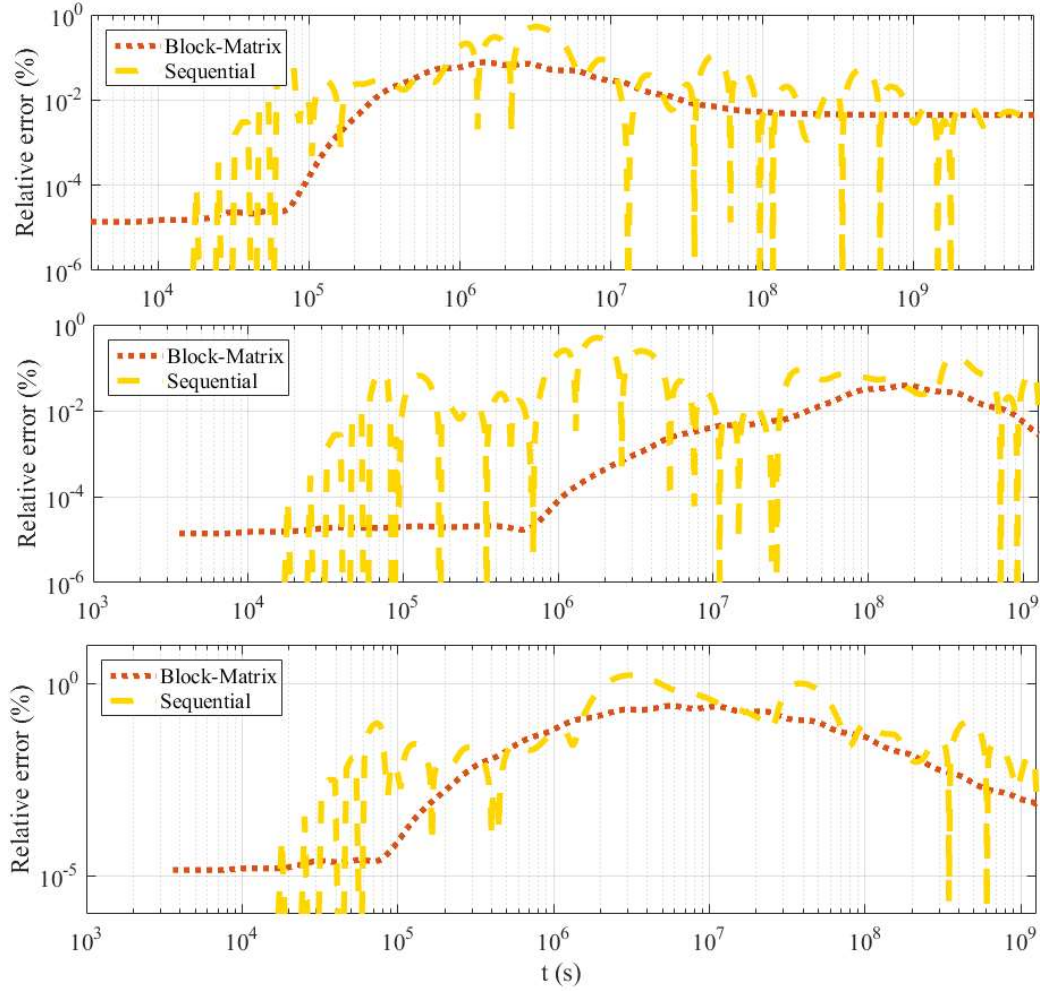


Figure 4.8 Relative errors committed during all three scenarios (from top to bottom, 1 to 3).

As shown in Table 4.4, the maximum error of the block matrix method (0.038 to 0.258 %) is always an order of magnitude lower than the sequential method (0.496 to 1.615 %). This is explained easily by looking at the time steps used by each method in Figure 4.8. The block matrix formulation uses a continuous geometric progression of its time steps that results in a continuous error that grows when the g -function (see Fig. 4.7) derivative changes. On the other hand, the error of the sequential method shows a cyclic pattern linked to the time steps used (see section 4.4.3). However, although the block matrix method is more accurate, both approaches can be used for practical applications.

Tableau 4.4 Computation times and maximum relative errors for Scenario 1, 2 and 3. The sequential method is used without Chebyshev polynomials while the block matrix formulation is used with Chebyshev polynomials.

Scenario	Method	Computation time (s)	Maximum relative error (%)
1	Iterative	6.96	-
	Sequential	0.55	0.525
	Block matrix	0.37	0.077
2	Iterative	31.36	-
	Sequential	1.18	0.496
	Block matrix	2.36	0.038
3	Iterative	67.04	-
	Sequential	2.13	1.615
	Block matrix	0.46	0.258

It is worth noting that from our experience, the accuracy and computation time of the block matrix formulation are mainly function of the temporal progression used, number of boreholes and grid layout. It was however difficult to assess precisely the influence of each variable separately. We however observed a higher accuracy and computation time with smaller time steps and a rather slow accuracy degradation when using 28 time steps instead of 38. For a fixed number of time steps, we observed no clear trend between the accuracy and the number of boreholes in a GHE. However, our tests indicated clearly that the computation time was more impacted by the number of unique distances within the field than by the number of boreholes, which is unsurprising since the number of unique distances modifies the number of matrices $\mathbf{G}$ constructed.

4.7 Conclusion

This article provided a novel simultaneous approach to calculate g-functions for ground heat exchangers with the use of a block matrix formulation. This approach, while being faster and more accurate than its sequential predecessor for complex and regular GHEs, is also more compact and easier to implement. Indeed, its implementation no longer requires the use of Fourier or Laplace transforms nor loops to account for time steps. These characteristics, its precision, speed and compactness, could be useful to design ground heat exchangers through optimization-based algorithms that require several thousand simulations. A method to ap-

proximate the FLS model using Chebyshev polynomials was also introduced. Although it suffers from a fixed cost in terms of computational time, this approach allows to speed up the g-function assessment even further when used jointly with the block matrix formulation on large GHEs. Used together, and because of the geometric time stepping in which the block matrix formulation is implemented in, they produced accurate g-functions when considering the base assumptions of the method. By using both strategies, constructing the g-function of a GHE composed of 50 regularly spaced boreholes can be achieved in less than a half-second while committing only a small relative error.

4.8 Acknowledgments

The authors wish to thank three anonymous reviewers who provided constructive and useful suggestions. This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the grant RGPIN-2014-05877.

4.9 Appendix A. Numerical example

The following section aims to present the reader with a small numerical example to help illustrate the methodology brought forth in the present article. The parameters used for this simulation are given in Table 4.5. Three boreholes located at $(0,0)$, $(0,1)$ and $(2,0)$ m are considered, which leads to the following distance matrix $\mathbf{r}$:

Tableau 4.5 Parameters used in the numerical example of Annex A

Parameter	Symbol	Value	Unit
Buried depth	D	2	m
Soil thermal conductivity	k_s	2.5	$W/m^\circ C$
Borehole radius	r_b	0.08	m
Soil volumetric heat capacity	C_s	2	$MJ/m^3{}^\circ C$
Borehole length	H	150	m
Number of boreholes	m	3	-
Initial ground temperature	T_g	0	$^\circ C$

$$\begin{bmatrix} r_{1 \rightarrow 1} & r_{1 \rightarrow 2} & r_{1 \rightarrow 3} \\ r_{2 \rightarrow 1} & r_{2 \rightarrow 2} & r_{2 \rightarrow 3} \\ r_{3 \rightarrow 1} & r_{3 \rightarrow 2} & r_{3 \rightarrow 3} \end{bmatrix} = \begin{bmatrix} 0.08 & 1 & 2 \\ 1 & 0.08 & 2.236 \\ 2 & 2.236 & 0.08 \end{bmatrix} \text{ m} \quad (4.40)$$

Computations are illustrated through a four time steps geometric progression with t equal to 1000, 2000, 4000 and 8000 hours. All the sub-matrices $\mathbf{G}_{j \rightarrow i}$ are assembled using Chebyshev polynomials to estimate the FLS model. For the sake of determining the g-function, each borehole is set to emit a linear heat flux of 1 W of heat per meter.

In this example, to calculate the transfer function of this GHE, one must resolve the linear system of equations of Eq. 4.22, which is made of several sub-matrices. Once all the expressions are evaluated with their respective numerical values, the sub-matrix $\mathbf{G}_{j \rightarrow i}$ in the system looks as following :

$$\begin{bmatrix}
 233.4 & 0 & 0 & 0 & 74.7 & 0 & 0 & 0 & 35.8 & 0 & 0 & 0 \\
 255 & 233.4 & 0 & 0 & 95.5 & 74.7 & 0 & 0 & 54.2 & 35.8 & 0 & 0 \\
 276.4 & 267.6 & 255 & 0 & 116.4 & 107.7 & 95.5 & 0 & 74 & 65.7 & 54.2 & 0 \\
 297.4 & 293.4 & 288.8 & 276.4 & 137.2 & 133.3 & 128.6 & 116.4 & 94.2 & 90.3 & 85.8 & 74 \\
 \hline
 74.7 & 0 & 0 & 0 & 233.4 & 0 & 0 & 0 & 30.3 & 0 & 0 & 0 \\
 95.5 & 74.7 & 0 & 0 & 255 & 233.4 & 0 & 0 & 48 & 30.3 & 0 & 0 \\
 116.4 & 107.7 & 95.5 & 0 & 276.4 & 267.6 & 255 & 0 & 67.4 & 59.2 & 48 & 0 \\
 137.2 & 133.3 & 128.6 & 116.4 & 297.4 & 293.4 & 288.8 & 276.4 & 87.4 & 83.5 & 79 & 67.4 \\
 \hline
 35.8 & 0 & 0 & 0 & 30.3 & 0 & 0 & 0 & 233.4 & 0 & 0 & 0 \\
 54.2 & 35.8 & 0 & 0 & 48 & 30.3 & 0 & 0 & 255 & 233.4 & 0 & 0 \\
 74 & 65.7 & 54.2 & 0 & 67.4 & 59.2 & 48 & 0 & 276.4 & 267.6 & 255 & 0 \\
 94.2 & 90.3 & 85.8 & 74 & 87.4 & 83.5 & 79 & 67.4 & 297.4 & 293.4 & 288.8 & 276.4
 \end{bmatrix} \times 10^{-3} \quad (4.41)$$

Note that the previous matrix must be completed with 4×4 identity matrices and a 4×4 matrix of zeros. For this four time steps example, the right-side vector of Eq. 4.22. looks as

follows :

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

(4.42)

Once the matrix system is assembled (Eq. 4.22), its solution gives a vector containing the incremental heat extraction rates emitted by each borehole and the time evolution of $-(T_b -$

T_g) through a vector assembled as follows :

$$\begin{bmatrix} f_1(t_1) \\ f_1(t_2) \\ f_1(t_3) \\ f_1(t_4) \\ \hline f_2(t_1) \\ f_2(t_2) \\ f_2(t_3) \\ f_2(t_4) \\ \hline f_3(t_1) \\ f_3(t_2) \\ f_3(t_3) \\ f_3(t_4) \\ \hline -(T_b(t_1) - T_g) \\ -(T_b(t_2) - T_g) \\ -(T_b(t_3) - T_g) \\ -(T_b(t_4) - T_g) \end{bmatrix} = \begin{bmatrix} 305.16 \\ -1.82 \\ -0.94 \\ -0.48 \\ \hline 318.19 \\ -0.12 \\ -0.01 \\ 0.00 \\ \hline 376.64 \\ 1.93 \\ 0.95 \\ 0.47 \\ \hline -108.50 \\ -128.27 \\ -148.69 \\ -169.24 \end{bmatrix} \times 10^{-3} \quad (4.43)$$

In this particular example, since the initial ground temperature is zero, the term T_g is neglected in Eq. 4.13 and the g-function of this GHE is easily obtained through :

$$\tilde{\mathbf{g}} = 2\pi \times 2.5 (W/m^{\circ}C) \times 3 (-) \times \begin{bmatrix} 108.50 \\ 128.27 \\ 148.69 \\ 169.24 \end{bmatrix} \times 10^{-3} (^{\circ}C)/1 (W/m) = \begin{bmatrix} 5.11 \\ 6.04 \\ 7.01 \\ 7.98 \end{bmatrix} (-) \quad (4.44)$$

Finally, Table 4.6 shows the values of the g-function according to both the time as well as the characteristic time.

Tableau 4.6 g-function presented according to the time and characteristic time of the simulation for the numerical example of Annex A.

Time (h)	$\ln(t/t_s)$ (-)	g-function (-)
1000	-14.5	5.11
2000	-13.8	6.04
4000	-13.1	7.01
8000	-12.4	7.98

Nomenclature

$\mathbf{\Lambda}$	: Constraint vector (-)
ρ	: Time dependent and source-specific coefficient for the iterative method ($W/m^\circ C$)
a, b	: Shifting coefficients of the Chebyshev polynomials (-)
c	: Coefficient of the Chebyshev polynomials (-)
C	: Volumetric specific heat ($MJ/m^3^\circ C$)
D	: Boreholes' buried-depth (m)
$f, \mathbf{f}$	: Incremental heat-flux function (W/m)
$g, \mathbf{g}$	: Finite-line source response under a unit heat load of $\tilde{q}$ ($m^\circ C/W$)
$\mathbf{G}$	: Lower-triangular matrix of $\mathbf{g}$ ($^\circ C$)
$\tilde{g}$	: g-function (-)
h	: Historical vector of temperature perturbations ($^\circ C$)
H	: Boreholes' length (m)
$\mathbf{I}$	: Identity matrix (-)
k	: Thermal conductivity ($W/m^\circ C$)
m	: Number of boreholes (-)
n	: Number of time steps (-)
q	: Heat load (W/m)
Δq	: Heat flux correction (W/m)
$r, \mathbf{r}$	: Distance or matrix of distances (m)
t	: Simulation time (s)
T	: Temperature ($^\circ C$)
T_c	: Chebyshev approximation of borehole wall temperature ($^\circ C$)

$\Delta T, \Delta \mathbf{T}$: Temperature perturbation or matrix of temperature perturbations ($^{\circ}C$)
 $\mathbf{T}$: First type Chebyshev polynomials (-)
 U : Second type Chebyshev polynomials (-)

List of subscripts

0 : Initial conditions
 b : Borehole
 g, s : Surrounding ground
 i : Receiving borehole
 j : Emitting borehole
 $j \rightarrow i$: Influence of borehole j on borehole i
 k : Time step
 m : Number of boreholes
 n : Number of time steps

List of acronyms

erf : Error function
 FLS : Finite-line source model
 GHE : Ground heat-exchanger

CHAPITRE 5 ARTICLE 2 : EFFICIENT G-FUNCTION APPROXIMATION WITH ARTIFICIAL NEURAL NETWORKS FOR A VARYING NUMBER OF BOREHOLES ON A REGULAR OR IRREGULAR LAYOUT

Bernard Dusseault^{1*}, Philippe Pasquier¹

Science and Technology for the Built Environment , Accepted author version posted online : 25 Jun 2019, Published online : 08 Aug 2019.

¹ Civil, Geological and Mining Department, Polytechnique Montreal, C.P. 6079 Succ. Centre-ville, Montreal H3C 3A7, Canada.

* Corresponding author : bernard.dusseault@polymtl.ca.

5.1 Abstract

A g-function links heat extraction rates to the mean borehole wall temperatures in geothermal heat exchangers. New computation methods are proposed continuously, becoming more accurate as they rise in complexity. In this work, it is shown how artificial neural networks can be assembled and trained to approximate g-functions rapidly and accurately for varying numbers of boreholes placed on a regular or irregular layout. This article focuses on geothermal heat exchangers made up of one to ten boreholes in a field of 30 by 30 m with various borehole geometries and soil thermal properties. After a training phase that used a database of 500,000 g-functions obtained with a block matrix formulation, a mean computation time of 11.18 ms was reached. This is 37% faster than using the block matrix formulation and represents a small fraction of the time required to simulate the fluid temperature over a 5-year period. The accuracy of the network against regular and irregular configurations of boreholes was evaluated and median errors of 0.18% and 0.39% were noted. The methodology introduced in this work is general and can be used together with any method allowing construction of g-functions.

5.2 Introduction

Geothermal heat exchangers (GHE) are useful infrastructures that allow heating and cooling of buildings at a fraction of the operating costs of conventional systems. The most common type is made of vertical boreholes through which a heat-carrying fluid flows. Calculating the

thermal response of a GHE to a time varying heating or cooling load can be accomplished using the elegant concept of g-function (Eskilson, 1987), a smooth and monotonic transfer function that links a heat extraction rate to the mean borehole wall temperatures. Such function varies according to the ground thermal properties and geometry of the borehole field and is different for each GHE. By convoluting a g-function with a time varying heat load signal, the evolution of the fluid temperature can be simulated and the operation of a ground-coupled heat pump system simulated. This process is useful for a variety of applications such as sizing of GHEs, interpretation of distributed thermal response tests and even model-based control of ground-source heat pump systems (Atam et Helsen, 2016a,b).

In its pioneering work, Eskilson (1987) used a numerical model to construct g-functions. Nowadays, numerical methods such as finite element or finite volume methods are generally much slower than analytical solutions, but have the advantage to provide accurate solutions for complex problems where an analytical solution does not exist. This is especially true for many phenomena affecting GHEs such as groundwater advection (Nguyen *et al.*, 2017), borehole capacity and geometry (Yavuzturk et Spitler, 1999; Yavuzturk *et al.*, 1999; Zanchini et Lazzari, 2014) or borehole wall boundary conditions (Priarone et Fossa, 2016).

For engineering tasks that require either iterative design or modeling of large and complex GHEs, analytical models are generally preferred to calculate g-functions due to their better effectiveness. The last decade has been marked by many improvements with the development, in the early 2000s, of the finite line source (FLS) model (Zeng *et al.*, 2002) and its subsequent improvements (Bandos *et al.*, 2009; Claesson et Javed, 2011; Diao *et al.*, 2004; Lamarche et Beauchamp, 2007). Indeed, versions of the FLS model now exist to account for inclined boreholes (Lazzarotto, 2016b,a) and groundwater flow (Molina-Giraldo *et al.*, 2011). Great steps forward were made to reduce the computational time associated with g-function construction and use either by adopting heat load aggregation schemes (Bernier *et al.*, 2004; Claesson et Javed, 2012; Liu, 2005), geometric sampling and spectral convolution (Marcotte et Pasquier, 2008) or by reformulating the problem using matrices (Cimmino *et al.*, 2013; Dusseault *et al.*, 2017).

Recent contributions are now proposing to segment a borehole by a series of FLS (Cimmino et Bernier, 2014; Lazzarotto, 2016b) or to take into account the connection layout between the boreholes (parallel, series or mixed arrangement) (Cimmino, 2015; Lamarche, 2017; Lazzarotto, 2014; Marcotte et Pasquier, 2014). The downside of these recent improvements is the increased complexity of the solution methods, which has a direct impact on their usefulness for a professional designer. Indeed, despite their increased accuracy, some of these approaches fall more into the category of numerical methods than easy to use analytical solutions readily

implemented into a spreadsheet.

In parallel, the ever-increasing computational capability of computers is leading to the democratization of artificial neural networks (ANNs). Such networks, whose primary objective is to emulate the learning capabilities of a biological brain, can be trained to perform different tasks. In the field of geothermal energy, ANNs have been used to predict heat pumps' COP (Esen *et al.*, 2008; Sun *et al.*, 2015), efficiency of a district geothermal system (Arat et Arslan, 2017), model a direct expansion system (Fannou *et al.*, 2014) and contribute to the control of district (Yabanova et Keeba, 2013) and hybrid (Gang *et al.*, 2014) systems.

A key feature of ANNs is their ability to approximate smooth and monotonic multivariable functions, such as g-functions. Recently, Pasquier *et al.* (2018) demonstrated the important gain of using ANNs to approximate short-term transfer functions. In their work, a thermal resistance and capacity model (Pasquier et Marcotte, 2012, 2014) was used to assemble a database of 15,000 short-term transfer functions. Distancing themselves from the traditional definition of g-functions, which focuses on borehole wall temperatures, the database of short-term g-functions was used to train a network designed to predict the borehole outlet temperature directly, that is the heat pump entering water temperature (EWT). The architecture of the ANN developed was made up of a 8 neurons input layer, two hidden layers of 100 neurons each and an output layer sampling the short-term g-function through 85 neurons. Pasquier *et al.* (2018) reported that the construction of a short-term g-function, as the one illustrated in Fig.1, was achieved in less than two milliseconds on a standard computer.

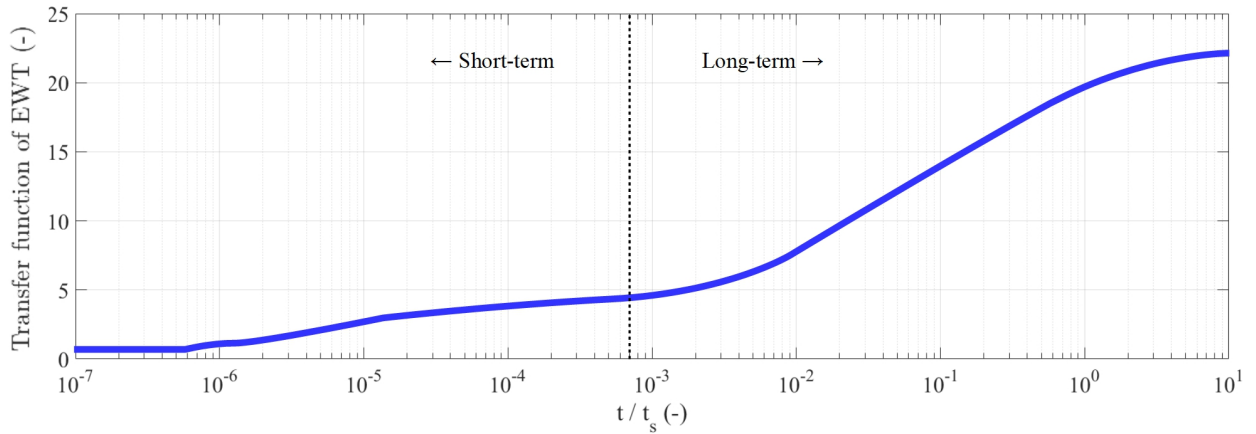


Figure 5.1 Short- and long-term transfer functions according to the characteristic time t_s designed to provide directly the EWT of a heat pump. The short-term transfer function is obtained by the ANN presented by Pasquier *et al.* (2018) while the long-term transfer function is obtained after correction of a g-function constructed by the block matrix formulation of Dusseault *et al.* (2017). Adapted from Pasquier *et al.* (2018).

ANNs have also been used by Dusseault et Pasquier (2018) to approximate long-term g-

functions, as the one illustrated in Fig. 5.1. In their paper, Dusseault et Pasquier (2018) proposed an ANN flexible enough to construct the g-function of GHEs composed of 1 to 10 boreholes placed on regular grids. A training database of 27,000 g-functions was used to train a network composed of an input layer of 8 neurons, three hidden layers of 50 neurons each and an output layer sampling the long-term g-functions using 40 neurons. Although limited to only 1 to 10 boreholes on a regular grid, the ANN had a computation time of a few milliseconds only and demonstrated good agreement with reference g-functions obtained with a proven reference method.

This paper aims to extend the initial work presented by Dusseault et Pasquier (2018) and to improve the generalization capabilities of ANNs for the construction of long-term g-functions for GHEs composed of 1 to 10 boreholes regularly and irregularly located on a field of 30 by 30 metres. In addition of this technical advancement, this work is also motivated by the desire to develop easy to implement and fast g-function generators. Indeed, once trained the weights and biases of a trained ANN can be published easily (for instance see Pasquier *et al.* (2018)) and integrated in a spreadsheet to construct rapidly a g-function without the need of a complex programming environment. Although reducing further the computing time of g-functions could seem worthless, optimization-based design methods (Hénault *et al.*, 2016; Pu *et al.*, 2017; Xia *et al.*, 2018) are promising, but are numerically demanding as they require thousands of g-function constructions. In such cases, even small time savings adds up to significant overall savings, justifying the need to propose more rapid g-function generators such the one proposed in this work.

This paper is structured as follows. The paper first presents the block matrix method used to efficiently assemble a large database of 500,000 g-functions. This collection of examples is then used to train the feedforward ANN described in Section 5.4. The results and discussion, presented in Section 5.5, support the parameterization used to code the GHE layout and show that the ANN developed can reproduce accurately and in a few milliseconds a set of reference g-functions.

5.3 Construction of a g-function with a block matrix formulation

To create a database of g-functions used to train the ANN, the block matrix formulation (BMF) presented by Dusseault *et al.* (2017) was selected mainly for its speed and efficiency. Indeed, this approach can construct a g-function for a GHE made of 50 boreholes over 40 years in only 0.46 second. This efficiency was a necessity to obtain, under a reasonable computation time, a large database of hundreds of thousands of g-functions. The core concept behind the block matrix formulation is to construct a linear system of equations accounting

simultaneously for the thermal interactions between all m boreholes in a GHE at n selected time steps. The thermal interactions between the boreholes (ΔT) are calculated using the finite line-source model of Claesson et Javed (2011). The latter, described by Eqs. 5.1 to 5.3, is evaluated with a unit heat load of $\tilde{q} = 1 \frac{W}{m}$.

$$\Delta T(r_{j \rightarrow i}, t_k) = \frac{\tilde{q}}{4\pi k_s} \int_{1/\sqrt{(4k_s t_k)/C_s}}^{\infty} \exp(-r_{j \rightarrow i}^2 s^2) \frac{Y(Hs, Ds)}{Hs^2} ds \quad (5.1)$$

where

$$Y(h, d) = 2 \operatorname{ierf}(h) + 2 \operatorname{ierf}(h + 2d) - \operatorname{ierf}(2h + 2d) - \operatorname{ierf}(2d) \quad (5.2)$$

and where

$$\operatorname{ierf}(X) = X \operatorname{erf}(X) - \frac{1}{\sqrt{\pi}} (1 - \exp(-X^2)) \quad (5.3)$$

To illustrate the block matrix formulation, first consider a GHE composed of only one borehole. To calculate its g -function, one would first have to calculate the mean borehole wall temperature T_b at each time step of the simulation. This can be done using Eq. (5.4) with $m = 1$ (one borehole) :

$$T_{b_i} = T_g + \sum_{j=1}^m \Delta T_{j \rightarrow i} \quad (5.4)$$

where j and i stand for emitting and receiving boreholes, which in this case is the same borehole, and where T_g is the initial ground temperature. All values of ΔT can be found simultaneously using a convolution product in the spectral domain (Marcotte et Pasquier, 2008; Pasquier et Marcotte, 2013). Convolution in the time domain can, however, be accomplished by a simple summation as described by the right-end part of Eq. (5.5) :

$$\Delta T(r_{1 \rightarrow 1}, q, t) = (\mathbf{f} * \mathbf{g})(t) = \sum_{k=1}^n f(t_k) g(t - t_{k-1}) \quad (5.5)$$

where f represents the incremental heat load emitted by a borehole throughout the simulation. In Eq. (5.5), the function $g(t - t_{k-1})$ is used to indicate the solution of the FLS model (Eq. 5.1-5.3) under a heat load over the duration specified in the parentheses.

For convenience, Eq. (5.5) can also be presented in a matrix form as shown by Eq. (5.6).

$$\begin{bmatrix} \Delta T(t_1, r_{1 \rightarrow 1}) \\ \Delta T(t_2, r_{1 \rightarrow 1}) \\ \Delta T(t_3, r_{1 \rightarrow 1}) \\ \vdots \\ \Delta T(t_n, r_{1 \rightarrow 1}) \end{bmatrix} = \begin{bmatrix} g(t_1 - t_0) & 0 & 0 & \dots & 0 \\ g(t_2 - t_0) & g(t_2 - t_1) & 0 & \dots & 0 \\ g(t_3 - t_0) & g(t_3 - t_1) & g(t_3 - t_2) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g(t_n - t_0) & g(t_n - t_1) & \dots & g(t_n - t_{n-2}) & g(t_n - t_{n-1}) \end{bmatrix} \begin{bmatrix} f(t_1) \\ f(t_2) \\ f(t_3) \\ \vdots \\ f(t_n) \end{bmatrix} \quad (5.6)$$

Using a more compact format and generalizing for any emitting (j) and receiving (i) borehole, the matrix formulation of Eq. (5.6) becomes :

$$\Delta \mathbf{T}_{i \rightarrow j} = \mathbf{G}_{j \rightarrow i} \mathbf{f}_j \quad (5.7)$$

where $\Delta \mathbf{T}$ is a vector containing the time evolution of the temperature perturbation, $\mathbf{G}$ is a lower triangular matrix and $\mathbf{f}$ is the vector of heat load increment.

To account for the thermal interactions between the m boreholes of a field, matrix $\mathbf{G}$ and vector $\mathbf{f}$ can be replicated $m \times m$ or $m \times 1$ times in a compact block matrix to integrate the spatial superpositions. The g-function is obtained by solving a linear system of equations under the constraint of a unit heat load expressed in matrix form by :

$$\mathbf{I} \mathbf{f} = \mathbf{\Lambda} = [1 \quad 0 \quad 0 \quad \dots \quad 0]' \quad (5.8)$$

where $\mathbf{I}$ is a $m \times m$ identity matrix. Adding the previous constraints to guarantee the uniqueness of the solution, the resulting linear system is given by :

$$\begin{bmatrix} \mathbf{G}_{1 \rightarrow 1} & \mathbf{G}_{2 \rightarrow 1} & \mathbf{G}_{3 \rightarrow 1} & \dots & \mathbf{G}_{m \rightarrow 1} & \mathbf{I} \\ \mathbf{G}_{1 \rightarrow 2} & \mathbf{G}_{2 \rightarrow 2} & \mathbf{G}_{3 \rightarrow 2} & \dots & \mathbf{G}_{m \rightarrow 2} & \mathbf{I} \\ \mathbf{G}_{1 \rightarrow 3} & \mathbf{G}_{2 \rightarrow 3} & \mathbf{G}_{3 \rightarrow 3} & \dots & \mathbf{G}_{m \rightarrow 3} & \mathbf{I} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{G}_{1 \rightarrow m} & \mathbf{G}_{2 \rightarrow m} & \mathbf{G}_{3 \rightarrow m} & \dots & \mathbf{G}_{m \rightarrow m} & \mathbf{I} \\ \mathbf{I} & \mathbf{I} & \mathbf{I} & \dots & \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \mathbf{f}_3 \\ \vdots \\ \mathbf{f}_m \\ -(\mathbf{T}_b - T_g) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{\Lambda} \end{bmatrix} \quad (5.9)$$

where $\mathbf{0}$ is either a null matrix or vector. Once assembled, Eq. (5.9) is solved for $\mathbf{T}_b$ and T_g

and the adimensional g-function is simply obtained with :

$$\tilde{g}(t) = 2\pi k_s m (\mathbf{T}_b - T_g) / \tilde{q} \quad (5.10)$$

5.4 Methodology

The objective of this work is to develop a single ANN having the ability to approximate a g-function over a period of 100 years for GHEs made of 1 to 10 boreholes regularly or irregularly spaced on a field of 30 m x 30 m. The methodological approach used relies on the universal approximation theorem (Cybenko, 1989; Hornik *et al.*, 1989), which states that a network with a single hidden layer that contains a finite number of neurons can approximate any continuous function on compact subsets, under mild assumptions on its activation function. A g-function being a smooth and monotonic multivariable function, it is well suited for approximation by ANNs.

In an ANN, each neuron (Rosenblatt, 1958) receives an input vector x from all the signals sent by the neurons of the previous layer and passes its own output y to all the neurons of the next layer as illustrated in Fig. 5.2.

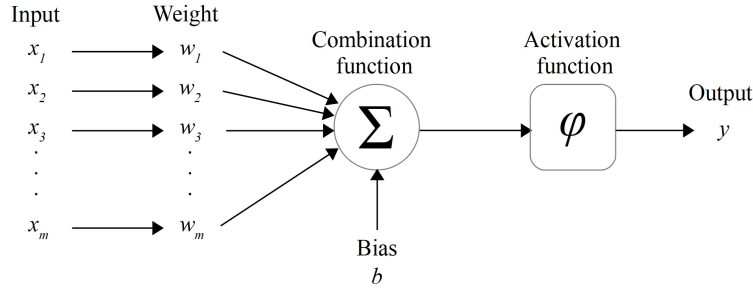


Figure 5.2 The mathematical model used for the creation of single artificial neuron.

Every time an artificial neuron receives some input vector x , it must decide, just like its biological counterpart, which output signal to transmit to the next layer. This is achieved by the relation expressed by Eq. (5.11) :

$$y_k = \phi \left(\sum_{j=1}^m w_{kj} x_j + b_k \right) \quad (5.11)$$

where k is the neuron index, x_j is the input provided by neuron j , w_k and b_k are the neuron' weights and bias respectively. The summation result is submitted to an activation function

ϕ to introduce an element of non-linearity between the input and output of the neuron. There are many types of activation function in the literature that are meant to be used for this specific task. As of today, the choice of an activation function is empirical at best and is usually driven by the experience of the user, by trial and error and by the type of phenomenon being modeled.

5.4.1 Architecture

The universal approximation theorem only requires a single hidden layer of a finite number of neurons to approximate any continuous function. However, the number of neurons in the hidden layer can grow rapidly for complex processes such as heat transfer in a GHE. Recent works indicate that use of ANNs having more than one hidden layer can help reduce the total number of neurons in an ANN. Consequently, after a few tests an ANN composed of two hidden layers made up of 200 neurons was selected as illustrated in Fig. 5.3. Furthermore, contrarily to the simpler feedforward neural network proposed by Dusseault et Pasquier (2018) to assemble g-functions for regular grid, an intertwined architecture was selected after testing as it offered a good trade-off between accuracy and training time. In the architecture retained, the input and output of the previous layers are all connected to the subsequent layers. Such architecture allows for more degrees of freedom and we found that the ANN can be trained to perform more complex tasks while having a smaller total number of neurons. For the hidden layers, the hyperbolic tangent sigmoid function was used as activation function due to its transitional nature. The identify function was chosen as the activation function for the output layer.

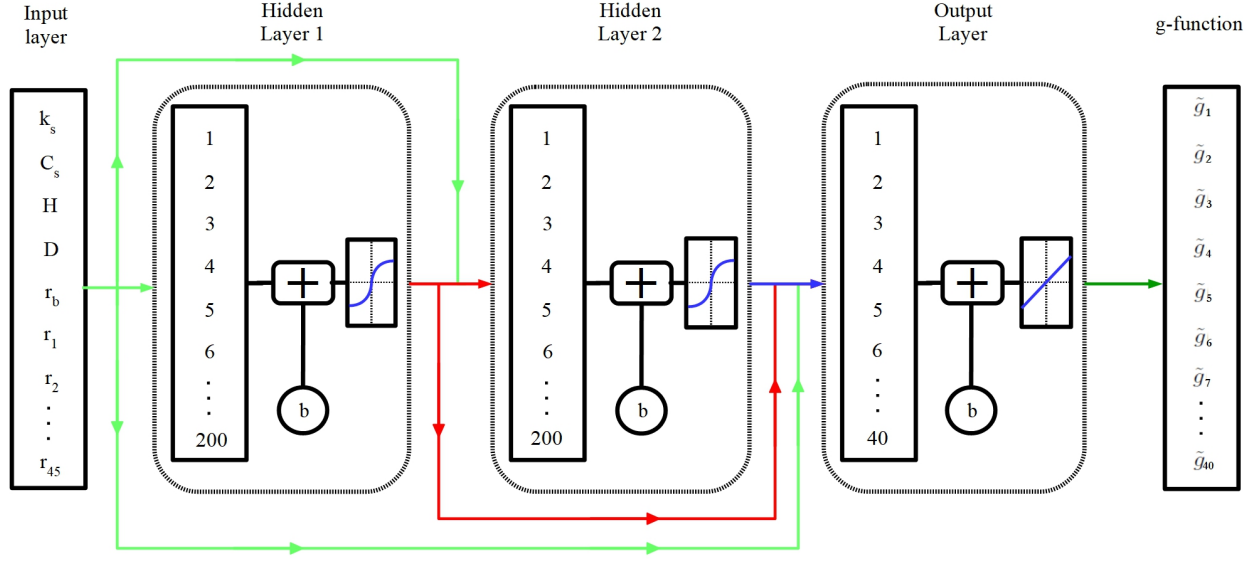


Figure 5.3 Proposed artificial neural network with 2 fully connected hidden layers used to approximate a g -function.

5.4.2 Inputs of the ANN

Input preparation is one of the key factors deciding if an ANN will be able to perform a certain task. The network needs to construct linear combinations of the values presented as input in order to learn to mimic the associated output. In that regard, the quantity of values in the input vector, their nature, the way they are ordered as well as their pre-processing have a decisive impact on the success or failure and training time of the ANN. As shown in Fig. 5.3, the variables used by the input vector x are the thermal conductivity k_s , the volumetric heat capacity C_s , the borehole length H , the buried depth D , the borehole radius r_b as well a vector $\mathbf{r}$ made up of 45 different distances between the boreholes. To properly train the ANN, all input parameters were centered and scaled between the interval $[-1; 1]$. This pre-processing prevents, for example, that some combinations of high inputs parameters will not saturate a neuron that has otherwise been trained to respond to small values. The possible values that these variables can take, before centering and scaling, are presented in Tab. 5.1 in Section 5.4.4.

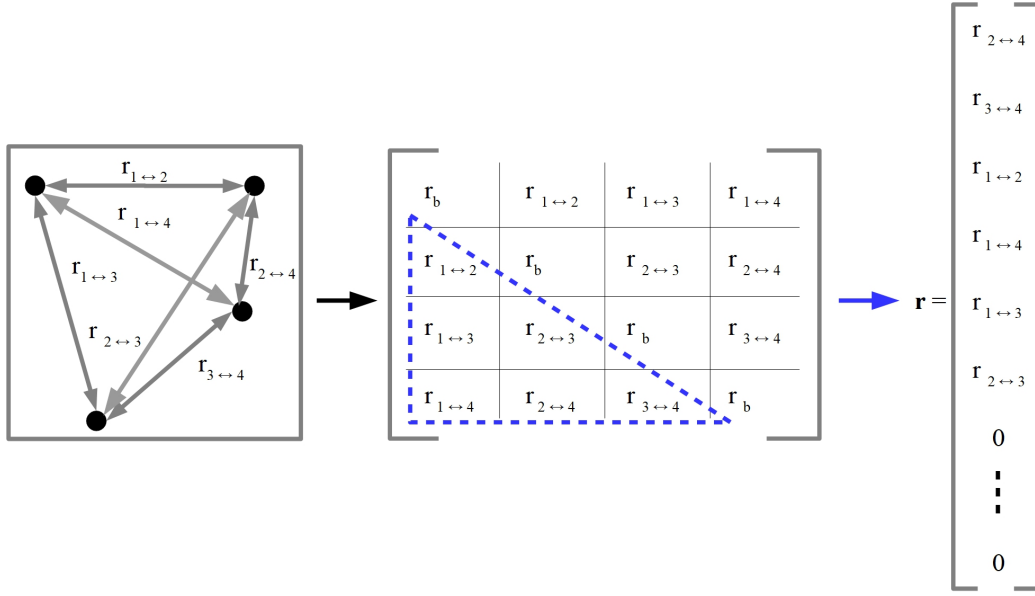


Figure 5.4 Illustration of how the distances within a GHE are extracted and formatted to construct input vector x .

The most difficult part of input formatting proved to be the construction of vector $\mathbf{r}$. Indeed, the method used previously by Dusseault et Pasquier (2018), based on the number of boreholes per row and column, and on the spacing between rows and columns, was deemed impractical to describe simultaneously the positioning of boreholes for both regular and irregular GHEs. After many unfruitful attempts based on Cartesian and polar coordinates, quite a simple and efficient method was found to assemble vector $\mathbf{r}$ from any borehole configuration. The approach is summarized in Fig. 5.4. The first step is to extract the distances between all boreholes within a field no matter the GHE' configuration. Then, the distances are collected within a 10 by 10 matrix of distances. Fig. 5.4 shows that the matrix diagonal (that represents the distances between the boreholes and themselves) serves only to collect the borehole radii. Since the radius is already communicated to the ANN through another input neuron, the central diagonal is of no use. Also, there are several redundancies in the distance matrix and using only the lower-triangular part of the matrix is enough to grasp the GHE configuration. The next step consists to place the distances of the lower triangular matrix in ascending order in a vector called $\mathbf{r}$.

In this work, the ANN constructed is trained to approximate the g-function of GHEs made of 1 to 10 boreholes. The maximum size of the lower triangular matrix, excluding its central diagonal, amounts 45 values (10 by 10, minus 10, divided by 2). By comparison, a GHE of only one borehole lead to an empty vector $\mathbf{r}$ (1 by 1, minus 1, divided by 2) and a strategy had to be identified to treat such situations as well. There were no clear guidelines in the

current literature as for how to handle the unused neurons when working with smaller GHEs. After some testing, it was found that the best way to present the distances to the ANN is to fill the remaining unused neurons with zeros at the bottom of the $\mathbf{r}$ vector. With this procedure, when approximating the g-function a GHE made of 8 boreholes for example, the vector $\mathbf{r}$ is composed of 28 ascending values of distances followed by 17 zeros.

5.4.3 Outputs of the ANN

The number of output neurons has a direct impact on the training time as well as the format in which the g-functions are produced by the ANN. To limit their number without losing information on the g-functions, the geometrical sampling scheme proposed by Marcotte et Pasquier (2008) was implemented. The progression that was selected starts at one hour and doubles every two time steps, allowing 100 years to be covered in only 40 steps. Note that the latter are regularly spaced on a logarithmic scale. The resulting output layer contains 40 neurons, each one corresponding to the g-function at a given time step. The hourly g-function can later be interpolated using cubic spline interpolation with a minor accuracy loss as found by Marcotte et Pasquier (2008).

$$y = \log_{10} \tilde{g} - \overline{\log_{10} \tilde{g}} \quad (5.12)$$

Before the training of the ANN, the mean values corresponding to each output neurons were subtracted from the decimal logarithm of the original database (Eq. 5.12). This procedure is done to reduce the span of possible values for each output neurons and simplify the ANN's training. Pasquier *et al.* (2018) observed that the first hidden layer of an ANN trained to approximate a short-term g-function was mostly returning the mean of the training database. Therefore, the mean of the log-transformed database was removed from the original database to assemble the final training set. In theory, removing the mean could get rid of a complete layer and lead to a simpler and more compact architecture. Before training, the resulting training data set was again centered and scaled between the interval $[-1; 1]$ to avoid neuron saturation.

5.4.4 Randomization and coverage of the training data

After some testing, it was observed that a training database containing only one type of configuration (regular or irregular) will always have trouble approximating the other type. To circumvent this problem, we found that it is important that the training database contains g-functions constructed using both types of GHE configurations. Also, because the ANN

needs to approximate accurately g-functions while subjected to varying values of ground thermal conductivities, volumetric heat capacities, borehole lengths, buried depths, number and configurations, it is important that the training database provides g-functions that span evenly across input range. The ranges of these parameters are presented in Tab. 5.1. Note that the range associated with vector $\mathbf{r}$ in Tab. 5.1 goes from 0, when the neuron is unused, to 42.43 meters. This value corresponds to the maximal distance possible within a square field of 30 by 30 meters (the square root of 1800).

Tableau 5.1 Ranges of input parameters of the ANN.

Parameters	Units	Lower bound	Upper bound
k_s	$\frac{W}{(m^oC)}$	1.0	3.5
C_s	$\frac{MJ}{(m^3oC)}$	1.0	3.5
H	m	50	250
D	m	0	2.5
r_b	m	0.05	0.1
$[r_1, r_2, r_3, \dots, r_{45}]$	m	0	42.43

Stratified random-sampling

When constructing the training database for regular GHEs, Dusseault et Pasquier (2018) used a method akin to stratified random-sampling to make sure every unique regular configuration was evenly represented amongst it, and consequently to prevent from overtraining certain configurations with respect to others. The approach used was as follows. Since 27 unique configurations and 1,000 g-functions were created for each configuration, a database made of 27,000 g-functions was assembled. During the generation process, the number of boreholes in each case, as well as their distribution between rows and columns, were imposed while the continuous parameters, including the spacing between rows and columns, were selected randomly each time.

In the present article, the same procedure is used and further extrapolated to the irregular cases. This is done by making sure the same number of g-functions is generated for every number of boreholes regardless of their configuration. For the regular configurations, the number of unique configurations for regular grids made of 1 to 10 boreholes was trimmed down from 27 to 15 by getting rid of mirror, equivalent configurations. For example, 3 boreholes can be placed on grids in two different ways, 1×3 and 3×1 , that mirror each other and therefore

only count for one unique configuration. On the other hand, 8 boreholes can be placed in four different ways, 1×8 , 2×4 , 4×2 and 8×1 , that only count for two unique configurations. These 15 different arrangement patterns then translate to an infinite number of possible GHE since any discrete values of spacing between rows and columns can be selected within a given field. On the other hand, irregular GHEs are not even limited by borehole configuration patterns since they can be placed literally everywhere on a given field of 30 m x 30 m.

Halton sequence and quasi-random sampling

When generating databases that contain lots of data placed in a high-dimensionality space, a poor coverage of the multidimensional space can sometimes become an important issue. Indeed, the use of random number generators based on sampling methods often leads to data clustering. In a context of clustering of the training data, this leads to poor generalization capabilities and low accuracy when dealing with inputs located away from the clusters. For discrete variables, such as the spacing between rows and columns or the distribution of boreholes in regular configurations, this problem can be avoided by simply listing all the possible unique combinations (see previous subsection). For the other variables in regular configurations as well as all variables in irregular configurations, this problem can be solved with the use of a quasi-random sampling methods such as the Halton sequence (Halton, 1964).

The Halton sequence is a method that is used in statistics to generate distributed sequences of points in spaces of high dimensionality. Although purely deterministic, the way these sequences are generated creates the illusion of randomness through their good distributions and coverage of the available space. The Halton sequence, along with other quasi-random number sequences, is often used alongside numerical methods such as Monte Carlo simulations. Even though we failed to identify a reference demonstrating the gain at using a Halton sequence, our tests showed that their use during the generation of training data can help significantly to reduce the training time by reducing the size of the training database itself.

The general principle behind the construction of Halton sequences is to start with as many coprime numbers as there are dimensions in a dataset (see example in Table 5.2). Each coprime number will serve as a base in its own dimension and will be used as a multiple of the denominators to assemble a series of fractions ranging from 0 to 1. Once all dimensions have their series of fractions based on their own coprime numbers, the coordinates of the sampling points are created by combining fractions from all these series to pinpoint to specific locations. These newly constructed coordinates, all ranging from 0 to 1, can then be modified by simple multiplications to fill the available real multidimensional space.

Tableau 5.2 Example of Halton coordinates.

Coprime number	Series of fraction	Halton coordinates	Corrected coordinates
2	$[\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, \frac{7}{8}, \frac{1}{16}, \frac{9}{16}, \dots]$	$(\frac{1}{2} \frac{1}{3}),$	$(5; 6.66),$
		$(\frac{1}{4} \frac{2}{3}),$	$(2.5; 13.33),$
		$(\frac{3}{4} \frac{1}{9}),$	$(7.5; 2.22),$
3	$[\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \frac{7}{9}, \frac{2}{9}, \frac{5}{9}, \frac{8}{9}, \frac{1}{27}, \dots]$	$(\frac{1}{8} \frac{4}{9}),$	$(1.25; 8.88),$
		$(\frac{5}{8} \frac{7}{9}),$	$(6.25; 15.55),$
		$\dots$	$\dots$

As a demonstration of how Halton sequences allow a good coverage of the available space, Fig. 5.5 shows how 250 boreholes are placed within a 50 by 50 meters field. The left-hand part of Figure 5.5 used the Halton sequence to distribute the boreholes while the right-hand part used a built-in random generator from MATLAB. It is clear by looking at it that the Halton sequence offers far better coverage of the available space and help prevent data clustering.

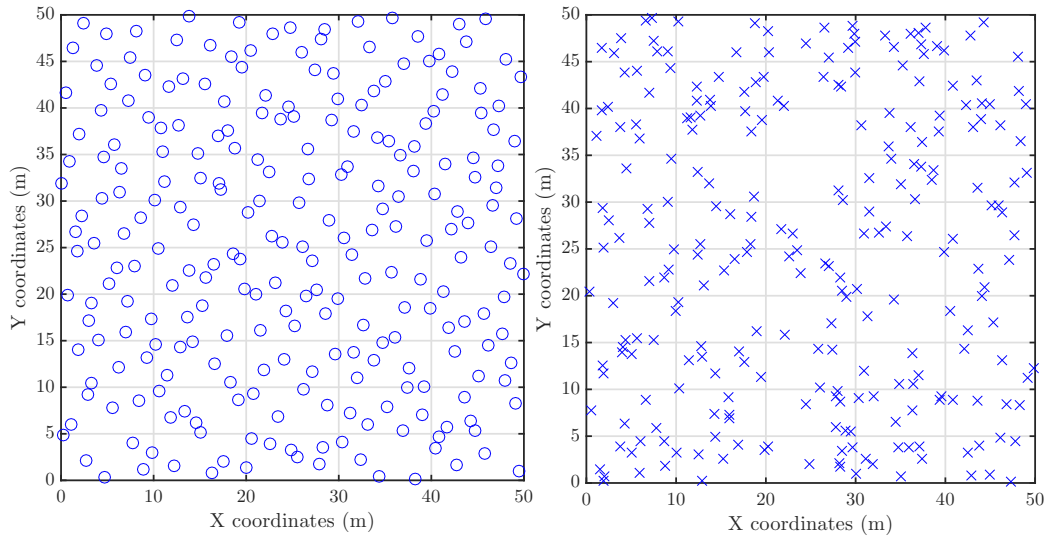


Figure 5.5 Comparison of 250 points distributions obtained with Halton sequence (left) and uniform number generator (right).

5.4.5 Size of the training database and training algorithm

As a rule of thumb, about 1,000 training examples are needed per parameter to train an ANN. In practice, the size of the database can vary widely from one problem to another and must be fine-tuned based on trials and errors or on the experience of the user. Dusseault et al. (2018) observed that when training an ANN for regularly gridded GHEs, the number of regular configurations is more of a driving force than the number of input parameters. In their work, 27,000 training g-functions were used, but concluded that a bigger database could potentially help mitigate some of the errors their ANN experienced.

In this work, the database used for the training of the ANN needs to include g-functions obtained both with regular and irregular configurations. A total of 10,000 g-functions were included in the database for each of the 15 possible unique regular configurations (for 1 to 10 boreholes). This amounts to 150,000 g-functions. As for the irregular GHEs, based on the rule of thumb introduced earlier and the 50 input neurons to the ANN, a logical choice would have been to go with 500,000 g-functions. However, the resulting database was believed to be too massive. Instead, after some testing, the number of g-functions generated for irregular GHEs was limited to 350,000. This amount proved to be enough for the ANN to reach a good accuracy. The combined database is therefore made up of 500,000 g-functions generated randomly to cover the ranges presented in Table 5.1.

The ANN was trained (Mathworks, 2015) using the scaled conjugate gradient with backpropagation (Hagan *et al.*, 1996) along with random initialization of the weights and biases of the ANN. The mean squared residuals (mse) was used as the objective function. Note that the training phase was accomplished on a Quaddro P6000 GPU produced by Nvidia. The final training time was around 70 hours.

5.5 Results and discussion

This section illustrates how the g-functions approximated by the ANN developed in the previous section are constructed rapidly and accurately with respect to reference g-functions produced by the BMF presented in Section 5.3. Note that to clearly assess the approximation capabilities of the ANN for regular and irregular GHEs, the results of both configurations are presented and discussed separately even if the same ANN was used to construct all the g-functions. It is important to note that the accuracy of an ANN is not verified using the same training data, as overfitting can sometimes exaggerate the accuracy of the network and lead to poor generalizations capabilities. Therefore, the trained ANN is validated using 5,000 independent g-functions generated after the training phase (2,500 for regular GHEs and 2,500

for irregular GHEs).

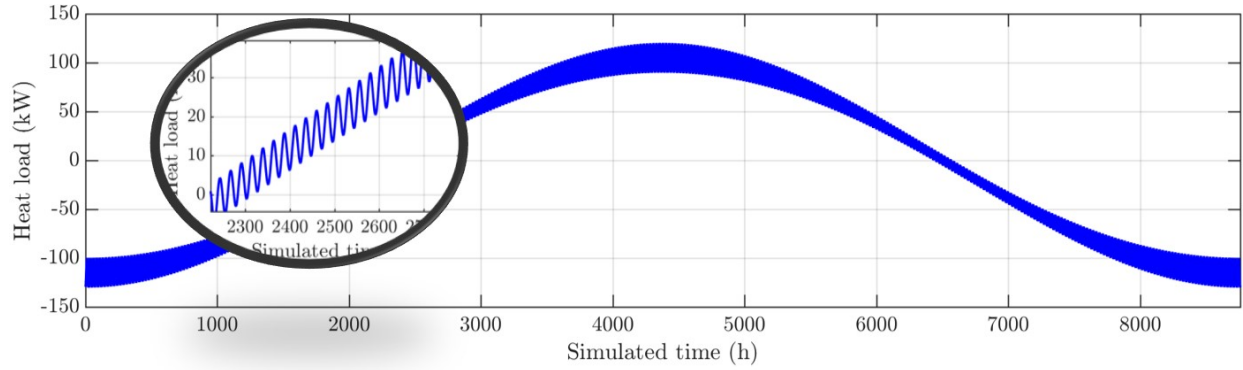


Figure 5.6 Synthetic hourly heat load signal presented for one year and used to assess the error made by the ANN on the fluid temperature. This signal is replicated five times to simulate over 5 years the heat carrying fluid temperatures of the Case Study.

It should be noted that to prevent poor approximations near the borders, the 5,000 independent g-functions were not drawn from the ranges introduced earlier in Table 5.1, but from the reduced ranges shown in Table 5.3. In this new table, all the ranges, except for the lower bound of $\mathbf{r}$ that remains 0, have been truncated by 5% at their lower and upper ends. Therefore, by shortening the possible span for input values by only 5%, the overall accuracy of the ANN is greatly improved as well as its overall reliability to produce consistent approximations.

Tableau 5.3 Reduced ranges of input parameters during the validation phase.

Parameters	Units	Lower bound	Upper bound
k_s	$\frac{W}{(m^{\circ}C)}$	1.125	3.375
C_s	$\frac{MJ}{(m^3{}^{\circ}C)}$	1.125	3.375
H	m	60	240
D	m	0.125	2.375
r_b	m	0.0525	0.0975
$[r_1, r_2, r_3, \dots, r_{45}]$	m	0	40.305

In addition, both the approximated and reference g-functions are convolved with a synthetic heat load signal to calculate hourly fluid temperatures over 5 years and assess the error made

by the ANN on fluid temperature. The yearly synthetic signal used over a 5-year period is illustrated in Fig. 5.6 and was generated following the generic equation proposed by Bernier *et al.* (2004). As mentioned in Section 5.4.3, the output of the ANN is a vector of 40 values that represent a g-function over 100 years using a geometric progression. Therefore, using only the 40 values of the g-function or only the mean computational time to assess the accuracy and rapidity is misleading. To be fair, the comparison must include the hourly cubic spline interpolation as well. Finally, to serve as a comparison tool, the mean time taken by the convolution product to construct a 100 years hourly heat carrying fluid temperatures signal is also calculated.

5.5.1 Validation of the ANN for regular GHEs

Figure 5.7 shows the g-functions computed by the ANN and the BMF for three of the 2,500 g-functions tested for the regular GHEs. Visually, the best case in blue, as well as the median case in green, seem like perfect fits when superimposed onto their reference g-function. Even the worst case (in red) looks like a fairly good fit up until the end of its first year (approximately 8500 hours) at which point it begins to diverge slightly from the BMF solution. For the worst case, the largest errors are in fact committed by the last neurons of the output vector (last time step). As illustrated in Fig. 5.8, the approximation error committed by the ANN is proportional to the number of boreholes in the GHE. Indeed, the lowest error is observed for a GHE composed of only one borehole, while the largest error is observed for a very compact field made of ten boreholes. This result is a good indication that use of ANNs to approximate the g-function of larger field could be challenging in the future and that other strategies could be envisioned to ease the construction of g-functions by ANN for large fields.

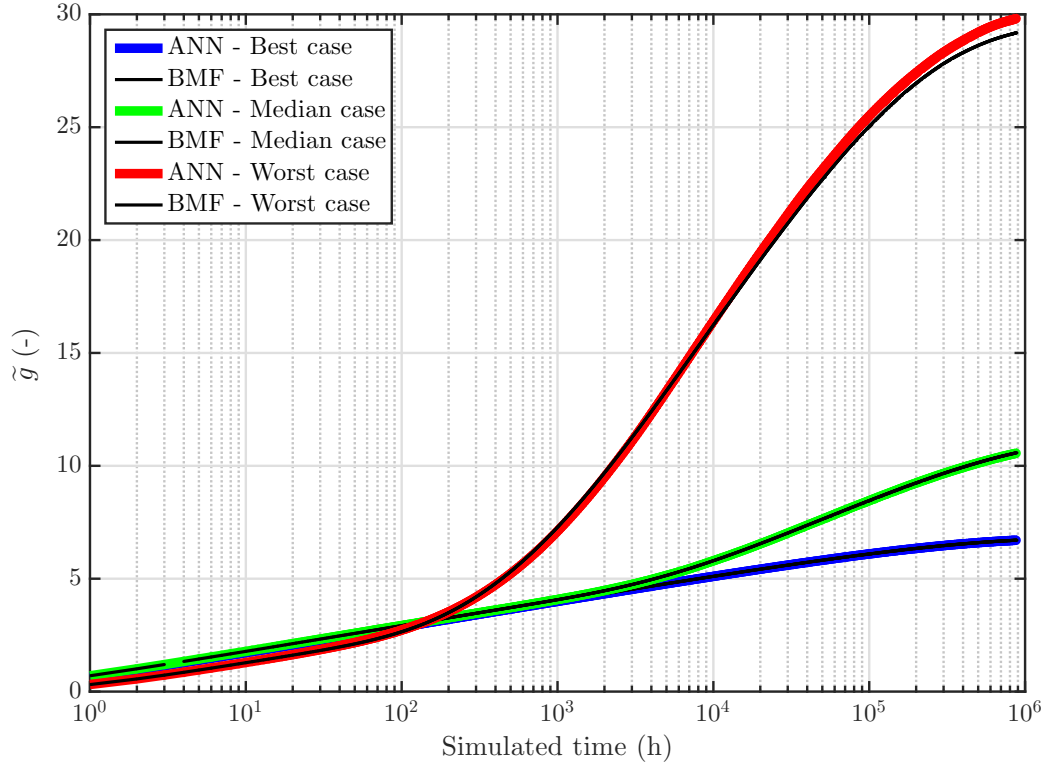


Figure 5.7 Visual comparison of the g-functions provided by the ANN and the BMF for the best, worst and median case amongst the 2500 regular.

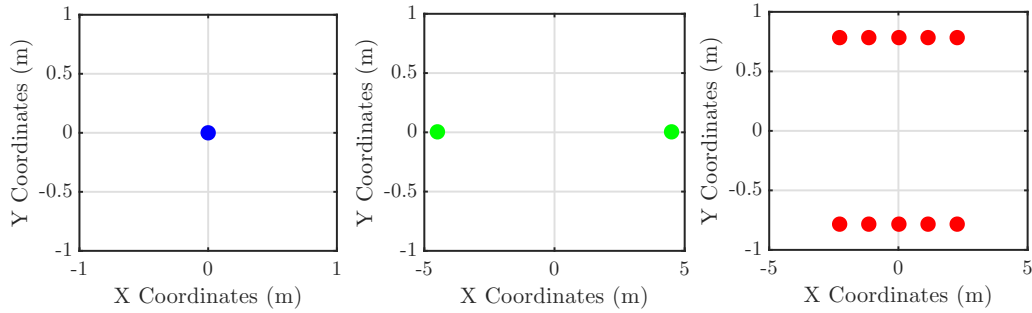


Figure 5.8 Geometry of the GHE for the best (blue), median (green) and worst (red) case.

Figure 5.9 shows the distributions of the maximum relative errors and maximum absolute errors committed over 100 years on the g-functions approximated by the ANN with respect to the one constructed by the block matrix formulation. The histograms present the maximum error for 2,500 borehole fields located on a regular grid. Figure 5.9 shows a clear pattern in which the vast majority of the g-functions were approximated with a maximum error far below 0.5%, with the approximations never deviating from the reference g-functions by more

than 0.2, no matter the number of boreholes in the field or the distance between the rows and column of the GHE. As a matter of fact, the worst case has a maximal inaccuracy around 2.2%. Meanwhile, the best case has an error of 0.035% and the median case has an error of only 0.18%.

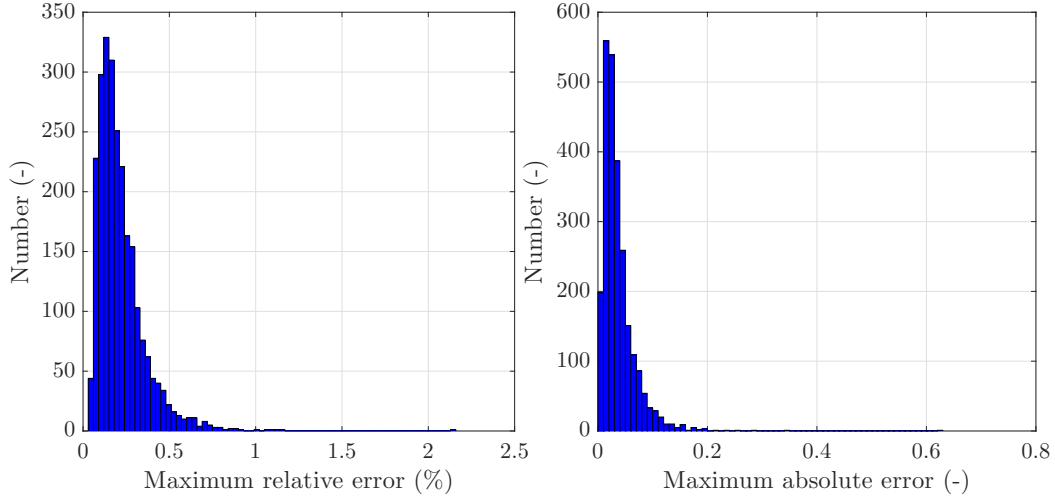


Figure 5.9 Distribution of the maximal absolute relative errors and maximal absolute errors between the g-functions provided by the ANN and the block matrix formulation for regular GHEs.

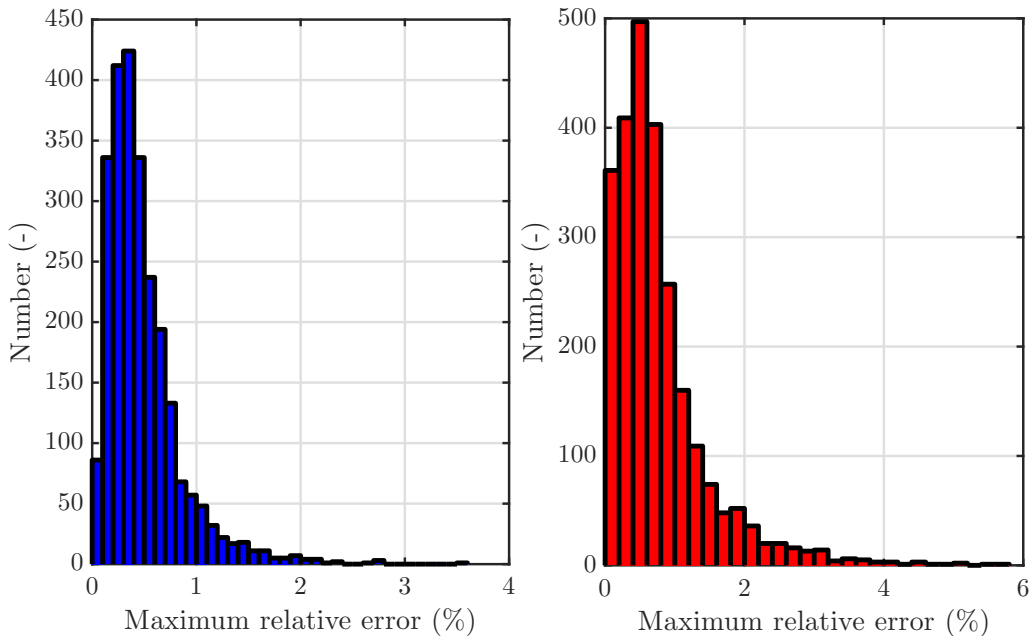


Figure 5.10 Distribution of the maximal relative errors between the fluid temperatures provided using the ANN approximated g-functions and the block matrix formulation provided g-functions over 100 years for regular (blue) and irregular (red) GHEs.

The maximum relative error on the fluid temperature is shown on the left-hand side (blue) in Fig. 5.10 and indicates that the vast majority of the maximum relative errors are still below 1%. However, errors on fluid temperature span across a wider spectrum of maximum relative errors than the g-functions. This result is opposite to that observed by Dusseault et Pasquier (2018) who also reported that their ANN produced g-functions that were less accurate for the first output neurons (broadly matching the first day on the g-function). This apparent contradiction simply stems from the time location of the approximation errors made on the g-function. Indeed, the convolution process propagates the errors made at the last output neurons (late times) while the errors made at the first output neurons cancel each other out due to the cyclic nature of the synthetic heat load signal used in both works.

5.5.2 Validation of the ANN for irregular GHEs

A comparison of the g-functions provided by the ANN and the BMF is illustrated in Fig. 5.11 for the best, median and worst cases for irregular GHEs. Again, the best and median case seem like perfect fits and are almost perfectly superimposed onto their reference g-function. The worst case in red fits well the BMF for the first four months (before approximately 1×10^7 seconds), but diverge slightly from the reference after that moment. Figure 5.12 illustrates the three irregular GHEs corresponding to the aforementioned best, median and worst cases. Note that this time the best case is not a GHE made of a single borehole, which indicates that the approximation errors made by the ANN do not simply increase with a higher number of boreholes for irregular fields. Our observations indicate that a small distance between two boreholes, which is more likely for irregular GHEs, is more detrimental to the approximation of a g-function than a higher number of boreholes.

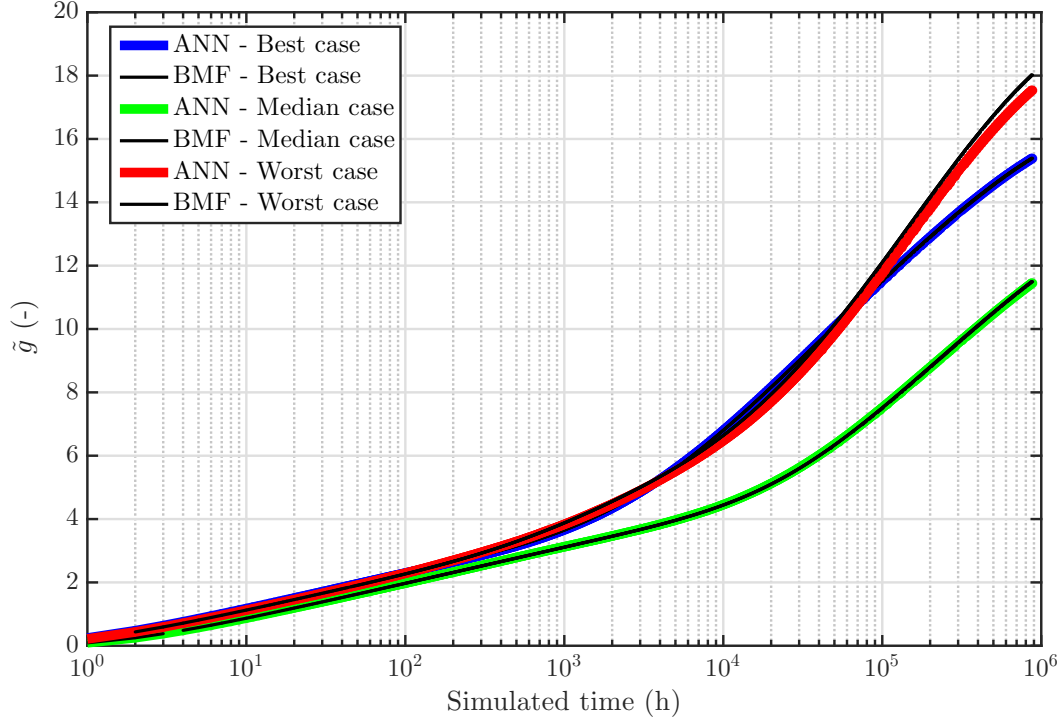


Figure 5.11 Visual comparison of the g-functions provided by the ANN and BMF for the best, worst and median cases amongst the 2500 irregular GHEs.

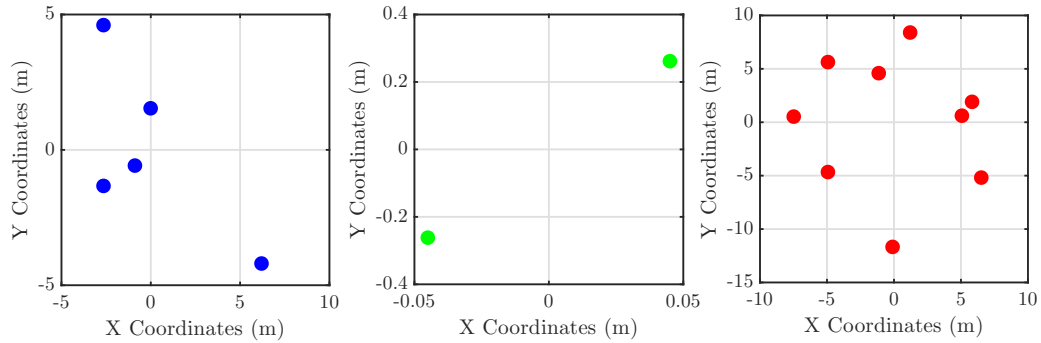


Figure 5.12 Configurations of the GHE on the best (blue), median (green) and worst (red) cases for irregularly placed boreholes.

The distributions of the maximal relative errors and absolute errors committed by the ANN for irregular GHEs are shown in Fig. 5.13. Comparing Fig. 5.13 and Fig. 5.9 we can note that the errors are distributed in a similar fashion, whether we are looking at regular or irregular configurations. However, the errors span on a wider range. Indeed, the maximum relative error for the worst case is around 2.7%, while the errors for the best case is 0.037% and for the median case around 0.39%. Although the best and worst cases are fairly similar for both

types of configurations, the median case is more than twice what it was for regular GHEs in the previous section. It is clear by looking at the right-hand side (red) of Fig. 5.9 that the phenomenon observed earlier for the error made on the heat carrying fluid temperatures also applies to irregular GHEs and reaches for some validation test almost 6%. However, most cases present an error inferior to 2%.

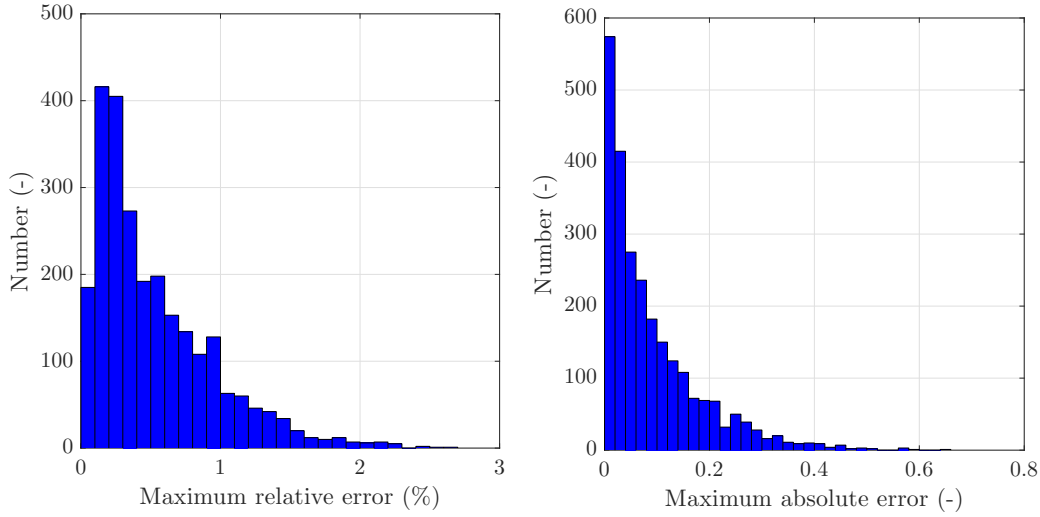


Figure 5.13 Distribution of the maximal absolute relative errors and maximal absolute errors between the g-functions provided by the ANN and the block matrix formulation for irregular GHEs.

5.5.3 Computation times

Once trained and fully validated, an interesting feature of ANNs is their computational efficiency. Table 5.4 lists the mean computation time (averaged over 5,000 test cases) of the ANN approximation, cubic spline interpolation and convolution in the spectral domain for a 100-year period on an hourly basis. The computation time devoted to the construction of a single g-function by the ANN presented in this work is 11.18 ms and represents a very small fraction (0.31%) of the overall simulation time. By comparison, the mean time required by the BMF to construct the 5,000 reference g-functions is around 15 ms, which is only 37% slower than the ANN proposed in this work. Although we tried to develop a more compact and fast network with good approximation and generalization capabilities, we narrowly beat the BMF, which is already very fast for small fields as the ones tested in this work. The gain could however be higher for larger GHEs and/or for more demanding g-function algorithms than the one used here.

Note in Table 5.4 that the convolution in the spectral domain by fast Fourier transforms (FFTs) is however almost 310 times longer than the g-function construction, which repre-

sents a spectacular gain given the recognized effectiveness of FFTs. g-functions can now be approximated so rapidly that most of the total computation time required to simulate the heat carrying fluid temperature is now used by the spectral convolution, an unthinkable situation until recently.

Tableau 5.4 Computation time for a 100-year simulation

	Time, ms
g-function construction by ANN	11.18
Cubic spline interpolation	176.72
Spectral convolution	3445.67
Total time	3633.57

The average ANN's approximation time, albeit extremely low, seems relatively high when compared to the average ANN's approximation time of 0.518 ms reported by Dusseault et Pasquier (2018) for a more specialized network. This can be explained by the fact that original ANN, because it focused only on the approximation of g-functions for regular GHE's, was much simpler : 198 neurons instead of 485 and 7,400 total connections between neurons instead of 78,000. The network proposed in this work is more general and can capture the essence of both regular and irregular GHEs. This is, however, achieved at the cost of higher network complexity and computation time.

5.6 Conclusion

In this article, the feasibility of using ANNs to approximate g-functions for GHEs composed of 1 to 10 boreholes randomly placed in a field was demonstrated. We covered the architecture of the network developed, how it was trained as well as its efficiency and accuracy by comparison to a reference method. A benchmark of 5,000 reference g-functions, not previously used for the training of the ANN, was assembled. Our results showed that, on average, the computation time devoted to the construction of a single g-function by the ANN presented in this work was 11.18 ms, a gain of 37% over the already fast block matrix formulation. This computation time now represents only a small fraction (0.31%) of the overall simulation time. By comparison, convolution in the spectral domain was almost 310 times longer, which represents a spectacular gain given the recognized effectiveness of spectral methods such as FFTs.

The largest errors, although very small, were always committed by the last neurons of the output vector which corresponds to the last time steps. For regular configurations, the median

error was 0.18% and was the smallest for GHEs made up of a single borehole. This result is a good indication that the use of ANNs to approximate the g-function of larger fields could be challenging in the future and that other strategies could be envisioned to ease their construction in such cases. For irregular configurations, the median estimation error was 0.39%. This time, our observations indicated that a small distance between two boreholes, which is more likely for irregular GHEs, was more detrimental to the approximation of a g-function than a high number of boreholes.

The methodology developed in this article shows the power and efficiency of ANNs for the approximation of a transfer function such as the g-function. Anyone can use it and easily adapt it to construct and train ANNs that suit their needs in terms of the number of boreholes and even include other training variables. In that regard, ANNs show great potential for demanding applications such as the design of ground coupled heat pump systems by optimization methods.

Acknowledgments

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Quaddro P6000 GPU used for this research. We also thank three anonymous reviewers who provided constructive comments during the revision of this work.

Block matrix Nomenclature

$\mathbf{\Lambda}$	: Constraint vector (-)
C	: Volumetric heat capacity ($MJ/m^3\ ^\circ C$)
D	: Buried-depth (m)
$f, \mathbf{f}$	: Incremental heat-flux function (W/m)
g	: Finite-line source solved with $\tilde{q}$ ($m^\circ C/W$)
$\mathbf{G}$	: Lower-triangular matrix of $\mathbf{g}$ ($^\circ C$)
$\tilde{g}$	: g-function (-)
H	: Borehole length (m)
$\mathbf{I}$	: Identity matrix (-)
k	: Thermal conductivity ($W/m^\circ C$)
m	: Number of boreholes (-)
n	: Number of time steps (-)
$\tilde{q}$	: Unit heat load ($1W/m$)
r	: Distance (m)
t	: Simulation time (s)
$\mathbf{T}$	: Temperature ($^\circ C$)
$\Delta T, \Delta \mathbf{T}$	: Temperature perturbation ($^\circ C$)

ANN nomenclature

Σ	: Combination function (-)
ϕ	: Activation function (-)
b	: Neuron's bias (-)
w	: Neuron's weight (-)
x	: Neuron's input (-)
y	: Neuron's output (-)

Subscripts

b	: Borehole
g, s	: Surrounding ground, soil
i	: Receiving borehole
j	: Emitting borehole
m	: Number of inputs in a neuron

CHAPITRE 6 ARTICLE 3 : USAGE OF THE NET PRESENT VALUE-AT-RISK TO DESIGN GROUND-COUPLED HEAT PUMP SYSTEMS UNDER UNCERTAIN SCENARIOS

Bernard Dusseault^{1*}, Philippe Pasquier¹

Submitted to Applied Energy : 19 June 2020.

¹ Civil, Geological and Mining Department, Polytechnique Montreal, C.P. 6079 Succ. Centre-ville, Montreal H3C 3A7, Canada.

* Corresponding author : bernard.dusseault@polymtl.ca.

6.1 abstract

Hybrid ground-coupled heat pump systems can efficiently heat and cool buildings by exchanging heat with underground geological materials. Their sizing is, however, complex and their financial profitability is hard to establish during the design phase due to uncertainties that taint important design parameters. Historically, impacts of those uncertainties are assessed a posteriori after sizing completion using sensitivity analyses. Unfortunately, these analyses cannot mitigate design risks, should the consequences of uncertainties be significant. Here, we show how the net present value-at-risk, a stochastic financial indicator inspired by a metric used in the financial sector, can weave the impacts of uncertainties throughout the design phase. The net present value-at-risk is compared to the net present value in a case study that considers uncertainty of the construction costs of the ground-coupled heat pump system, building's heat load, energy tariffs and ground thermal conductivity. The results show that this financial indicator, although it comes at a higher computational price than traditional net present value indicator, leads to shorter payback periods, greatly reduces the risks of unforeseen financial losses and does not require further sensitivity analysis. By describing uncertain parameters with statistical distributions during the sizing, the proposed designs are more conservative, still efficient and financially viable while being shielded from worst case scenarios.

6.2 Introduction

Geothermal heat pumps, when linked to ground heat exchangers (GHE), allow for efficient heating and cooling of buildings. Indeed, ground-coupled heat pump (GCHP) systems can

have coefficients of performance (COP) reaching three to four, which puts them amongst the most energy-efficient ways to heat and cool residential and commercial buildings (Omer, 2008). Such systems are also amongst the cleanest way to provide thermal comfort to a building's occupants when electricity is produced from a low-emitting source of greenhouse gases (Self *et al.*, 2013). Even when these systems are sized without detailed knowledge of the soil properties and installed in cold climate with a suboptimal mechanical room, Ozyurt et Ekinici (2011) show that their COPs can still reach between 2.65 and 3. Although their upfront costs can be quite significant compared to other heating and cooling systems, when correctly sized, GCHP systems can yield superior return on investment in the long run due to their lower operational costs (Hénault *et al.*, 2016).

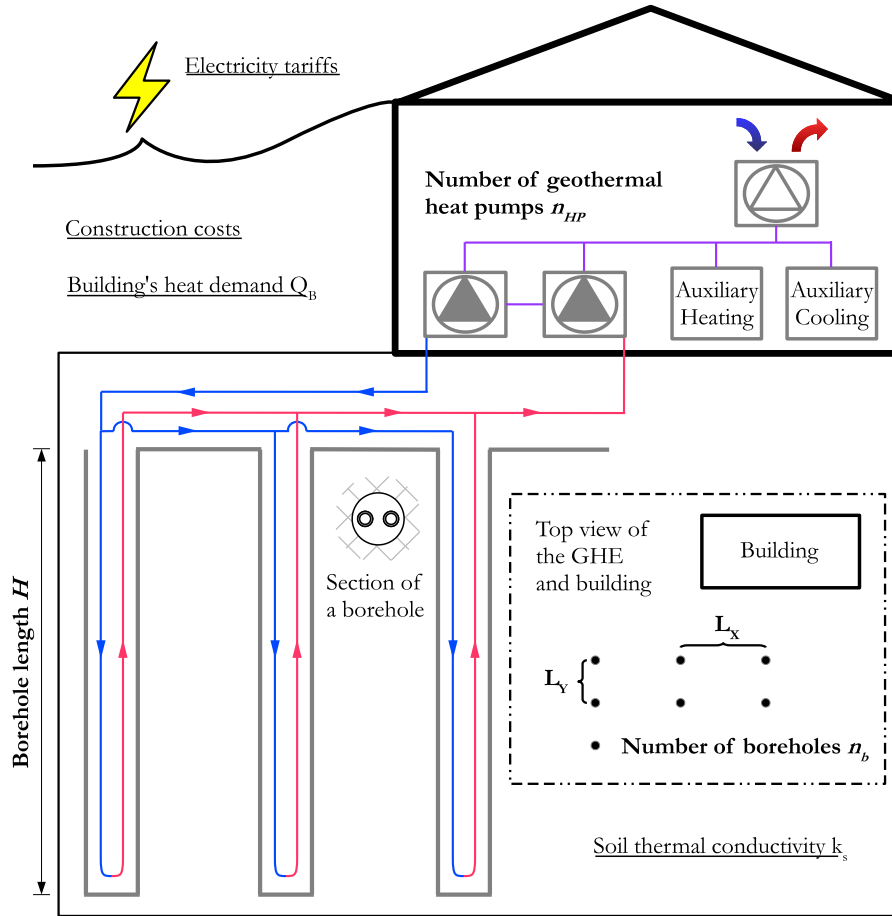


Figure 6.1 Main components of a simplified hybrid ground-coupled heat pump system where the five design parameters (H , n_b , n_{HP} , L_x , L_y) are in bold and the four uncertain parameters considered in this work are underlined.

To save on initial costs and provide redundancy, it is common practice to couple auxiliary heating and cooling devices to a GCHP system, thus forming a hybrid GCHP system as the

simplified system shown in Fig. 6.1. In such systems, the expensive GHE and GCHPs are designed to cover the building's base loads, while cheaper auxiliary systems are used to cover rare but demanding and expensive peak periods. For instance, it was shown that for a cooling dominated building, designing a GHE covering only the peak heating loads allows lowering of the initial investment (Hackel *et al.*, 2009). However, downsizing too much a GHE can potentially lead to longer payback periods (Nguyen *et al.*, 2014). Other more complicated approaches that exist to compensate for an undersized GHE include the combined use of thermal solar in heating dominated environment and the heating of domestic water as a precooling means in cooling dominated buildings (Yang *et al.*, 2010).

Designing a hybrid GCHP system is a complex task that brings into play dozens of parameters and quantities. Some parameters are connected to the building and its mechanical equipment (peak loads, heat pump capacity, etc.) or the GHE (borehole depth, number, etc.) and can be modified to some extent by a designer. Other important quantities, such as the thermal properties of the local geological materials, the hydrogeological context (Schelenz *et al.*, 2017) or the cost of the land (Tsagarakis, 2019), cannot be modified in practice but have a significant impact on the system's performance and profitability.

Historically, GHE's design has been achieved either by the ASHRAE sizing equation or by combining simulation and optimization algorithms. Several variants of the ASHRAE equation have been proposed (ASHRAE Handbook Ch.34, 2011; Bernier, 2006; Bernier *et al.*, 2008; Kavanaugh et Rafferty, 1997; Philippe *et al.*, 2010). These contributions differ by the way the thermal resistances and temperature penalties are computed, but the various sizing equations are all relatively easy to use and to implement. However, to establish accurately the operating cost of a specific GHE, one still has to simulate the time evolution of the GCHP system. It is why more complex design methodologies aiming at identifying the optimum of a cost or objective function have been proposed continuously in the last decade. The approaches are numerous, and their strong suits and weaknesses vary greatly. For example, sweeping the parameter space allows covering all possible cases and identify the global optimum of a given cost function. This approach was used by Ni *et al.* (2011) who used sweeping on a single variable to determine the optimal heating load that should be supplied by a gas boiler in combination to a GHE. Bayer *et al.* (2014) optimized directly both the heating and cooling loads represented by two variables. Alavy *et al.* (2013), in contrast, optimized the percentages of the peak loads coverage in heating and cooling. Hénault *et al.* (2016) used a sweeping on the number of geothermal heat pumps as part of a five variables, non-linear optimization algorithm to force this particular variable to behave like an integer.

Sweeping the parameter space is, however, not suited for the optimization of more than two

variables. This is because the numerical burden increases exponentially with the number of design variables. To take a more direct route and close-in more efficiently on the optimum, optimization methods relying on the gradients and Hessians of the cost function can be used to identify the best set of design variables. Amongst the most utilized algorithms in geothermal energy are the Newton (Newton, 1969) and quasi-Newton methods (Broyden, 1967), the steepest descent (Curry, 1944) and the interior trust region algorithm (Coleman et al., 1996) for non-linear minimization problems.

To search in high-dimensional space, heuristic and meta-heuristic approaches are more suited but are known to often propose suboptimal designs to complex optimization problems and may require lots of fine-tuning from the user (Pearl, 1984). On that subject, Das et al. (2011) did a great survey of the state-of-the-art in the field of differential-free optimization method applied to engineering problems. As an example of heuristic use in the design of plate-fin heat exchangers, a combination of a differential evolution algorithm and a free search allowed Ayala *et al.* (2016) to incorporate 9 design variables in a multi-objective optimization problem. Evolutionary algorithms have also been used for the design of vertical GCHP systems, but were found getting stuck more often on local optimum by Retkowski et al. (2014) depending on the guesses they used to initialize their search.

All aforementioned optimization strategies enable the deterministic design of hybrid GCHP systems, but none of them accounts for the fact that some design parameters are uncertain and should rather be described by a statistical distribution rather than a simple constant value. Indeed, uncertainties on key parameters can have significant impacts on the design proposed for geothermal systems (Daniilidis *et al.*, 2017). The effect of uncertainty is traditionally assessed either by ranking, screening or mapping (Pianosi *et al.*, 2016). Ranking aims to quantify the impacts and to sort the design variables according to their effect on the optimum of the cost function. This can be seen for instance in the work of Hadidi (2015) who disturbed alternatively each of his design variables and then assessed visually the impacts that resulted on the Pareto front for the design of plate fin heat exchangers using a biogeography based optimization algorithm. Zurmühl *et al.* (2019) used a similar procedure to observe the individual effect on the optimum of changes induced on three uncertain design parameters : the soil thermal conductivity, the cost of electricity and the cost of borehole drilling. Their study concluded that the cost of drilling has an impact that is 1.35 to 8 times greater on the optimum than the other two parameters. Finally, Escalante Soberanis *et al.* (2018), albeit applied to the design of flow batteries, also used ranking to compare the impacts of interest rates, capital costs and fuel prices. Screening is used to identify design variables that have small impacts on the solution so they can be left out of the optimization process. This operation is achieved to save computational time and has been used to limit the optimization

process to five design variables by Huang *et al.* (2015) (number, length, radius and spacing of boreholes and the outer diameter of the pipes). Fernandez *et al.* (2017) used three different methodologies to perform a screening : Morris (Morris, 1991), Sobol-Jansen (Jansen, 1999) and Bacco (O’Hagan, 2006). Based on all the results, they concluded that out of their 20 original design variables, only five of them (storage and filling thermal conductivity, fluid specific heat, initial storage temperature and flow) were truly required as they accounted for almost 90% of the output variance. Finally, mapping looks for regions in the search space that produce extreme or incoherent results so they can be avoided during the optimization process. Explicit examples of mapping were not found in the field of geothermal energies, as it is probably performed *de facto* in practice.

Ranking, screening and mapping are implemented using different methodologies (Pianosi *et al.*, 2015). Most technical contributions in geothermal energy implement these methods following one-at-a-time (OAT) schemes as opposed to all-at-a-time (AAT) ones. This means that sensitivity analyses rarely consider the combined effects at the optimum of modifying more than one design variable at a time. This is unfortunate because the complexity of energy systems often ends up linking uncorrelated variables together (Charitopoulos et Dua, 2017). One of the fastest way to obtain an AAT sensitivity analysis of a system is to resort to surrogate models. Any model that produces the same outputs as the original model at a fraction of the computational time by using simplifications, assumptions or algebraic relations that are calibrated using a small dataset qualifies as a surrogate model. A good example of this procedure can be found in Wen *et al.* (2016). They used a simpler kriging based model to assess the sensitivity of their three design variables instead of reusing their more complex multi-objective genetic algorithm. This holistic, simplified surrogate model was used to rapidly estimate *a posteriori* new optimums, thus evaluating the implications of a change in variables post-sizing.

Uncertainties issues are not confined to the realm of GCHP systems nor to engineering problems. However, other research fields tackle them quite differently and avoid using sensitivity analyses. In the financial sector, uncertainties are present in the form of uncertain yield on investments, bonds, shares or inflation to name only a few indicators. The conditional value-at-risk (CVaR) (Rockafellar et Uryasev, 1999) is an AAT stochastic method largely used in the financial sector that integrates uncertainty at the design stage. The CVaR aims to minimize the average losses of worst case scenarios instead of maximizing the average return on investment. The optimal portfolio is then assembled by a linear combination of individual stocks of which the average potential losses have been considered. Because it focuses on the distribution tails of uncertain parameters, this indicator proves to be safer for stocks with unknown, volatile or undefined history (Chikhaoui, 2015). Unfortunately, the CVaR cannot

be used directly to design GCHP systems because the design criteria, such as the net present value (NPV) of a system, cannot be assembled through a linear combination of its components. For instance, increasing the number of boreholes in a GHE by 10% does not necessarily provide a 10% increase of the NPV.

Traditional stochastic financial indicators have been adapted for such engineering projects. The net present value-at-risk (NPVaR) (Ye et Tiong, 2000), developed by Ye and Tiong for the financial evaluation of infrastructure projects, is a good illustration of such adaptation. As shown in Fig. 6.2, the NPVaR represents the average of the worst NPVs of a distribution obtained when considering the uncertainty of some input parameters. To prove its superiority against conventional deterministic financial indicators, Ye and Tiong designed a case study of two almost identical engineering projects and assessed their profitability using multiple financial indicators. They found that two plant projects can have similar NPVs, internal rates of return and payback periods, and still have considerably different NPVaRs. By selecting design variables that maximize the NPVaR, one minimizes considerably the risk that a project loses significant amount of money.

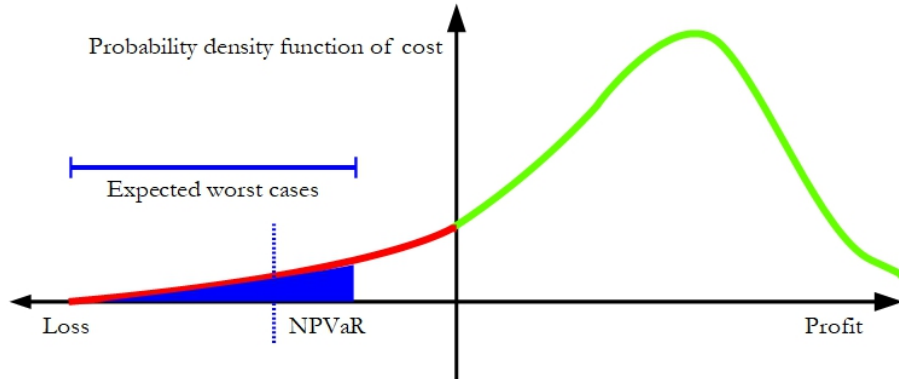


Figure 6.2 The net present value-at-risk (NPVaR) corresponds to the average of the worst fraction (α) of the outcomes of a NPV distribution composed of ω samples. The smaller the α value, the more cases that feature extreme losses are used to compute the NPVaR.

There is currently a clear gap of knowledge regarding the design of GCHP systems when one or more parameters are plagued with uncertainty. A methodology is thus needed to incorporate directly into the design the different uncertainties instead of simply quantifying their impacts afterward. Towards that end, we propose in this work to use the NPVaR, a stochastic financial indicator that minimizes risks of losing money, and adapt it for the design by optimization of hybrid GCHP systems. Four uncertain parameters are considered in this work (future cost of energy, GCHP system first cost, building heat loads and ground thermal conductivity) while five design variables (number of GCHPs, number and depth of boreholes,

horizontal and vertical spacing between the boreholes) are optimized taking into account the continuous and discrete nature of the variables.

The next section of this article presents the optimization and simulation strategies used to simulate efficiently the operation of a GCHP system. The second section also describes the elements considered in the computation of the NPVaR. The third section presents the case study used to demonstrate the utility of the NPVaR. The fourth section deals with the validation and results of the proposed methodology. A thorough analysis of the results follows in the fifth section. The sixth section presents a parametric study of the parameters that are relevant for the implementation of the NPVaR. Finally, the seventh section concludes the article.

6.3 Design-based Optimization Under Uncertainty

This section first introduces the spectral approach used for the fast simulation of an hybrid GCHP system. Then, it presents briefly the two methods used to construct the various g-functions used to simulate the GCHP system, namely an artificial neural network and a block matrix formulation. Then, it is shown how the NPVaR is used within a mixed-integer optimization algorithm to design GCHP systems under uncertainty. Finally, the sub-sampling scheme developed to speed up the computation of the NPVaR is presented.

6.3.1 Direct Simulation of Hybrid GCHP Systems

To simplify and speedup the simulation process, a spectral simulation method was used (Pasquier et Marcotte, 2013; Pasquier *et al.*, 2013). The approach used works under the assumption that all GCHPs and boreholes are fed in parallel and thus share the same inlet fluid temperature. Thus, the temporal evolution of the entering water temperature (*EWT*) on the source side of the n active heat pumps is given by :

$$EWT(t) = T_g + \frac{Q_G(t)}{n_b H} R_b + (f * g)(t) - \frac{Q_G(t)}{2\dot{V}C_f} \quad (6.1)$$

where t is the simulation time, T_g is the initial ground temperature, Q_G is the total ground thermal load, H is the borehole length, n_b is the total number of boreholes, R_b is the equivalent borehole thermal resistance, $\dot{V}$ is the total volumetric flow rate and C_f is the fluid volumetric heat capacity. Finally, g is the g-function of the GHE (Eskilson, 1987), f is the excitation function at $t \equiv t_j$ and $*$ stands for convolution performed here in the spectral domain. For a ground thermal conductivity k_s , the excitation f is a function of the ground load variation

between two successive time steps and is provided by :

$$f(t) = \frac{Q_G(t_j) - Q_G(t_{j-1})}{2\pi k_s n_b H} \quad (6.2)$$

In Eq. 6.1 and 6.2, Q_G corresponds to the thermal load extracted or transferred to the ground by the active GCHPs. If n heat pumps are operating at a coefficient of performance COP and at a capacity CAP , the total ground load is given by :

$$Q_G(t) = \pm n(t) \cdot CAP(t) \left(1 \pm \frac{1}{COP(t)} \right) \quad (6.3)$$

Note that in Eq. 6.3, the plus sign is used when the heat pumps are operated in cooling while the minus sign is used during heating mode. It is worth noting that Eqs. 6.1 to 6.3 integrate the main characteristics of the GHE (borehole depth, location and diameter ; ground and grout thermal properties ; pipe spacing, etc.) and GCHPs (performance and capacity curves).

Since the COP , CAP and number of active heat pumps n are either function of EWT or unknown, this results in a non-linear problem for which a solution is found iteratively (see Pasquier et Marcotte (2013); Pasquier *et al.* (2013) for details). The iterative process used to solve Eqs. 6.1 and 6.3 ensures that the building energy needs are covered by the GCHPs or the auxiliary system, that the EWT stays within the prescribed operating limits of the GCHPs, and that the number of active heat pumps does not exceed the number of installed GCHPs ($n \leq n_{HP}$). Finally, the approach allows to calculate which portion of a building's energy needs are covered by the n GCHPs and what must be provided by the auxiliary heating or cooling system. This also provides valuable information about the electricity consumption of the heat pumps and auxiliary systems and allows evaluation of the monthly energy bill. A simulation output is illustrated in Fig. 6.3 that presents n , EWT , COP and CAP , the hourly electrical power required to operate the heat pumps and the auxiliary devices and the cumulative monthly operating cost.

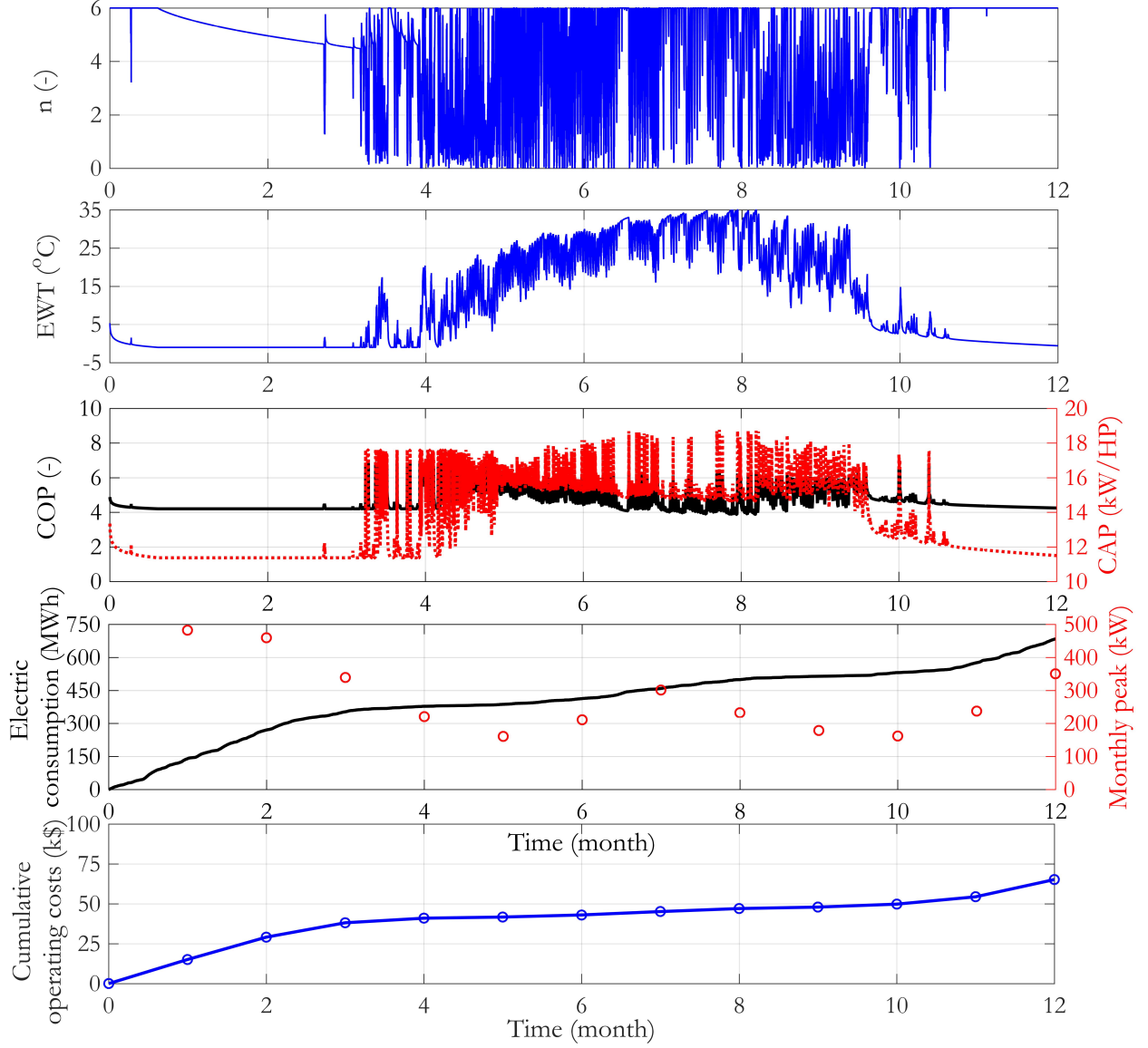


Figure 6.3 Hourly simulation output obtained with the method described in Section 2.1 for a case where the number of installed GCHPs is 6, the minimum temperature limit on the source side is -2°C and where the GHE is composed of 9 boreholes of 150 m. From top to bottom, the figures show the number of active heat pumps in operation, the EWT, the COP and CAP, the hourly cumulative consumption and peak monthly electric consumption and finally the cumulative monthly operating costs of the whole system.

6.3.2 Efficient Computation of g-functions

Design by optimization of GCHP systems requires a large amount of g-function evaluations. Indeed, for an optimization problem of just a few design parameters, tens of thousands of g-functions can easily be required, which represents a significant computation time (Hénault

et al., 2016). For efficiency purposes, we used two complementary approaches we developed recently for this task.

Artificial Neural Network

The first approach uses an artificial neural network designed and trained to approximate accurately the g-function of a GHE composed of 1 to 10 boreholes (Dusseault et Pasquier, 2018, 2019). This neural network is composed of one input layer of 45 neurons, three fully-connected hidden layers made up of 200 neurons and one output layer of 40 neurons. The first layer receives inputs such as ground thermal conductivity (k_s), ground volumetric heat capacity (C_s), borehole length (H), buried depth (D), as well as the distances between the boreholes of the GHE. The output layer provides the g-function at 40 specific times over a period of 100 years. A cubic spline interpolation is finally used to interpolate g at each simulation time step. It is worthy noting that the neural network was trained over the course of several weeks using a training database of 500,000 g-functions obtained using the method described below. For regular GHE layouts, the trained network can approximate a g-function with a median relative error of 0.18% and in less than 0.2 s, which is a significant gain by comparison to most published algorithms.

Block Matrix Formulation

The neural network described previously is limited to GHEs of 10 or less boreholes. For fields made of more than 10 boreholes, the block matrix formulation (Dusseault *et al.*, 2018) was used to obtain g . This method uses block matrices organized in the following system of linear equations :

$$\begin{bmatrix} \hat{\mathbf{G}} & \hat{\mathbf{I}} \\ \hat{\mathbf{I}}' & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ -(\mathbf{T}_b - T_g) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{0}} \\ \mathbf{\Lambda} \end{bmatrix} \quad (6.4)$$

where $\hat{\mathbf{I}}$ is a collection of identity matrices, $\mathbf{0}$ and $\hat{\mathbf{0}}$ are respectively a null matrix and vector, $\mathbf{\Lambda}$ is a constraint vector, $\hat{\mathbf{G}}$ is a collection of lower-triangular convolution matrices obtained using the finite line source model (Claesson et Javed, 2011) under a unit heat load $\tilde{q}$, $\hat{\mathbf{f}}$ is the vector of heat fluxes and $\mathbf{T}_b$ is a vector that captures the evolution of the average borehole wall temperature. This linear system is solved once for $\hat{\mathbf{f}}$ and $\mathbf{T}_b$ as it accounts simultaneously for temporal superposition and spatial interactions between the n_b boreholes of the field. The g-function is finally obtained using the following equation :

$$g = 2\pi k_s n_b (\mathbf{T}_b - T_g) / \tilde{q} \quad (6.5)$$

The computation time required by the block matrix formulation to construct a g-function for

a GHE of one to 50 boreholes is typically less than 0.5 s. Although slightly slower than the neural network, the block matrix formulation is still amongst the fastest techniques available. The compactness of both methods allow for an efficient implementation within a complex algorithm such as the one presented here.

6.3.3 Net Present Value-at-Risk

The computation of the NPVaR requires the calculation of a series of NPVs under various conditions of which only the worst cases are considered. Indeed, obtaining the NPVaR requires the sorting of all these cases and the calculation of the average of the worst fraction. Using the $< >$ symbol to represent the ascending sorting operation, ω to express the sample size and α to quantify the fraction of the worst outcomes, the *NPVaR* is given by Eq. 6.6. The impact of using both small and large values for α and ω are explored in an in-depth analysis in Section 6.8 that deals with the fine-tuning of the NPVaR related parameters.

$$NPVaR = \frac{1}{\lceil \alpha \cdot \omega \rceil} \sum_{j=1}^{\lceil \alpha \cdot \omega \rceil} < NPV >_j \quad (6.6)$$

For the calculation of a given NPVaR, all corresponding NPVs involve the same set of design variables. However, each individual NPV is computed with different uncertain parameters selected at random within their respective distributions. Furthermore, to deal in positive values, NPVs in this work express the difference between the NPV of a reference system that uses no geothermal heat exchanger to cover the building heating and cooling loads, and the NPV of a hybrid GCHP. This definition is captured by Eq. 6.7 :

$$NPV = \sum_{k=0}^n \frac{CF(k)}{(1+i)^k} \quad (6.7)$$

where i is the discount rate and CF is the cash-flow that arises from the savings generated by operating the hybrid system instead of the reference system during month k . Using this definition, by optimizing the NPV, what is really done is maximizing the monetary savings stemming from using the hybrid system, not the amount of energy generated by it. Therefore, these savings need to be calculated using the power consumption of the reference and hybrid systems and the local pricing of electricity (energy and power premium). Those considered in this work as well as all the other costs are introduced in Table 6.1 alongside the case study later on.

The first term in the summation of Eq. 6.7, when $k = 0$, represents the unavoidable extra

costs of building a costlier hybrid system instead of the all-electrical reference system. It can be expanded using the following equation :

$$CF(k=0) = -CF(k=0)_{GHE} - CF(k=0)_{Mec} + CF(k=0)_{Ref} \quad (6.8)$$

In the previous equation, $CF(k=0)_{GHE}$ involves all costs related to the construction of the GHE of the hybrid system and is given by

$$CF(k=0)_{GHE} = n_b [H(c_1 + 1.816c_4) + (c_2 + c_3 + c_7 + L_Y c_5 c_6)] + L_X c_5 c_6 \lceil n_b/5 \rceil + c_8 \quad (6.9)$$

Equation 6.9 integrates 8 costs, c_1 to c_8 , which are respectively the cost of : borehole drilling per meter, geothermal vault per borehole, borehole commissioning, heat-carrying fluid per liter, trenches excavation by volume and width, GHE commissioning and of thermal response test. It considers that the GHE is constructed in successive rows that never exceed 5 boreholes, hence the $\lceil n_b/5 \rceil$. Finally, it considers u-tube pipes of 3.4 centimeters in diameter in all boreholes, whence the 1.816 which corresponds to the number of liters of heat carrying fluid that fits per meter of borehole.

$CF(k=0)_{Mec}$ involves all costs related to the mechanical parts of the hybrid system and is given by :

$$CF(k=0)_{Mec} = (c_9 + c_{10}) \max(|Q_G|) + c_{11} \max(Q_{Aux_{heat}}) + |c_{12} \min(Q_{Aux_{cool}})| \quad (6.10)$$

It makes use of c_9 to c_{12} which corresponds to the cost of the geothermal heat pump per kW, additional equipment per kW associated with the mechanical room, electric heating and electric cooling both per kW of auxiliary equipment. The final costs depend ultimately on the maximum heating and cooling power supplied by the auxiliary equipment ($\max(Q_{Aux})$) and by the geothermal heat pumps ($\max(|Q_G|)$) throughout their yearly operations.

Finally, $CF(k=0)_{Ref}$ considers the costs of the electric heating and cooling machines of the reference system. To that end, it uses the same costs, c_{11} and c_{12} , that were used for the auxiliary equipment previously. That way, neither reference nor auxiliary gets unduly favoured during sizing solely based on their costs. Again, the final costs for the reference equipment scale with maximum heating and cooling powers that are required by the reference system ($\max(|Q_B|)$) throughout its yearly operation by :

$$CF(k=0)_{Ref} = c_{11} \cdot \max(|Q_{B_{heat}}|) + |c_{12} \cdot \min(|Q_{B_{cool}}|)| \quad (6.11)$$

where Q_B is the hourly energy needs of a building. For the case study introduced later on in this work, Q_B is depicted in Fig. 6.5.

All other terms in the summation of Eq. 6.7, where $k > 0$, represent the monetary savings which stem from operating a hybrid system instead of the all-electrical reference system. It can be expanded using the following equation :

$$CF(k > 0) = -CF(k > 0)_{Mec} + CF(k > 0)_{Ref} \quad (6.12)$$

Again, Eq. 6.12 can be divided into two equations : one for the term $CF(k > 0)_{Mec}$ which represents the operating costs of the heat pumps and auxiliary equipment and one for $CF(k > 0)_{Ref}$ which represents the operating costs of the reference system. These equations are

$$CF(k > 0)_{Mec} = \begin{cases} (c_{13} \oplus c_{14}) \sum_{j=1}^{730} (Q_{HP}(i) + Q_{Aux}(i)) + c_{15} \max(Q_{HP} + Q_{Aux}) & \text{if } k \in \text{summer} \\ (c_{13} \oplus c_{14}) \sum_{j=1}^{730} (Q_{HP}(i) + Q_{Aux}(i)) + c_{16} \max(Q_{HP} + Q_{Aux}) & \text{if } k \in \text{winter} \end{cases} \quad (6.13)$$

and

$$CF(k > 0)_{Ref} = \begin{cases} (c_{13} \oplus c_{14}) \sum_{j=1}^{730} (Q_{Ref}(i)) + c_{15} \max(Q_{Ref}) & \text{if } k \in \text{summer} \\ (c_{13} \oplus c_{14}) \sum_{j=1}^{730} (Q_{Ref}(i)) + c_{16} \max(Q_{Ref}) & \text{if } k \in \text{winter} \end{cases} \quad (6.14)$$

Both Eq. 6.13 and 6.14 follow the same structure. They use either a base price per kWh for the first 1200 kWh per month (c_{13}) or an inflated price for the excessive consumption (c_{14}) that gets multiplied on a pro rata basis with the total monthly consumption of electricity. Additionally, a summer (c_{15}) or a winter (c_{16}) premium per kW gets multiplied by the system's peak power consumption during the month to penalize power demand above 50 kW. This corresponds to the DP electricity rates charged by Hydro-Quebec for domestic uses where demand exceeds 50 kW every month.

6.3.4 Optimization of a Mixed-Integer Programming Problem

In this work, five design variables are modified to optimize the NPVaR. Three variables are continuous (L_x , L_y and H) while two others are discrete design variables (n_b and n_{HP}), leading to a mixed-integer non-linear programming (MINLP) problem. A solution is obtained in this work by combining a constrained nonlinear optimization algorithm for the continuous variables, a branch-and-bound algorithm (Lawler et Wood, 1966) for n_b and a variable sweep for n_{HP} .

The optimization process is first initialized as if integers were not mandatory for n_b . This process is akin to a relaxation of the original MINLP problem into a simpler one. A first optimum is found for all variables. If, by any chance, the optimum on n_b is an integer, then no branch and bound is performed and the optimization process stops for this variable. However, if n_b is a fractional number, two new branches will be created and queued. Both new optimization problems will be almost identical to the original one, the only difference being their intervals allowed on n_b . One problem will get a set of interval that goes from the rounded up previous solution to the maximum. The other problem will get an interval that starts at the minimum and goes up the previous solution rounded down. For example, lets consider that a first optimization reveals the optimum on n_b to be 7.8 within a possible range of 1 to 20 boreholes. The branch and bound algorithm would create a set of two problems : one that constraints n_b between 1 and 7 and the other between 8 and 20. The branch-and-bound algorithm will continue until one of the two following conditions are met. The obvious first condition is that all branches created by the algorithm have been explored and an optimum have been found. This could represent a significant computation burden so a second condition allows to cut short the optimization process by avoiding unpromising branches (Clausen, 1999).

Even though branch-and-bound could be used also on n_{HP} , we found after some testing that a significant computational gain could be achieved by using a simple variable sweep on n_{HP} . This is due to the fact that, for optimizations were n_b oscillates between 1 and 20 boreholes, a parameter sweep constrained by simple rules of thumb requires on average less iterations than performing a full-on branch-and-bound. These rules of thumb consider that the limits of the parameter sweep on n_{HP} are expressed as a function of the current guess on n_b (see the last row of Table 6.2). As part of the overall optimization strategy used in this work, a sweeping is performed each time a new iteration on n_b is tried by the optimization algorithm.

6.3.5 Subsampling of the NPV Distribution

For a typical value of ω equal to 1000 samples, the computation of the NPVaR requires 1000 simulations of the GCHP system each time the optimization algorithm iterates. Although we used the efficient simulation strategy presented in Section 6.3.1, this represents a significant simulation burden. A programming ploy was then used during the implementation phase to greatly bolster the calculation speed. The approach consists to first perform ω simulations, sort the resulting NPVs and keep the IDs of the lower $(\alpha + \beta)\omega$ outcomes. During the next iterations, the NPVs are computed only for the saved IDs, sorted and the lower $\alpha\omega$ values are used to obtain the new NPVaR. Simply put, $\beta \cdot \omega$ is a buffer of the worst cases stored on

top of the first $\alpha \cdot \omega$ NPVs. This strategy assumes that the cases corresponding to the worst NPVs will always be located in the lower tail of the distribution and present in the pool of the lower $(\alpha + \beta)\omega$ NPVs. The robustness of this assumption is championed and discussed in Section 6.8. The whole algorithm is summarized in Fig. 6.4 with a focus on the subsampling scheme used.

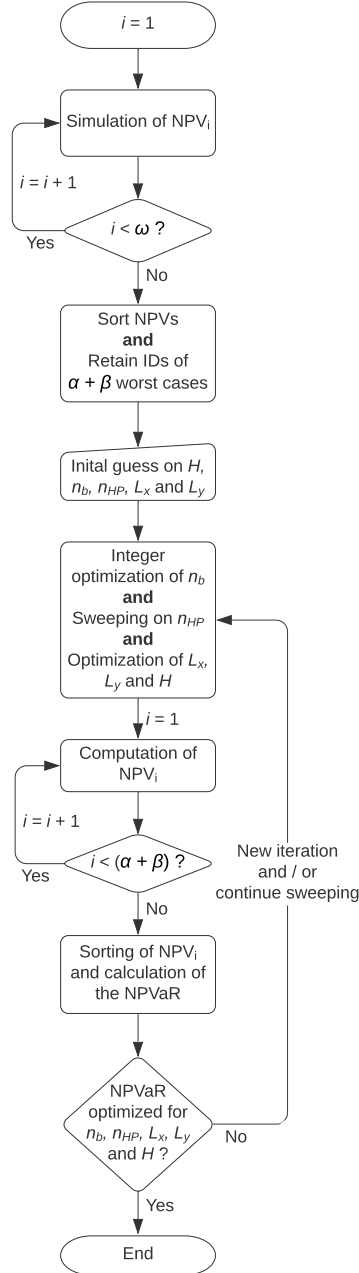


Figure 6.4 Simplified flow-chart of the algorithm used to optimize the NPVaR with a detailed breakdown of the subsampling scheme retained to speedup its evaluation.

6.4 Case Study Description

The algorithm presented in the previous section has been used to optimize the design of a hybrid GCHP system. The building used for this case study is a five-storey long-term care center located in Montreal, Canada. The thermal loads of this building ($2322\text{ m}^2/\text{storey}$) have been modelled in the *Simeb* (Millette *et al.*, 2011) environment considering a typical climate and construction for the city of Montreal. The resulting building loads of a single year, shown in Fig. 6.5, have peak loads of 534.6 kW in heating and 368.8 kW in cooling. With only electric heating and cooling, it costs 89451\$ in electricity using local tariffs per year accounting for power consumption and peak power demands each month. In January, the coldest month, it costs 18809\$ and in July, the hottest month, it costs 2494\$. Such thermal profile is good to showcase the robustness of the proposed algorithm because of two features although it is not typical of buildings that would usually get considered for hybrid systems. First, the thermal loads are heavily imbalanced in favor of heating, requiring 7.8 GWh of energy in heating and only 3.0 GWh of energy in cooling annually. Second, it has multiple sharp peak loads, both in heating and cooling. This prevents heat pumps to provide a good coverage of the building's energy needs and leads to heavy use of auxiliary systems. To illustrate this, Fig. 6.6 shows the relationship between the maximum power supplied by the hybrid GCHP system and the resulting energy coverage. As illustrated, if half of the annual peak load is covered by the GCHPs, almost 95% of the energy needs are covered.

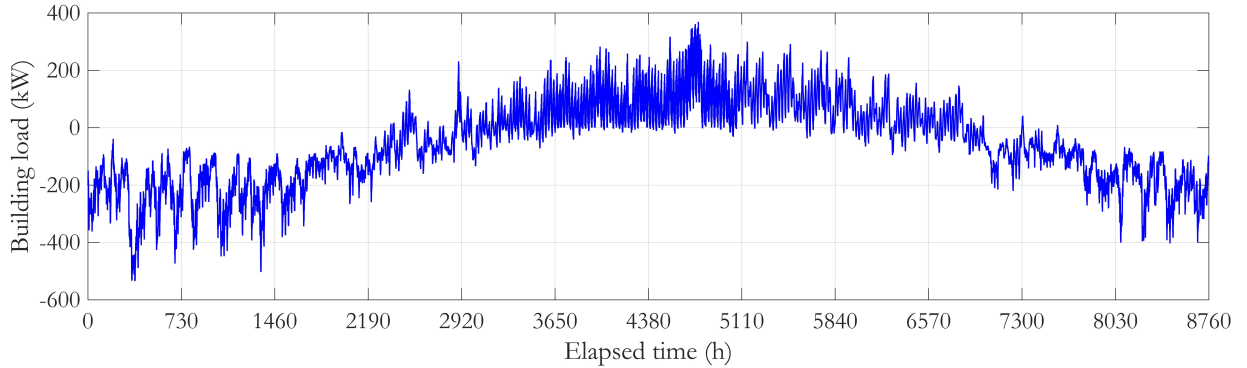


Figure 6.5 Building loads of a 5-storey long-term care facility located in Montreal, Canada. The first value corresponds to January 1st, while the negative and positive values correspond to heating and cooling loads respectively.

It is worth noting that the system considered in this work is relatively typical of new buildings in eastern Canada. When natural gas is not available locally or when building owners want to reduce their greenhouse gas emissions and peak power demand, use of hybrid GCHP system is a logical choice. For such system, a series of GCHPs provide heating or cooling to a

primary water loop, while electrical boilers and air-cooled chillers cover the remaining peak loads as shown in Fig. 6.1. Note that these two auxiliary devices are usually preferred to water-to-air heat pumps in northern climates due to the efficiency drop of the latter at very low air temperatures. The performances of the GCHPs used are typical of commercial units with a COP and CAP function of EWT as shown in Fig. 6.7. For simplicity, the COP of the auxiliary electrical boiler and chiller were considered to be constant at 1 and 3.3 respectively. Note that the competing reference system considered in this work uses only a combination of the same types of boilers and air-cooled chillers to provide heating or cooling to the primary water loop.

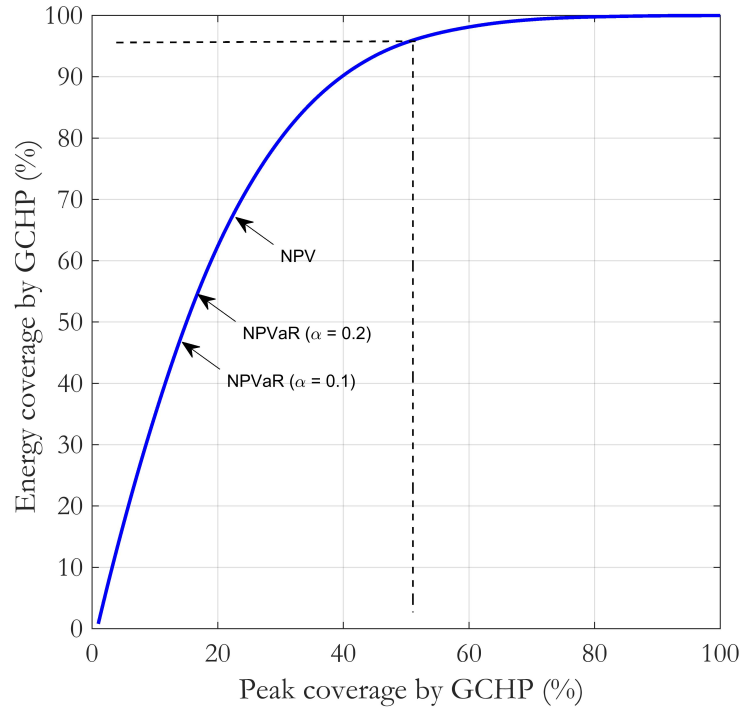


Figure 6.6 Percentage of the energy needs shown in Fig. 6.5 that can be covered according to the relative capacity of the GCHP system. The arrows points towards the three systems analyzed in Section 6.6.

6.4.1 Fixed design parameters used to compute the NPVaR

To compute the NPV of each system, the cash flow was evaluated based on the electricity tariffs applied by the local electricity company to single-payer and multi-residential buildings. This pricing includes a cost for energy consumption and a cost for power demand. For instance, the first 1200 kWh per month are charged at 0.0588 \$/kWh while the remaining electricity consumption is charged at 0.0894 \$/kWh. In addition, the highest power demand

of a month is charged at 4.59 \$/kW in summer and 6.21 \$/kW in winter. Note that the previous costs are in Canadian currency.

The values associated with the construction costs variables (c_1 to c_{12}) introduced with the NPVaR formula (Eq. 6.9 to 6.11) are presented in Table 6.1. These values were gathered by Hénault (2015) as part of its master thesis and used in a similar analysis (Hénault *et al.*, 2016). To assemble this database, a market study among geothermal equipment retailers and distributors, energy providers, designers and installers of GCHP systems was achieved by Hénault to obtain reliable costs for the greater Montreal area in 2015. These costs were used as-in in the current work.

Items	Variables	Values
Borehole drilling (\$/m)	c_1	50
Geothermal vault (\$/borehole)	c_2	1666
Borehole commissioning (\$/borehole)	c_3	200
Heat-carrying fluid (\$/L)	c_4	1.75
Trench, volume (\$/m ³)	c_5	60
Trench, width, (\$/m)	c_6	2
GHE commissioning (\$/borehole)	c_7	400
Thermal response test (\$)	c_8	10000
Geothermal heat pump (\$/kW)	c_9	733
Additional costs (\$/kW)	c_{10}	366
Auxiliary heating (\$/kW)	c_{11}	50
Auxiliary cooling (\$/kW)	c_{12}	586
Sub 1200 kWh consumption (\$/kWh)	c_{13}	0.0588
Above 1200 kWh consumption (\$/kWh)	c_{14}	0.0894
Summer premium (\$/kW)	c_{15}	4.59
Winter premium (\$/kW)	c_{16}	6.21

Tableau 6.1 Costs considered for the calculation of the NPV and NPVaR of the case study analyzed in Sections 4 and 5.

Depending on authors and types of project, the discount rate i used in Eq. 6.7 to compute the NPV can range anywhere between the expected inflation rate for a period to a few tens of percents. The inflation is almost never used directly to gage the profitability of a project because private individuals, companies and loan givers usually look forward on making a profit when investing. On the opposite end of the spectrum, high values for i are usually used when assessing the financial viability of risky investments to account for potential losses.

This discount rate i can make or break a project when not chosen carefully. We decided to select the discount rate in the same way it would be in large civil engineering infrastructure projects like bridges and hospitals. Thus, the GCHP systems are treated like necessary in-

vestments for future generations but still have to be efficient and affordable as they will not be constructed at any costs. Therefore, a discount rate of $i = 8\%$ was retained as suggested in Montmarquette et Scott (2007).

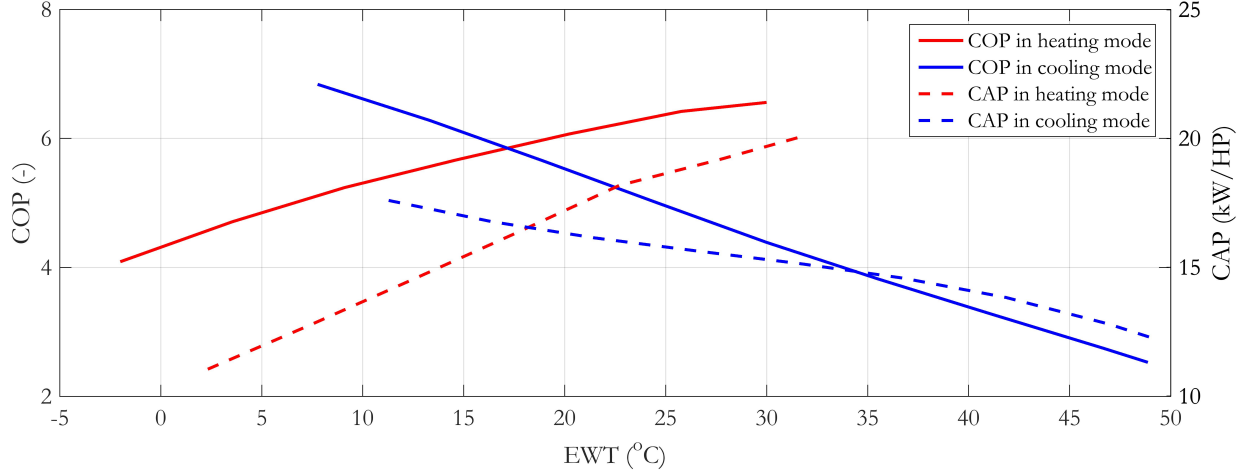


Figure 6.7 Coefficients of performance (COP) and capacity (CAP) of the GCHPs used in the case study.

6.4.2 Uncertain Design Parameters

In this work, four design parameters are considered uncertain. These parameters are the overall construction cost of the hybrid GCHP system, the inflation of the electricity cost, the ground thermal conductivity (k_s) and the building heat loads (Q_B) as illustrated in Fig. 6.8. To ease the analysis, this uncertainty is expressed by the probability density function (pdf) of a positive value that, with the exception of the monthly electricity cost, multiplies the value retained in the base case scenario. For the construction cost, this strategy avoids dealing with distributions for the individual costs shown in Tab. 6.1. The construction cost multiplier drawn from the PDF gets multiplied by the total upfront costs (CF_0) each time an NPV is calculated. Although the distribution shown in Fig. 6.8 is arbitrary and is not based on empirical data, the goal was to shape a distribution that has greater chances of inflating construction costs rather than lowering them. In this case, the lowest value the multiplier can take is 0.6, and its highest is 2.

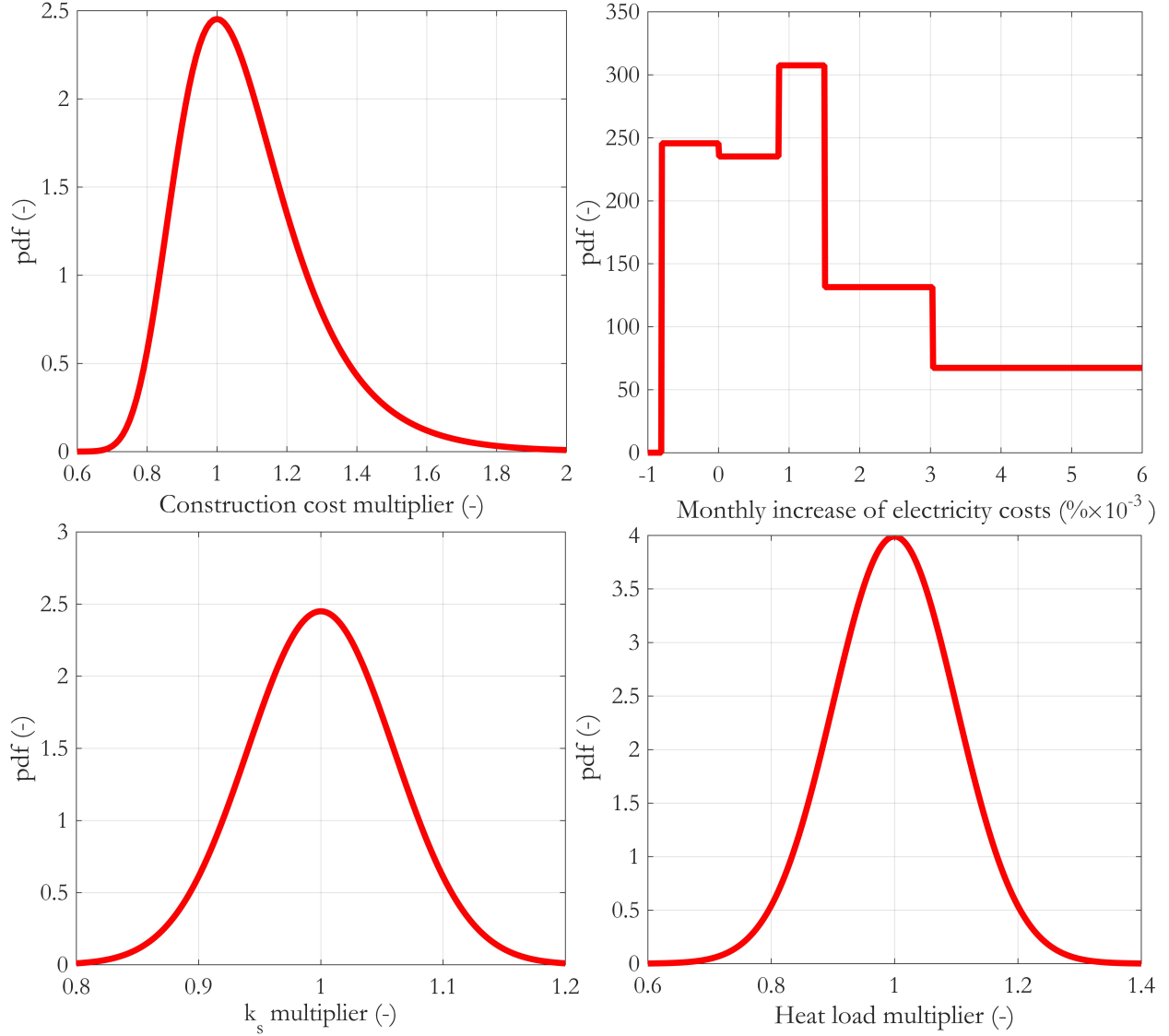


Figure 6.8 Probability density functions used to describe the four uncertain design parameters used in this work. The PDFs express, with the exception of the monthly electricity cost, the probability that the values retained in a base case scenario be multiplied by a given positive value.

The PDF of the soil thermal conductivity multiplier corresponds to a gaussian truncated distribution. With a base case scenario of $k_s = 2.5 \text{ W}/(\text{m}^\circ\text{C})$, a multiplication by this distribution allows to consider k_s ranging from $2 \text{ W}/(\text{m}^\circ\text{C})$ to $3 \text{ W}/(\text{m}^\circ\text{C})$. Indeed, such values span across the range of validity of many geological environments, such as granite (Cho *et al.*, 2009), frozen soils (Penner, 1970), saturated soil (Johansen, 1977), sandstone and marble (Clauser et Huenges, 1995).

A similar strategy was used to integrate the uncertainty of the building's thermal demand.

Instead of dealing with recalculating a new building's load signal each iteration, the thermal load presented in Fig. 6.5 was multiplied by a constant drawn from the PDF each time an NPV was calculated. Although the distribution itself is arbitrary and is not based on empirical data, the goal was to shape a distribution that has equal chances of inflating or lowering the demand. In fact, this represents a building having higher/lower insulation or occupancy than planned when the hybrid GCHP system was designed. In this case, the lowest value the heat load multiplier can take is 0.65, and its highest value is 1.35.

The distribution of the relative monthly increases and decreases in electricity costs corresponds to the step function shown in the upper right quadrant of Fig.6.8. This data-driven distribution was modeled using real electricity tariffs from Hydro-Quebec, the local electricity supplier, between the year 1989 and 2014 (MERN, 2016). The original data were yearly costs expressed in $\text{¢}/kWh$ and were converted into monthly relative variations using the monthly inflation rate. The lowest possible variation on an annual basis is a decrease of -0.97% while the highest possible increase is 7.44% . In practice, the PDF is used to assemble a vector of monthly cost variation for a period of 10 years. This vector is then used to compute the monthly electricity bill.

6.4.3 Design Variables Optimized

Five design variables are optimized in this work, These variables are the length of the boreholes H , the spacing between the rows L_x and columns L_y of the boreholes located on a regular grid, the number of boreholes n_b and GCHP n_{HP} . The limits used to constrain the variables can be found in Table 6.2. As mentioned previously, the sweeping range of n_{HP} varies throughout the optimization process as a function of n_b according to the set of rules presented in the last row of Table 6.2. Note in the table the initial values used to initialize the optimization process.

	Lower boundary	Initial guess	Upper boundary	Variable type
H (m)	50	150	150	Non-negative real
L_x (m)	3	5	12	Non-negative real
L_y (m)	3	5	12	Non-negative real
n_b (-)	1	10	20	Natural integer
n_{HP} (-)	$\text{Max}(1, \lfloor n_b/2 \rfloor)$	5	$\lceil 3n_b/4 \rceil$	Natural integer

Tableau 6.2 Boundaries and initial guesses for the 5 design parameters optimized in this work.

6.5 Algorithm Validation

In this section, the methodology described previously is validated using the case study introduced earlier. Since it is a complex methodology that consists of many parts linked together, the validation performed is two-fold. First, the NPVaR is implemented in a deterministic context assuming that no uncertainty taints the design parameters. This comes down to using a single NPV obtained with multipliers worth 1 (k_s , construction cost and heat load) and average electricity increases of 0.016% each month. In such situation, where $\omega = 1$, $\alpha = 1$ and $\beta = 0.0$, the NPV and NPVaR are equal. Using this objective function, the results of a parameter sweep and of an optimization are compared to make sure the optimization algorithm is correctly implemented and allows to find the global optimum of the objective function. These evaluations use the original heat load signal, the parameters, the electricity tariffs and initial values introduced previously. Second, the NPVaR is then used in a stochastic framework to check that the implementation of the objective function does not prevent the algorithm of reaching the global optimum. The following Table 6.3 shows the ranges and increments used during sweeping of the five design parameters.

Parameters	Starting point	Increment	End point
H (m)	50	50	150
L_x (m)	3	3	12
L_y (m)	3	3	12
n_b (-)	1	2	19
n_{HP} (-)	$\max(1, \lfloor n_b/2 \rfloor)$	1	$\lceil 3n_b/4 \rceil$

Tableau 6.3 Ranges and increments used during the sweeping of the five design variables.

As Table 6.3 shows, H can be in one of three states, both L_x and L_y can be in one of four states, n_b accounts for all odd numbers of boreholes between 1 and 19 and n_{HP} , according to the rule introduced in the two formulas, can range between 1 state when $n_b = 1$ to 7 states when $n_b = 19$. This is because, for $n_b = 1$, $\max(1, \lfloor 1/2 \rfloor)$ and $\lceil 3 \times 1/4 \rceil$ equal a total of 1 state. For $n_b = 19$, the same formulas would read $\max(1, \lfloor 19/2 \rfloor) = 9$ and $\lceil 3 \times 19/4 \rceil = 15$, and going from 9 to 15 corresponds to a total of 7 different states. In total, this leads to 2112 different combinations that are calculated to define where the global optimum of the objective function lies.

6.5.1 Deterministic use of the NPVaR

The upper graphic of Fig. 6.9 illustrates the results of the first sweeping process performed for the deterministic case. Different frequencies are visible throughout the figure. These patterns are linked to the way the five variables were ordered before the sweeping. The order was n_b starting at one, then n_{HP} , H and finally L_x and L_y . This explains why the wavelength corresponding to n_b increases with higher values of n_b as it includes more and more heat pumps (n_{HP}). A close up view of the sweeping process of the NVP is shown in Fig. 6.10 and illustrates that the second longest wavelength is tied to n_{HP} . The sweeping of H is responsible for the step increase of the NPV for any pair of n_b and n_{HP} . Finally, the variations linked to L_x and L_y are smaller, but still represents a variation of a few thousands of dollars.

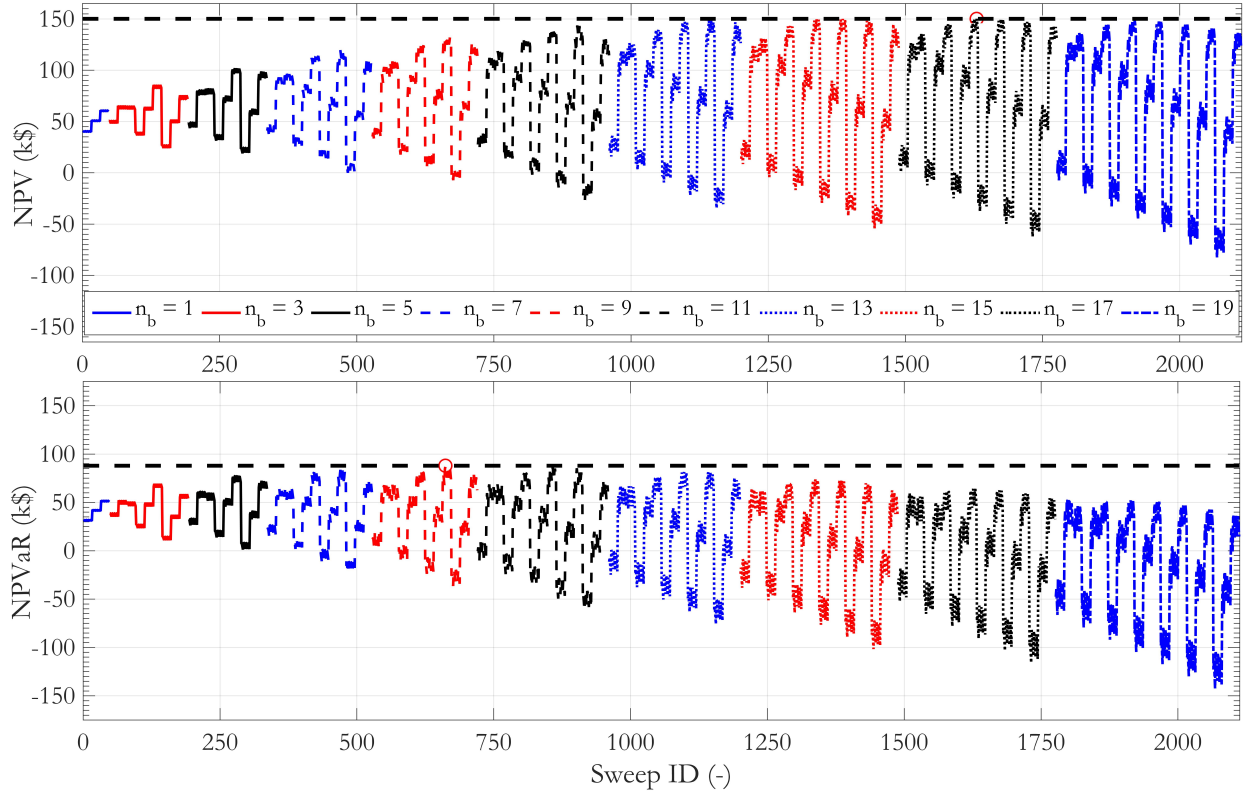


Figure 6.9 NPV (upper figure) and NPVaR (lower figure) obtained by sweep of the design variables n_b , n_{HP} , H , L_x and L_y to assess the global optimum of each cost function. The black horizontal dashed line represents the best NPV or NPVaR reached by optimization while the red circle shows the location of the sweeping optimum.

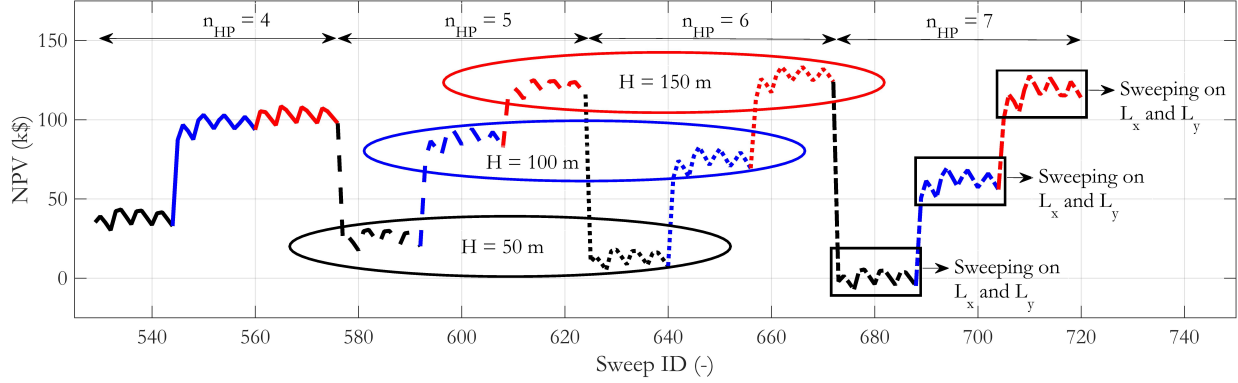


Figure 6.10 Close up view of the upper part of Fig. 6.9 for $n_b = 9$. Each color corresponds to a different borehole length (black for $H = 50$ m, blue for $H = 100$ m and red for $H = 150$ m as highlighted by the ellipses), each line style is for a different value of n_{HP} while the sweepings of L_x and L_y is illustrated by the boxes.

The maximum NPV found during the sweeping process is illustrated in Fig. 6.9 using the red dot. This optimum corresponds to $n_b = 17$, although the best cases associated with $n_b = 13$ to 19 are similar. This indicates that the global optimum in the deterministic case could possibly be somewhere between 13 and 19 boreholes if the sweeping was performed with a finer grid. The highest visible NPV is 150.1k\$ and is achieved during the 1630th sweep. This corresponds to $n_b = 17$, $n_{HP} = 10$, $H = 150$ m, $L_x = 12$ m and $L_y = 6$ m as shown in Table 6.4.

The sweeping is now confronted to the optimization results for the deterministic case. The optimum design variables are summarized in Table 6.4 and the best NPVs are represented by the black dashed lines in Fig. 6.9 for comparison purposes with the sweeping process. These results seem trustworthy in regard to the parameter sweep. Indeed, n_{HP} and H are the same, and n_b only differs by 1 borehole. However, L_x and L_y are the most different variables, but the sweeping process also proved that they contribute relatively little to the NPV. Furthermore, 16 boreholes seem like a reasonable candidate for the optimum, since the highest NPVs were almost equal for $n_b = 13$ to $n_b = 19$ and because $n_b = 16$ wasn't specifically covered by the sweeping process.

6.5.2 Stochastic use of the NPVaR

All the parameters used for the first sweeping are used exactly as-is for this second validation test. The only change is that the NPVaR is used as the objective function instead of the NPV. For its calculation, a sample of 100 NPVs ($\omega = 100$) was selected, paired with $\alpha = 0.1$ and

Objective function	Deterministic-NPV		Stochastic-NPVaR		
	Sweeping	Optimization	Sweeping	Optimization	$\alpha = 0.2$
Optimum NPV ($k\$$)	150.1	150.8	87.9	89.3	96.5
n_b (-)	17	16	9	9	11
n_{HP} (-)	10	10	6	6	7
H (m)	150.00	150.00	150.00	150.00	150
L_x (m)	12.00	8.83	6	7.28	7.69
L_y (m)	6.00	5.27	6	4.73	4.83
Computational time (hours)	4.7	3.8	75.9	48.2	158.5

Tableau 6.4 Results and computational time for the sweep, the deterministic and stochastic optimization that use the NPV and NPVaR as objective function.

$\beta = 0.05$. The results of the sweeping are shown in the lower graphic of Fig. 6.9. The whole process required calculating 2112 NPVaRs, which was achieved in 75.9 hours. Fig. 6.9 shows a more distinct optimum than it did for the deterministic sweeping. The best NPVaR occurred during the 662th sweep ID, which corresponds to an optimum of 87.9 $k\$$. This optimum was obtained in conjunction with $n_b = 9$, $n_{HP} = 6$, $H = 150$ m and $L_x = L_y = 6$ m (see Table 6.4).

The same types of oscillations observed previously can be seen on the lower graphic of Fig. 6.9. This shows that again, H , n_b and n_{HP} have a higher impact on the NPVaR than L_x and L_y . The sweeping performed with the NPVaR is also confronted to its optimization counterpart. Again, these results seem trustworthy in regard to the sweeping that was performed. This time, n_b , n_{HP} and H are identical, while L_x and L_y show relatively small variations. The optimization's optimum NPVaR is slightly higher than the one obtained by sweeping with a relative difference of only 1.64%. This is not surprising since the sweeping had less freedom of movement throughout the search space, a limitation the optimization algorithm had not. Since the maximum NPVaRs obtained by sweeping and optimization are almost the same, we conclude that the implementation of the algorithm developed in this work has been achieved thoroughly and thus can be used for the optimization of ground source heat pump systems.

6.6 Gains at Using the NPVaR

To clearly show the advantages of using the NPVaR over the traditional NPV, let's suppose that the uncertain parameters are revealed after construction of the system. This means that the probability density functions of the four uncertain parameters, that integrated uncertainty during the design phase, collapse into their real, final value. In this experiment, three systems will be compared : the two presented in Sections 6.5.1 and 6.5.2 obtained by NPV and NPVaR,

and a third one also obtained using NPVaR but with $\alpha = 0.2$. The optimum and associated variables values that correspond with this third system are shown in Table 6.4 alongside the other systems. Simulating these three hybrid GCHP systems with their real parameters allows assessing the real energy use of the GCHPs and auxiliary systems, and assess their NPVs with respect to the expected ones. For the interested reader, a sensitivity analysis that covers the impacts of varying the α , β and γ parameters when using the NPVaR can be found in Annex.

6.6.1 Actual NPV and Risk Management

To establish the distribution of the NPVs, the real parameters are drawn 1000 times from the distributions shown in Fig. 6.8 and another 1000 times from bounded uniform distributions to establish the impact of the a priori distributions. The uniform distributions are bounded by the same values as the other distributions. For instance, the construction cost multiplier and soil thermal conductivity multiplier have the same probabilities of getting assigned any value from 0.6 to 2 and 0.8 to 1.2 respectively. Figures 6.11 and 6.12 compare respectively the histograms and cumulative distribution functions of all three systems in post-design, real conditions. Recall that the NPVs shown are the differential NPVs, that is the difference between the NPV of the GCHP and reference systems.

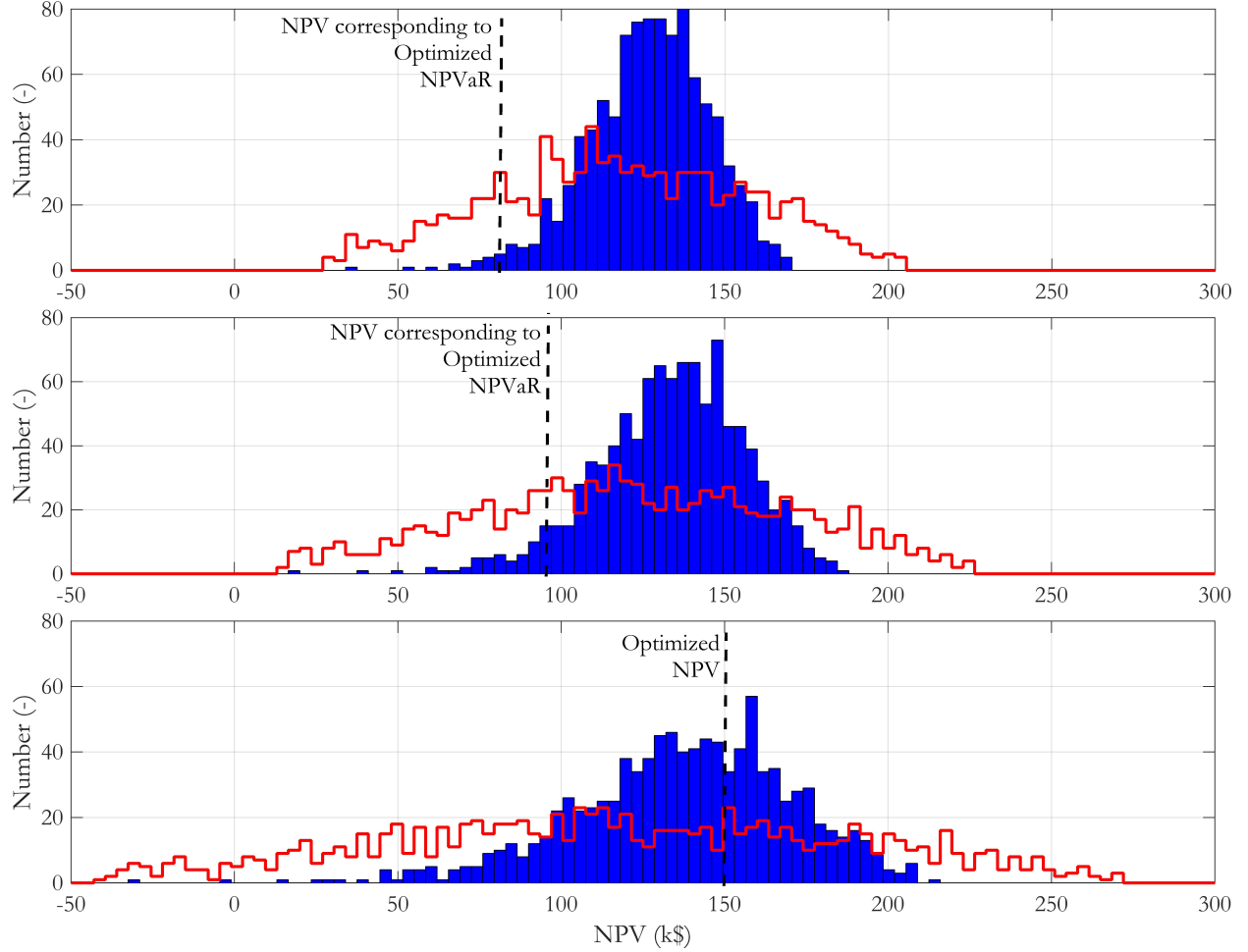


Figure 6.11 Actual NPVs after 10 years of operating the hybrid GCHP systems calculated using the design proposed by the stochastic NPVaR using $\alpha = 0.1$, $\beta = 0.05$ and $\omega = 100$ (top), using $\alpha = 0.2$, $\beta = 0.05$ and $\omega = 100$ (middle) and using the deterministic algorithm (bottom) based on the NPV as objective function. The blue histograms are associated with the statistical distributions of Fig. 6.8, and the red lines to the bounded uniform distributions.

First, note that in Fig. 6.11 a) and b), most NPVs obtained with the real parameters are higher than the predicted NPVs of 89.3k\$ and 96.5k\$, respectively, optimized with the NPVaR (dashed black lines). In fact, 97% of cases for the NPVaR ($\alpha = 0.1$) and 94.2% of cases for the NPVaR ($\alpha = 0.2$) are worth more than anticipated when considering the same distributions of uncertain parameters used during the optimization (blue histograms). Even when considering flat distributions for all parameters (red lines), they are still 75% of cases and 68.7% of cases that exceed expectations. This confirms that using the NPVaR to design a GCHP system will likely lead to higher NPVs during its real operation and thus is a safer financial design. By comparison, the histograms corresponding to the system designed with the NPV (see Fig.

6.11 c) are much wider and only 36.4% of the NPVs of the blue histogram cases and 34.3% of the red line cases exceed the optimized value of 150.8 k\$. This is not surprising since the maximum of a near-gaussian distribution is usually close to its median. The distribution's tails, that contains the worst returns, is also longer and stretch all the way into financial losses territory for both distributions.

On average, the NPV-based objective function yields a better return on investment than the NPVaR objective function by only 7.8% (for $\alpha = 0.1$) and 3% (for $\alpha = 0.2$) when considering the original distributions. This difference jumps to 20.2% and 13.3% when looking at the best cases : 213.6k\$ for the NPV and 170.4k\$ and 185.3k\$ for the NPVaRs. These results were, however, predictable since the NPVaR is not designed to be a "high risk-high reward" kind of investment tool. What it excels at is proposing safe designs, that consider the worst possible outcomes in order to best shield against losses. Therefore, the NPVaR should be expected to perform better than the NPV when considering worse cases.

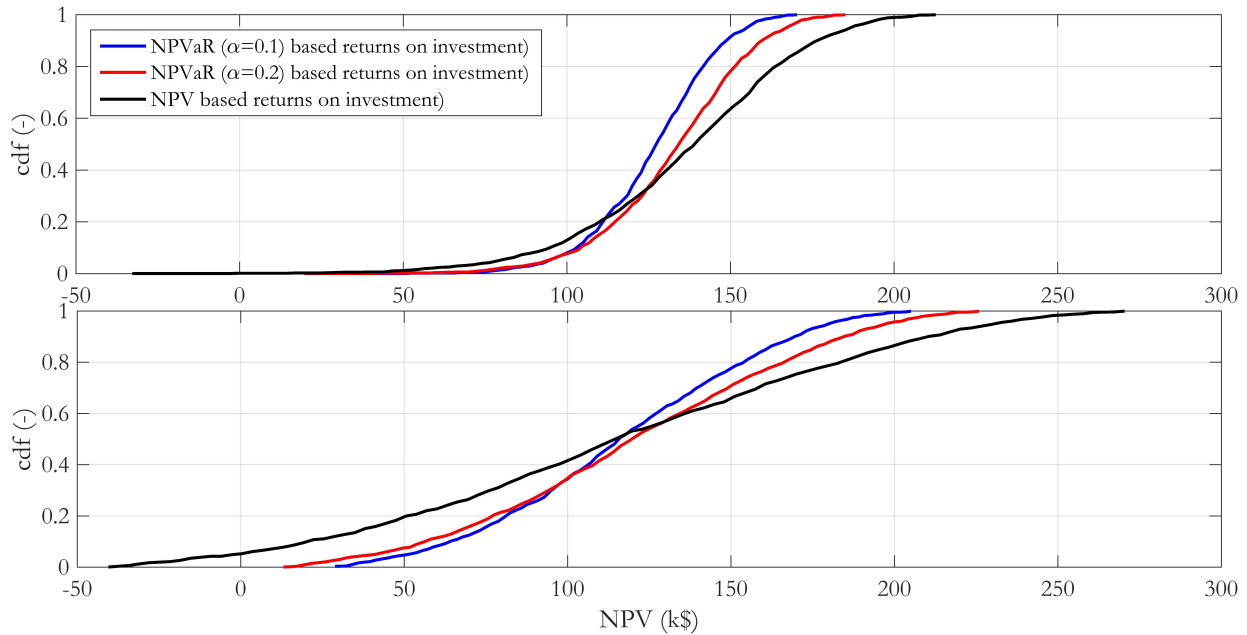


Figure 6.12 Cumulative distribution functions of the NPVs shown in the histograms of Fig. 6.11. The three above correspond to the original distribution of uncertain parameters and the three below are for the flat distributions.

This tendency can be observed when looking at the cumulative distribution functions of Fig. 6.12 that were assembled using the results presented in Fig 6.11. The cdfs obtained by NPVaR are all more compact than the ones obtained by NPV. Their maximum values achieved are not as high, but their worst cases are also not as low. For example, when comparing the upper

blue and black lines that correspond respectively to the design by NPVaR ($\alpha = 0.1$) and by NPV, for the percentiles 75 to 100, the NPV-based values yields on average 27.78 k\$ more in return than the NPVaR-based results, resulting in gains 15.7% higher. This difference goes down to 11.0% for the percentiles 50 to 75, and to only 5.5% for the percentiles 25 to 50. This trend finally inverts to 9.2% for the percentiles 0 to 25 in favor of the NPVaR. Therefore, the relative differences prove that the NPVaR outperforms the NPV for unfavorable cases, which was the intended goal in the first place.

However, as mentioned previously, the real power of the NPVaR as an objective function is shown when looking at the absolute worst conditions possible and the associated net present values of the system. Indeed, the average NPVaR's return on investment is two and a half times that of the NPV when considering the average worst 10 cases for $\alpha = 0.1$ and 2.25 times for $\alpha = 0.2$. Finally, when considering the absolute worst-case separately, the NPVaR can help turn considerable losses into gains, resulting in the design by NPVaR never losing money no matter harsh the circumstances and no matter the distributions considered. All these observations point to the fact that the NPVaR is a safer objective function since it is more likely to yield performances better than expected, while shielding against poor performances and losses.

6.6.2 Impact on the design and operation of GCHP systems

This subsection begins with a comparison of performances between designs in terms of energy and power coverages. The results of this comparison exercise are presented in Table 6.5. In theory, all designs should fall somewhere on the curve presented in Fig. 6.6. Once the analysis of these data is completed, the power distributions between GCHPs and auxiliary equipment effects on the monthly billing and the monthly cashflows of each design as well as the impacts of the objective function on the design variables will be investigated. From now on, the focus will be put on the designs obtained by NPV and NPVaR ($\alpha = 0.1$) to capitalize on their sharper contrast and lightened the analysis.

Objective function	NPV	NPVaR($\alpha = 0.1$)
Energy coverage (%)	67.41	47.96
Power coverage (%)	23.24	14.09

Tableau 6.5 Energy and power coverage of the designs optimized with the NPV and NPVaR ($\alpha = 0.1$).

As expected, all combinations of energy and power coverages fall precisely on the curve of Fig. 6.6. As Table 6.5 shows, the design based on the NPV makes a more ambitious use

of geothermal energy than the other. Overall, by investing in a GCHP system that can contribute 7.15 kW more in peak periods than the other hybrid system, the NPV-based design can cover 20% more of the building's energy needs during its yearly operation. With this setup, the NPV-based design can theoretically meet the building's energy needs using only its GCHPs during 1125 hours per year (i.e. 12.84% of the time) while the GCHPs of the NPVaR-based design ($\alpha = 0.1$) can only operate by themselves 726 hours per year (i.e. 8.29%). This shows that, for both systems, the GCHPs and the auxiliaries need to work together most of the time.

Some may be surprised of the low power and energy coverages of the optimized solutions and wonder whether this is a consequence of using the proposed algorithm as they differ quite a lot from the traditional figure of 60 % / 90 % (power/energy coverage). We believe the answer lies with the building that was used in the case study. As mentioned previously, the building's heat loads are heavily imbalanced in favor of heating and present multiple sharp peaks, both in heating and cooling. This situation is highly unfavorable for geothermal energy, even for hybrid systems, as auxiliaries are highly needed during winter and therefore the operating costs are not lowered by much when compared to the reference system. In that regard, an interesting observation concerning the power energy coverage versus peak energy coverage graphic can be made when redrawn in the following semi-logarithmic format.

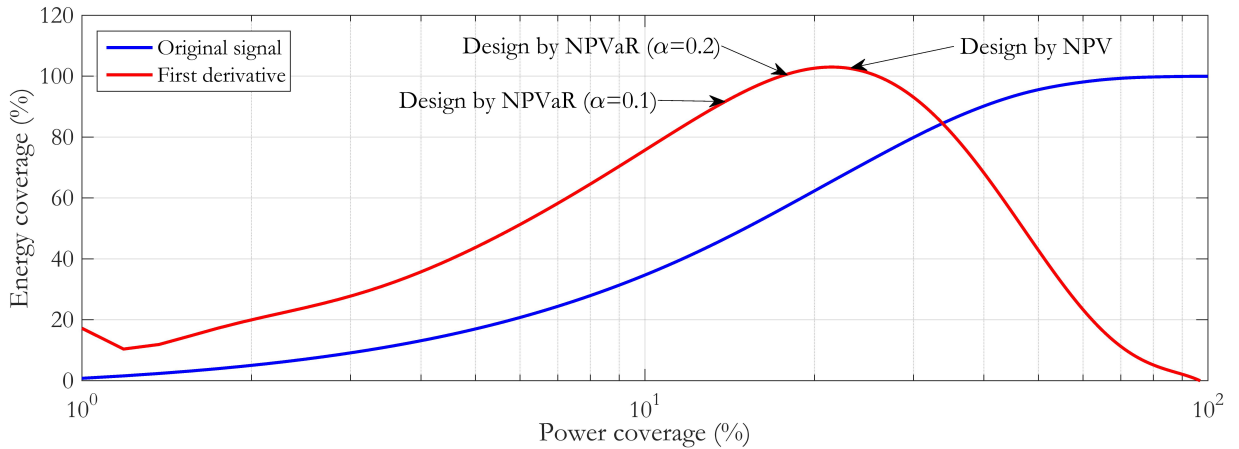


Figure 6.13 Semi-logarithmic presentation of the power versus energy coverage curve associated with the long term care center of the case study. In red, the derivative is presented alongside the original curve in blue without its own axis simply to show its behavior.

As it can be seen on Fig. 6.13, there is a clear point of inflexion in the blue curve, which is the original curve presented back in Fig. 6.6, around the low 20s of percents. As highlighted by the derivative of the signal in red, it corresponds to a point onward from which there are

diminishing returns. In other words, more power coverage translates to less and less energy coverage the more heat pumps and larger GHEs are installed after that point. Coincidentally, the optimized system obtained using the NPV sits right around that inflection point at a power coverage of 23.24 %. We believe that the position of the optimum on the blue curve is dependent on the relation between construction costs and operating savings and is visibly influenced by the uncertain parameters since harsher circumstances resulted in optimums getting deported leftward.

The consequences of the mismatches between heating and cooling can be seen in the monthly operating costs as well. As shown in Table 6.6, out of the almost 90 k\$ an all-electric, reference system that does not uses geothermal energy costs to operate yearly, more than 18 k\$ (21 %) are spent in January alone during the peak of winter while less than a seventh of that amount is required in July during the peak of summer. This greatly reduces the potential for savings during a large portion of the year and thus extends the payback time. This can be seen when comparing the operating costs for January and July of the hybrid systems that were sized using the NPVaR and NPV to the one of the reference system. Here, the payback period is defined as the number of months it takes to save the extra money that was required to construct a hybrid system instead of the reference system. If the goal was instead to reimburse the whole system, payback periods would be much longer.

Type	Annual (k\$)	January (k\$)	July (k\$)	Payback time (month)
1-NPVaR	65.2	15.1	2.1	36
2-NPV	51.0	12.5	1.8	55
3-Reference	89.5	18.8	2.5	N/A

Tableau 6.6 Annual and monthly operating costs for the months of January and July and payback time for the systems design by NPVaR ($\alpha = 0.1$), by NPV and the reference system.

Now, Fig. 6.14 presents the energy billing breakdowns of three systems. The yearly billing of the electricity used by the same reference system introduced in Table 6.6 is added to give a comparative base to the billings of the electricity used by the NPV and NPVaR-based designs. The major difference in the way electricity is charged for the three systems comes from the energy consumption that exceeds 1200 kWh each month represented by the blue bars. This was expected since GCHP systems are notoriously efficient and require less energy to heat and cool buildings than the auxiliary and reference equipment (Type 3). As shown by the red bars, all three systems cost the same amount to operate for their monthly sub-1200 kWh energy consumption. This is due to the large energy needs of the long-term care center used in the case study. Because of it, every month, even in spring and fall, requires at least 1200 kWh of electricity to power both heat pumps and auxiliary equipment. Surprisingly, the

black bars show that the power premium that penalizes power consumption above 50 kW are also relatively similar for the three systems. This is explained by the fact that most of the premiums occur in the worst of winter where the GCHPs of both systems only account for a small fraction of the total energy provided to the building while the rest is supplied using inefficient electrical heating. This again can be attributed to the much higher needs for heating than cooling of the building. As evidence, the variations in the premiums of the three systems (78k\$, 90k\$ and 108k\$ from left to right) shows a good correlation with the percentage of auxiliaries that are used to cover the building power needs during the peak period ($100\% - 23.24\% = 76.76\%$, $100\% - 14.09\% = 85.91\%$ and 100% respectively).

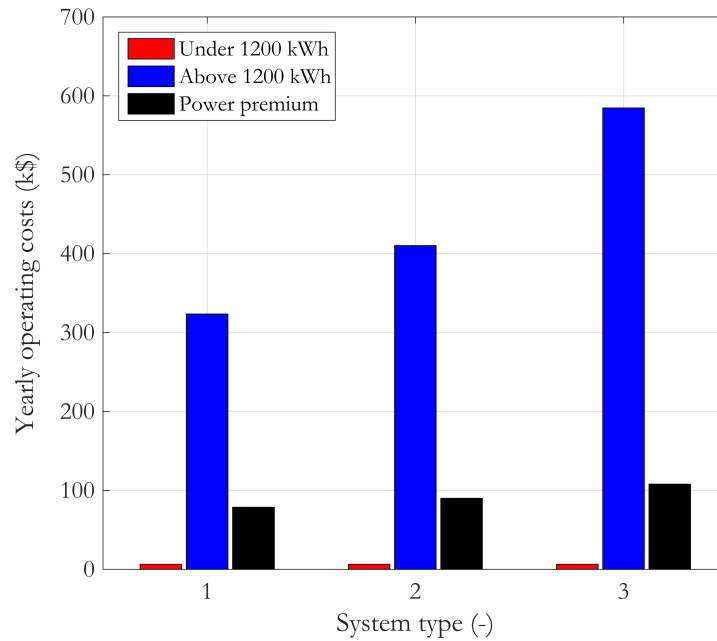


Figure 6.14 Yearly energy billing breakdowns for the operating costs of the NPV (System type 1) and NPVaR-based designs (System type 2) and a reference system without GCHPs (System type 3).

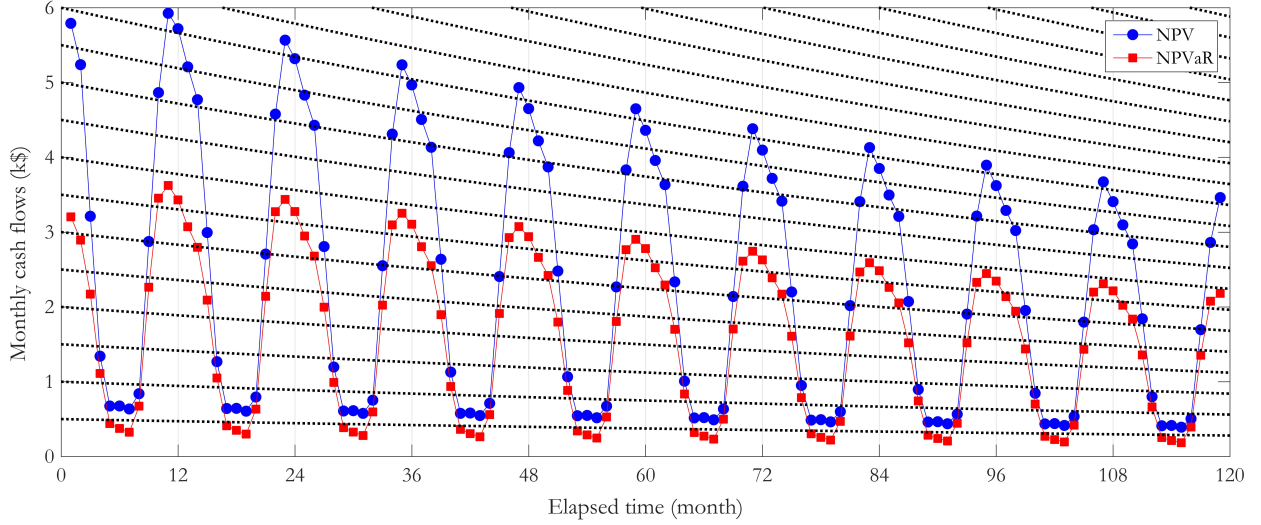


Figure 6.15 Cashflow differences between the NPV and NPVaR-based designs. The dotted black lines illustrates the impact of the discount rate and electricity tariffs changes on the value of cashflow.

Fig. 6.15 illustrates the differences in monthly cashflows during the 10 years operation of the NPV and NPVaR-based designed GCHP systems. As defined in Eq. 6.7, cashflows, not to be mistaken for the energy billings presented earlier, are really the difference in operating costs between the hybrid system and the full-electrical, reference system. Since the construction costs that occur during the 0^{th} month aren't shown, Fig. 6.15 holds only positive values. Unsurprisingly, the NPV-designed system generates more revenues than the other one as it uses more inexpensive to operate GCHPs. Although still significant at the end of the 10 years horizon, the absolute differences in cashflows between the two systems shrink continuously during their operation. The explanation for this is two-fold. The most important factor is the time variation of money value, as shown by the dotted lines on Fig. 6.15. As it can be seen, the cashflows that correspond to identical months for a system tend to progress parallel to the dotted lines throughout the years. Therefore, profits that are generated using GCHP systems are diminished accordingly as time goes by with respect to the 0^{th} year. The second and more minor factor is the ground thermal loads that are unbalanced because of the heating-dominant building. Because of this, both systems rely slightly more on inefficient auxiliary equipment as time progresses. This is true for both systems of course but is only clearly visible on the graphic for the NPV-designed system that relies more on GCHPs. Indeed, the most upper points on the graphic get notably closer to a dotted line as time progresses.

The reason the second factor contributes so little can be explained by looking at the number of active heat pumps and EWT on Fig. 6.3. Although this figure features only the data of the

first year of the system designed using NPVaR, important observations that apply to both systems can be made. From late November to late March, the EWT drops rapidly as all heat pumps are operated constantly to the maximal capacity. Beginning in the middle of January, at the year's highest heat demand, the EWT is as close to the minimum allowed by the heat pumps as possible and their constant operation is no longer possible. Since this is how the system is operated from the first year, there is almost no margin for future years when the yearly average ground temperature inevitably drops due to thermal imbalances between heating and cooling. Therefore, the operation of the system changes very little between the first and the tenth year, hence the small differences in cashflows throughout the simulations.

Now to investigate the impacts the objective function has on the design variables at the optimum. The difference in surface area taken by the two GHEs is quite significative whether the NPV or NPVaR is used for sizing. Because it requires 16 boreholes (two full columns of 5 boreholes and the remaining four placed on a third column) where rows are spaced by 8.82 meters (L_x) and columns by 5.27 meters (L_y), the GHE designed by NPV needs at least $558.4m^2$ of land. On the other hand, at only 9 boreholes (distributed in two columns of four and 5 boreholes), with row and column spacings of 7.28 and 4.78 meters respectively, the land area required goes down to only $137.74m^2$. That can be significant in dense urban areas where land is expensive as it could raise the upfront cost of the project sized using NPV significantly more than the other one and even prevent the construction of the GHE entirely for existing building.

As shown in Table 6.4, the optimized length of boreholes (H) in both GHEs wasn't influenced at all by the objective function used. From this fact, a conclusion can be reached. Either because there was no penalty for drilling as deep as possible or because the allowed bracket of borehole lengths was too shallow, the algorithm never settled for other solutions. However, during the optimization phase, the algorithm could clearly be seen trying possible solutions that used boreholes of less than 150 meters, which was time consuming and proved to be unnecessary. Therefore, for future works, H should either be switch to a design parameter instead of a variable and fixed to a single value or see its allowed range expanded and shift deeper. The ramifications of this choice should be investigated in due time and be based on local geological conditions, drilling costs and building's energy needs.

The impacts the objective functions have on n_b and n_{HP} are probably the hardest to assess individually as their contribution are really intertwined when calculating the return on investment on a project. One thing that can be said for sure about n_{HP} is that less heat pumps, 6 as opposed to 10 when using the NPVaR, results in a small mechanical room and less complex piping. As for n_b , 9 boreholes instead of 16 equals a shorter closed-loop system

that generates less pressure loss. This itself translates to less pumping power required and thus fewer operating costs, which was not considered in this study. These observations on the impacts of the choice of the objective function on design variables as well as all the financial considerations that have been discussed thus far lead to the conclusion that undersizing is a viable and safe approach to designing hybrid GCHP system when faced with uncertainties.

Now for some final observations. The optimum found using the deterministic implementation of the NPVaR leads to a maximum ΔT in and out of the heat-pumps of $4.5^{\circ}C$ in cooling season and of $2.25^{\circ}C$ during the heating season. Those values translate to 2.87 gpm US per ton in cooling and 5.72 pgm US per ton in heating, which is not very far from the value of 3 gpm US per ton commonly used for design. This optimum implies a maximal linear ground load of $58.56 \frac{W}{m}$. This falls in the normal operation range of between $50 \frac{W}{m}$ and $80 \frac{W}{m}$ for vertical closed-loop systems (Banks, 2012) and therefore provides another layer of confidence in the validation presented earlier.

Additionally, the optimum found using the stochastic implementation of the NPVaR corresponds to a maximal ΔT of $5.045^{\circ}C$ in cooling and of $3.123^{\circ}C$ in heating throughout the 10 years simulation. Those values translate to 2.54 gpm US per ton in cooling and 4.10 pgm US per ton in heating, which is even closer to 3 gpm US per ton figure than the other proposed design. These differences are also within the prescribed operational limits of the modeled heat-pumps. Finally, this resulting borehole field will transfer $62.22 \frac{W}{m}$ during peak performances, which again falls within the $50 \frac{W}{m}$ to $80 \frac{W}{m}$ recommended bracket for vertical closed-loop systems.

6.7 Conclusion

Traditionally, when automated on a computer, the sizing of hybrid ground-coupled heat pump systems is done using deterministic algorithms. This means that, under known parameters, an optimal set of variables that maximizes an objective function is found. Classical objective functions usually involve the financial value of the system after a fixed amount of years in operation. If some parameters are found to be subjected to an amount of uncertainty, sensitivity analyses are performed in hindsight to assess their impact on the optimum. Although these analyses can inform on the relative contributions of each variables to the optimum, they cannot make predictions on the best set of variables when faced with uncertainties during the design.

In this article, we show how a stochastic objective function called the net present value-at-risk (NPVaR) helps mitigate instead of quantifying the impact of uncertainties associated with

the design of a hybrid ground-coupled heat pump systems. This process takes place during design phase as it is an integral part of the optimization process. The benefits of using this method are shown in a case study that compares a design proposed using the NPVaR against one obtained using the classical net present value (NPV). In it, four parameters are treated as statistical distribution instead of fixed data : the construction costs, the building's heat loads, the electricity costs and the soil thermal conductivity. Our results show that this new objective function produces systems that are worth more when considering scenarios less favorable to geothermal energies, even when using different statistical distributions during the validation phase than the ones that were used during sizing. Furthermore, it helps shield against unforeseen losses to a high degree when considering the worst combinations of uncertain parameters by proposing designs that almost constantly overperformed their expected value, which is not the case when NPVs are used for sizing. When designed for 10 years, the NPVaR method generates on average similar revenues than the reference NPV proposed design. However, the average revenues are 9.22% higher when considering the worst 25% scenarios, 2.6 times higher for the worst 10 scenarios and losses of 32.2k\$ are transform into revenues of 35.2k\$ when considering the absolute worst-case scenario. Since the design proposed by NPVaR rely less on geothermal energy than its counterpart, the break-even point is reached in 36 months instead of 55 in the proposed case study.

This method of design is radically different than current approaches who rely on deterministic algorithms and then perform sensitivity analysis in retrospect to try to gauge the robustness of their design. Although NPVaR requires many times more calculations from a computer to construct than a corresponding NPV, we believe our method will be a welcome addition to the toolbox of designers, especially in regions where geothermal energies are less prevalent. With more robust designs, this will surely bolster the confidence of governments, industries and population alike towards geothermal energies and help for a quicker transition to sustainable and renewable energies. Finally, it proves beyond a reasonable doubt that in face of uncertainty regarding some parameters, under-sizing is a good, safe and still profitable approach to design.

Acknowledgments

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through a Discovery grant.

6.8 Appendix - Fine-tuning of the NPVaR related parameters

Since the parameters α , β and ω can impact the optimization process, this section aims to illustrate how their value can inform the solution found by the algorithm presented in this work. To do so, the section presents three simple OAT sensitivity analyzes that showcase how an increase or a decrease in a particular parameter can affect the optimum NPVaR when compared to a reference case. As a reminder, the reference case used a sample size $\omega = 100$, a fraction of the worst outcomes $\alpha = 0.1$ and a buffer $\beta = 0.05$. Note that all optimizations performed in this section were launched using the initial solution used previously.

6.8.1 Impact of the fraction of the worst outcomes α

The experiment designed to assess the sensitivity of α consists of two different cases, one with a low α of 0.01, and one with a high α of 0.25. The other two parameters remain constant at $\omega = 100$ and $\beta = 0.05$. The results summarized in Table 6.7 indicate clearly that the value of α does have an impact on the optimum location. When a small α value is used, the algorithm focuses on optimizing only the worst NPVs and therefore produces a smaller NPVaR. This smaller NPVaR corresponds naturally to a more conservative design that relies less on geothermal energies and more on auxiliaries heating and cooling. The GHE is therefore smaller, 8 boreholes instead of 9 in our case. On the contrary, when a larger α value is used, the resulting NPVaR is higher at the optimum. This difference is significant, more than 10% in this case, and allows the algorithm to suggest a design that incorporates more boreholes and geothermal heat pumps. Since the number of NPVs computed at each iteration is given by $(\alpha + \beta)\omega$, a smaller α value leads to faster computation times while larger ones to longer computation times. Of course, the choice of α should not be based on computation times but should be adjusted to match the risk tolerance of a designer.

Optimum found	Reference Case	Impact of α		Impact of β		Impact of ω	
		0.01	0.25	0.0	0.1	10	1000
NPVaR ($k\$$)	89.3	80.6	99.8	89.3	89.3	100.3	87.0
n_b (-)	9	8	10	9	9	10	9
n_{HP} (-)	6	5	7	6	6	7	6
H (m)	150	150	150	150	150	150	150
L_x (m)	7.3	7.15	7.25	7.3	7.3	7.6	7.35
L_y (m)	4.75	4.4	4.85	4.75	4.75	5.0	4.7
Computation time (h)	48.2	24.8	68.5	29.6	68.6	3.8	747.3

Tableau 6.7 Impact of the NPVaR related parameters on the optimum location.

6.8.2 Impact of the buffer β

The experiment designed to assess the impact of β consists of two different cases, one with a nil value, and one with a higher value of 0.1. Again, the other two parameters remain constant. As indicated in Table 6.7, β has no impact on the optimum found by the optimization algorithm. Additional tests, however, revealed that only small borehole fields, consisting of one to three boreholes, had really different sets of design variables at their optimum. For boreholes fields consisting of more than tree boreholes, the 10 worst combinations within a pool of 100 possible outcomes ($\alpha = 0.1, \omega = 100$) seem to be almost identical even though their order can change. For that reason, β could be set to a small value for small scale projects, but set to zero for medium to large scale projects to save computation time.

6.8.3 Impact of the sample size ω

The sample size ω does have a limited impact on the optimum. When a small value is used, the algorithm only has a few combinations of NPVs to work with. For $\omega = 10$ and $\alpha = 0.1$, this means that the algorithm only considers a single NPV taken within a pool of only 10 values when calculating the NPVaR at each iteration. Therefore, luck plays a bigger role in the optimization process than with larger ω values because this single NPV, taken as-is, will not be averaged out with other cases during the NPVaR calculation. With such a small pool to choose from, the corresponding NPVaR can either be higher or lower than the reference case, depending on the randomly generated uncertain parameters. In our case, it produced a relatively high NPVaR and thus proposed a more ambitious use of geothermal energies. On the other end of the spectrum, when a bigger $\omega = 1000$ is used, the resulting NPVaR is lower at the optimum than the reference NPVar of 89.3 k\$. This time, the difference is not significant, less than 3%, as it resulted in the same design being proposed by the algorithm for a much higher computation time of 31 days. Note that without the use of the efficient calculation methods presented in Section 2 for the construction of the g-functions, the calculation time would have been significantly higher.

Small values for ω should be envisioned for quick prefeasibility studies while larger values can be used for advanced designs. From this series of experiments, it is clear that the smaller α is, the more cases that feature extreme losses become over-represented. A large ω value smoothens the resulting distribution and ensures that a lot of different combinations are considered. However, the computation time is proportional to ω and a high value can rapidly lead to impractical computation times.

Variables

H	: Borehole length [m]
L_x	: Spacing between rows of boreholes [m]
L_y	: Spacing between columns of boreholes [m]
n_b	: Number of boreholes $[-]$
n_{HP}	: Number of heat pumps $[-]$

Parameters

c_1	: Borehole drilling costs [$($/m)$]
c_2	: Geothermal vault costs [$($/borehole)$]
c_3	: Borehole commissioning costs [$($/borehole)$]
c_4	: Heat-carrying fluid costs [$($/L)$]
c_5	: Trench, volume costs [$($/m^3)$]
c_6	: Trench, width costs [$($/m)$]
c_7	: GHE commissioning costs [$($/borehole)$]
c_8	: Thermal response test costs [$($)$]
c_9	: Geothermal heat pump costs [$($/kW)$]
c_{10}	: Additional costs [$($/kW)$]
c_{11}	: Electric heating costs [$($/kW)$]
c_{12}	: Electric cooling costs [$($/kW)$]
C_f	: Fluid volumetric heat capacity [$J/(K\ m^3)$]
CF	: Cash-flows between hybrid and reference system [$($)$]
C_s	: Ground volumetric heat capacity [$J/(K\ m^3)$]
D	: Buried depth [m]
f	: Excitation function [$^{\circ}C$]
$\hat{\mathbf{f}}$	: Vector of heat fluxes [W/m]
g	: Borehole field's g-function $[-]$
$\hat{\mathbf{G}}$	: Collection of lower-triangular convolution matrices obtained using the FLS under $\tilde{q}$ [$^{\circ}C$]
i	: Discount rate [%]
$\hat{\mathbf{I}}$	: Collection of identity matrices $[-]$
k_s	: Ground thermal conductivity [$W/(m\ K)$]
NPV	: Net present value [$($)$]
$NPVaR$	: Net present value-at-risk [$($)$]
ΔT	: Difference between EWT and LWT in the heat pumps [$^{\circ}C$]
$\tilde{q}$	: Unit heat load [W/m]
Q_{Aux}	: Power consumption of auxiliary equipment [kW]
Q_B	: Power requirement of the building [kW]

Q_G	: Total ground thermal load [kW]
Q_{HP}	: Power consumption of the heat pumps [kW]
R_b	: Equivalent borehole thermal resistance [$m\ K/W$]
t	: Simulation time [s]
$\mathbf{T}_b$	: Vector of average borehole wall temperatures [$^{\circ}C$]
T_g	: Initial ground temperature [$^{\circ}C$]
$\dot{V}$	: Total volumetric flow rate [m^3/s]
$\mathbf{0}, \hat{\mathbf{0}}$	: Null matrix and vector [$-$]

Greek letters

α	: Fraction of worst cases for NPVaR [$-$]
β	: Buffer of worst cases for NPVaR [$-$]
ω	: Total population of cases for NPVaR [$-$]
Λ	: Constraint vector [$-$]

Subscripts

GHE	: Ground heat exchanger
Mec	: Mechanical systems
Ref	: Reference system

Acronyms

CAP	: Capacity
COP	: Coefficient of performance
EWT	: Entering water temperature
FLS	: Finite line source
$GCHP$	: Ground-coupled heat pump
GHE	: Geothermal heat exchanger
HP	: Heat pump
LWT	: Leaving water temperature
MIP	: Mixed-integer programming
$MINLP$	: Mixed-integer non-linear programming

CHAPITRE 7 DISCUSSION GÉNÉRALE

Les travaux réalisés dans cette thèse visaient à la fois le développement d’une méthode rapide pour le calcul des fonctions de transfert de champs de puits géothermiques, ou *g-function*, ainsi que le dimensionnement sous incertitude de systèmes géothermiques hybrides. Ces objectifs ont fait l’objet de trois publications distinctes qui ont été présentées aux Chapitres 4 à 6. Les contributions originales de ces articles, une revue critique des hypothèses et limitations derrière ces publications ainsi qu’une ouverture sur de futurs travaux de recherches seront maintenant présentées ici.

7.1 Article 1 : A Block Matrix Formulation for Efficient g-Function Construction

Dans cet article, une méthode permettant le calcul de fonctions de transfert a été présentée. Elle utilise des matrices de bloc creuses pour résoudre simultanément les problèmes de superpositions spatiales et temporelles des impulsions de chaleur émises sur une base horaire par des puits géothermiques verticaux. De plus, pour accélérer la résolution du modèle de la source linéique finie (SLF) de Claesson et Javed (2011), des polynômes de Tchebychev ont été utilisés. Ces polynômes permettent d’approximer une version dérivée par rapport au temps du modèle de la SLF, ce qui permet par la suite son intégration analytique dans le domaine du temps, le tout dans le but de s’affranchir d’une coûteuse intégration numérique de l’équation originale. De plus, une implémentation astucieuse de cette méthode permet de tirer parti des redondances dans la matrice de distance entre les forages du champ au moment d’assembler la matrice de blocs, permettant ainsi de réduire davantage les temps de calcul.

Dans l’article, on recense des temps de calcul de 0.37 seconde, 2.36 secondes et 0.46 seconde pour le calcul de la fonction de transfert de champs à 7 forages (configuration irrégulière) sur 200 ans, 30 forages (configuration irrégulière) sur 40 ans et 50 forages (configuration régulière) sur 40 ans respectivement. Ces valeurs se comparent avantageusement à des exemples présentés par d’autres auteurs ayant développé d’autres méthodologies permettant de trouver la fonction de transfert de champs géothermiques. Citons notamment :

- 1) Pasquier et Marcotte (2013) dont la méthode itérative requiert 8.79 secondes pour le calcul de la fonction de transfert d’un champ de 4 par 2 forages sur un horizon de 10 ans et 0.40 seconde pour un champ de quatre forages placés aléatoirement sur un horizon de 2 ans.
- 2) Marcotte et Pasquier (2014) dont la méthode permettant de traiter des configurations en

série, en parallèle et mixtes d'alimentation du fluide caloporteur aux forages permet le calcul d'une fonction de transfert en 1.5 seconde pour un champ de 9 forages sur un horizon de 10 ans.

3) Cimmino et Bernier (2014) dont la méthode semi-analytique requiert 28 secondes de calcul pour un champ de 3 par 2 forages, 98 secondes pour un champ de 6 par 4 forages et 433 secondes pour un champ de 10 par 10 forages (tous en configurations régulières) sur un horizon d'environ 8 ans.

4) Lamarche (2017) dont l'algorithme permet d'obtenir la fonction de transfert d'un champ de 16 forages sur 10 ans en 20 secondes environ.

5) Cimmino (2018), après avoir amélioré son algorithme semi-analytique pour tenir compte des similarités lors de l'évaluation d'une fonction de transfert, obtient la *g-function* d'un champ régulier de 7 par 7 forages sur 1000 ans en 191 secondes et celle d'un champ contenant 512 forages placés aléatoirement en 27 minutes.

6) Chen *et al.* (2019) rapportent, comme mentionné durant la revue de littérature du Chapitre 2, 90 secondes pour le calcul de la fonction de transfert d'un champ de 40 puits sur 30 ans en ayant recours à une dérivée de la SLF par rapport au temps de Fourier avec des transformées de Fourier.

Ainsi, comme on peut le voir, la méthode développée pour cette thèse est plus rapide que les autres méthodologies recensées. De plus, elle est compacte, précise, facile à implémenter dans un logiciel de programmation et traite rapidement des configurations de forages régulières et irrégulières. Cependant, elle ne permet de traiter que des cas où les forages sont alimentés en parallèle comme c'était le cas dans les travaux d'Eskilson (1987). Cela dit, cette limitation serait relativement facile à contourner en adaptant, pour une écriture de matrice par blocs, la méthodologie utilisée par Marcotte et Pasquier (2014). Dans leur article, ils additionnent à leur matrice de perturbations thermiques $\mathbf{G}$ des matrices diagonales nommées $\mathbf{D}$ et $\mathbf{L}$ faisant intervenir la vitesse de circulation du fluide caloporteur ainsi que la chaleur spécifique du fluide lors des cas où les forages sont alimentés en série ou encore dans une configuration mixte. Cette procédure devant être répétée à chacun des pas de temps de leur méthode itérative, il est envisageable de concevoir que des matrices par blocs diagonales contenant toutes les sous-matrices $\mathbf{D}$ et $\mathbf{L}$ pourraient être assemblées et additionnées à la matrice par blocs $\mathbf{G}$ utilisée dans nos travaux.

En s'inspirant davantage de la méthodologie de Marcotte et Pasquier (2014), il serait également envisageable d'adapter notre méthode pour que la fonction de transfert calculée représente directement la température du fluide à la sortie du champ de captage géothermique

au lieu de la température moyenne aux parois des forages. Ainsi, une étape de calcul serait sauvée ce qui pourrait accélérer légèrement le temps total de détermination de fonctions de transfert.

Pour finir, il pourrait être intéressant de trouver un polynôme pour remplacer les polynômes de Tchebychev pour l'approximation de la dérivée de la SLF. En effet, comme mentionné dans l'article, ces polynômes souffrent d'un coût de base en temps de calcul, et ce, peu importe la complexité du champ modélisé. Ce coût, de l'ordre de quelques dixièmes de secondes, est négligeable pour les champs comptant beaucoup de forages ou des configurations irrégulières, mais compte pour un pourcentage significatif du temps total de calcul pour les champs plus simples. Ainsi, il serait intéressant de voir si des polynômes plus simples et rapides à utiliser pour un ordinateur, comme les splines cubiques par exemple, pourraient être une alternative viable lors de la détermination de fonction de transfert pour les champs comptant peu de forages ou encore des forages segmentés.

7.2 Article 2 : Efficient g-Function Approximation with Artificial Neural Networks for a Varying Number of Boreholes on a Regular or Irregular Layout

Dans cet article, un réseau de neurones artificiels (*artificial neural network*) est créé puis entraîné pour l'approximation de fonctions de transfert selon des configurations de forages régulières et irrégulières. Ce réseau est une amélioration d'un précédent réseau présenté lors d'une conférence Dusseault *et al.* (2018) qui ne permettait de traiter que des forages placés selon des agencements réguliers. À l'aide d'un bon formatage des paramètres d'entrées couplé à l'emploi de méthodes quasi aléatoires assurant une bonne couverture des exemples utilisés pendant l'entraînement du réseau, celui-ci permet l'approximation de fonctions de transfert 37% plus rapidement qu'avec la méthode présentée au Chapitre 4 en commettant des erreurs d'approximation largement inférieures au pourcent.

En date d'aujourd'hui, il constitue l'un des quelques réseaux ayant été publiés permettant le calcul de fonctions de transfert : un autre pour le long terme (Dusseault et Pasquier, 2018) et deux pour le court terme (Pasquier et Marcotte, 2020; Pasquier *et al.*, 2018). Cela est probablement dû entre autres au fait que la base de données nécessaires à son entraînement, à cause de sa taille, n'aurait pas pu être constituée dans un temps raisonnable il y a quelques années encore avant l'avènement d'une méthode comme celle présentée au Chapitre 4. À cause de la faible prévalence des méthodes d'intelligence artificielle appliquée à cette fin et à la nouveauté de ce genre d'outils en géothermie en général, il est important de réaliser qu'il reste beaucoup de recherche à faire dans ce domaine et que le réseau neuronal présenté dans cet article n'est que le premier pas dans cette direction. Une panoplie de réseaux de

neurones découleront probablement de celui-ci au courant des années à venir, en faisant intervenir différents paramètres d'entrée, des architectures différentes et de nouvelles façons de générer des données d'entraînement. Déjà en date d'aujourd'hui, quatre améliorations sont envisageables au travail qui a été présenté au Chapitre 5 :

Tableau 7.1 Taille de la base de données d'entraînement requise en fonction du nombre de forages maximal accepté par le réseau neuronal.

Nombre de forages	Ensemble des configurations possibles	Config. uniques (sous-total)	Nombre de <i>g-functions</i>
1	1x1	1 (1)	10000
2	1x2, 2x1	1 (2)	20000
3	1x3, 3x1	1 (3)	30000
4	1x4, 2x2, 4x1	2 (5)	50000
5	1x5, 5x1	1 (6)	60000
6	1x6, 2x3, 3x2, 6x1	2 (8)	80000
7	1x7, 7x1	1 (9)	90000
8	1x8, 2x4, 4x2, 8x1	2 (11)	110000
9	1x9, 3x3, 9x1	2 (13)	130000
10	1x10, 2x5, 5x2, 10x1	2 (15)	150000
11	1x11, 11x1	1 (16)	160000
12	1x12, 2x6, 3x4, 4x3, 6x2, 12x1	3 (19)	190000
13	1x13, 13x1	1 (20)	200000
14	1x14, 2x7, 7x2, 14x1	2 (22)	220000
15	1x15, 3x5, 5x3, 15x1	2 (24)	240000
16	1x16, 2x8, 4x4, 8x2, 16x1	3 (27)	270000
17	1x17, 17x1	1 (28)	280000
18	1x18, 2x9, 3x6, 6x3, 9x2, 18x1	3 (30)	300000
19	1x19, 19x1	1 (31)	310000
20	1x20, 2x10, 4x5, 5x4, 10x2, 20x1	3 (34)	340000
21	1x21, 3x7, 7x3, 21x1	2 (36)	360000
22	1x22, 2x11, 11x2, 22x1	2 (38)	380000
23	1x23, 23x1	1 (39)	390000
24	1x24, 2x12, 3x8, 4x6, 6x4, 8x3, 12x2, 24x1	4 (43)	430000
25	1x25, 5x5, 25x1	2 (45)	450000
26	1x26, 2x13, 13x2, 26x1	2 (47)	470000
27	1x27, 3x9, 9x3, 27x1	2 (49)	490000

1) À moins d'un algorithme d'optimisation dédié qui permettrait de faire intervenir des contraintes très précises liées au terrain, il est peu probable que les capacités d'approximation du réseau neuronal servant aux configurations irrégulières soient utilisées lors d'un processus de dimensionnement. Ainsi, à cause du pourcentage important consacré aux cas

irréguliers (350000 cas sur 500000), la base de données servant à l'entraînement pourrait être réduite substantiellement, diminuant ainsi le temps d'entraînement. Autrement, à capacité informatique égale, ces 500000 fonctions de transfert, prises uniquement à travers des configurations régulières, pourraient être utilisées pour couvrir un plus grand nombre variable de forages. D'après la progression suivie au Tableau 7.1 et à raison de 10000 exemples par configurations régulières possibles, on parle d'un nombre variant entre 1 et 27 forages au lieu de simplement 1 à 10 forages.

2) Récemment, un nouveau paradigme a commencé à émerger qui préconise l'apprentissage en continu des réseaux neuronaux après la phase initiale d'entraînement en fonction de nouvelles données sans compromettre les apprentissages initiaux. Le *continual learning* comme il l'est appelé dans les travaux de Aljundi (2019) sur le sujet, est un pas de plus vers la création de réseaux qui auront la capacité d'apprentissage des êtres humains. Ce genre de réseau pourrait représenter une alternative intéressante au réseau développé dans l'article du Chapitre 5 pour diminuer le temps d'entraînement initial qui se chiffre présentement en semaines. Bien sûr, cela augmenterait le niveau de complexité du réseau neuronal et nécessiterait qu'il soit couplé à une méthode secondaire de détermination de fonctions de transfert lorsqu'il n'est pas capable de procéder à une approximation, mais il pourrait malgré tout s'agir d'une piste intéressante à explorer lors de futurs travaux.

3) Si les cas irréguliers de configurations de forages reste un enjeu intéressant lors de futurs travaux, l'utilisation de réseaux neuronaux convolutifs (*convolutional neural network*) pourrait être envisagée. Ces réseaux, largement utilisés pour la reconnaissance vocale et d'image, permettent de traiter de grands jeux de données simultanément (donc des matrices de données) en adoptant une architecture où chaque neurones n'est pas complètement connectée avec l'ensemble des neurones de la couche suivante. En traitant comme des images la localisation des puits et en utilisant des couches conventionnelles, il serait possible de traiter plus rapidement et facilement les cas irréguliers à l'intérieur de grandes bloc matrices creuses contenant peu de redondance.

4) Pour finir, si jamais une méthode comme celle par matrices de blocs du Chapitre 4 était améliorée pour tenir compte de différentes configurations de fluide caloporteur en fournissant une fonction de transfert basée sur la température du fluide au lieu de la température aux parois du forage, il deviendrait alors possible de créer une base de données d'entraînement pour que de telles possibilités d'approximation soient transmises à un réseau neuronal. Le réseau pourrait ainsi fournir la fonction de transfert de la température d'entrée d'eau à la thermopompe comme l'ont fait pour un seul puits Pasquier *et al.* (2018). Il est possible que la complexité du réseau résultant augmente tellement que celui-ci ne puisse plus être entraîné

dans un temps raisonnable ou que le nombre maximal de forages pouvant être considérés doivent être réduits, mais cela mérite d'être étudié lors de futurs travaux.

7.3 Article 3 : Usage of the Net-Present-Value-at-Risk to Design Ground Source Heat Pump Systems Under Uncertain Scenarios

Dans cet article, une étude de cas est réalisée portant sur l'utilisation d'un nouvel algorithme de dimensionnement stochastique pour les systèmes géothermiques hybrides en boucles fermées verticales. Cet algorithme, inspiré d'indicateurs financiers utilisés pour l'assemblage de portefeuilles d'investissement en finance (Rockafellar et Uryasev, 1999) et la construction d'ouvrages en génie civil (Ye et Tiong, 2000), propose un design en se basant sur l'espérance de la valeur actuelle nette d'un projet advenant les pires scénarios envisageables (VANaR). La VANaR est nouvellement utilisée en géothermie dans le cadre de cet article et constitue une première dans le domaine pour l'intégration du risque durant la phase de dimensionnement. Cette façon de procéder contraste avec l'approche habituelle qui consiste à quantifier l'impact d'une variation des paramètres incertains (d'où provient le risque) une fois le processus d'optimisation terminé en utilisant des analyses de sensibilité.

La VANaR nécessitant le calcul itératif de la valeur actuelle nette à chaque itération de l'algorithme d'optimisation, cet indicateur n'aurait pas pu être utilisé dans un temps raisonnable s'il n'utilisait pas les méthodes développées durant les deux premiers articles pour le calcul des fonctions de transfert. En effet, comme l'a remarqué Hénault (2015) dans son mémoire, le calcul de l'ensemble des fonctions de transfert nécessaires à un processus de dimensionnement par optimisation peut consommer jusqu'à 75% du temps total de calcul.

En plus de l'utilisation de la VANaR, cet article se démarque dans la littérature actuelle de trois autres façons différentes. Premièrement, il se situe dans la courte lignée des travaux recensés au Chapitre 2 utilisant la programmation en nombres entiers pour rendre compte que certaines variables de design, le nombre de forages et de thermopompes en l'occurrence, ne peuvent logiquement pas être fractionnaires à l'optimum. Deuxièmement, il corrobore une des conclusions de Hénault *et al.* (2016) impliquant que le nombre de forages et le nombre de thermopompes sont les variables de design qui ont la plus grande influence sur le design d'un système géothermique hybride. Pour finir, en fournissant la preuve à l'aide d'une figure présentant deux balayages, il démontre hors de tous doutes que le sous-dimensionnement de l'échangeur de chaleur souterrain d'un système géothermique hybride est bénéfique lorsque les conditions externes au projet sont incertaines.

Malgré les retombées positives attendues à la suite de la publication de cet article, il est

évident que certaines des hypothèses posées lors de son écriture nécessiteront des travaux additionnels pour lever des questionnements. Notamment, bien qu'elle ait été une variable de design, la profondeur des puits H n'a jamais été sélectionnée par l'algorithme d'optimisation à une valeur inférieure à la valeur maximale autorisée. Cela est peut-être dû au fait que la gamme de variations permises pour H était trop restrictive, mais peut-être aussi que cela témoigne que cette variable pourrait être un paramètre à l'avenir (posé en fonction des équipements de forages disponibles par exemple), réduisant ainsi le temps total de calcul.

Aussi, les appareils auxiliaires considérés fonctionnent à l'électricité. Cette hypothèse de travail est logique dans le contexte québécois où l'électricité est à la fois abondante et peu dispendieuse, mais pourrait être rejetée dans d'autres régions. Même au Québec, bien que l'usage de l'électricité soit largement répandu pour le chauffage résidentiel, commercial et industriel, de nombreux ménages, commerces et entreprises utilisent le gaz naturel. Utiliser le gaz naturel, dont le prix est concurrentiel à l'électricité par exemple, dans l'algorithme développé aurait nécessité l'emploi d'une nouvelle valeur incertaine, cette fois sujette à des variations quotidiennes en fonction des fluctuations du marché. Cette façon de procéder n'a pas été considérée dans le présent travail entre autres à cause de la complexité de son implémentation. Cela dit, dans le cadre de travaux futurs situés à nouveau dans le contexte québécois, son introduction devrait être considérée et implémentée en balançant le déploiement de chaudières au gaz bon marché avec le coût attendu des crédits carbone du système de plafonnement et d'échange de droits d'émission (SPEDE) ou par une deuxième fonction objectif qui minimiserait les émissions de gaz à effet de serre.

Pour finir, il pourrait être intéressant de revisiter le solveur utilisé dans cet article pour résoudre le problème d'optimisation. Bien qu'il ait été sélectionné pour ses qualités mathématiques et pour la facilité de son implémentation dans l'environnement de programmation Matlab, il est également vrai qu'il l'a été pour sa fiabilité dans de précédents travaux effectués par notre groupe de recherche. Le même phénomène de l'utilisation *de facto* des outils d'optimisation habituels se produisant sûrement dans la majorité des groupes de recherches, il est probable que des outils plus performants et mieux adaptés au problème à résoudre existent dans d'autres branches de la recherche (mathématique, recherche opérationnelle, etc.) qui demeurent inconnus des chercheurs en géothermie. Maintenant que le problème de l'optimisation des systèmes géothermiques hybrides sous incertitude avec la VANaR a été étudié une première fois, l'avenir de l'optimisation de cet indicateur financier repose peut-être dans l'utilisation d'une méthode heuristique ou métaheuristique créée sur mesure pour en accélérer le calcul.

CHAPITRE 8 CONCLUSION ET RECOMMANDATION

Le dimensionnement d'un système géothermique hybride en présence d'incertitudes comporte un lot d'enjeux techniques et technologiques qui peuvent facilement compromettre la rentabilité d'un tel système. En effet, un sous-dimensionnement oblige l'achat d'équipements additionnels une fois l'installation terminée qui se rajoute aux coûts initiaux en plus d'augmenter les frais d'exploitation mensuels. Au contraire, un surdimensionnement, qu'il soit prévu pour des raisons de sécurité ou fortuit et découlant d'erreurs lors du design, engendre un système fonctionnel, mais dont le temps de retour sur investissement risque de se chiffrer en décennies. Présentement, les ingénieurs en industrie disposent de peu de moyens efficaces pour tenir compte des nombreuses sources d'incertitudes qui accompagnent naturellement ce genre de projet. Le but de cette thèse était de proposer une méthodologie permettant d'intégrer quatre des principales sources d'incertitude (la demande en chauffage et en climatisation d'un bâtiment, l'évolution du coût de l'électricité, la conductivité thermique du sol et les coûts de construction initiaux) et d'étudier l'impact de son implémentation par rapport à une méthode de design traditionnelle. Pour y arriver, trois sous-objectifs ont été ciblés en début de projet.

8.1 Synthèse des travaux

Premièrement, il a été constaté dans des travaux antérieurs (Hénault, 2015; Hénault *et al.*, 2016) que le temps de calcul des fonctions de transfert des champs géothermiques est un enjeu de taille lors du design par optimisation de systèmes hybrides. La première contribution originale de cette thèse a donc été de concevoir une nouvelle méthode pour obtenir ces fonctions de transfert plus rapidement qu'avec les méthodes disponibles. Cet objectif a été atteint par l'utilisation conjointe de matrices par bloc pour reformuler le problème de la superposition spatiale et temporelle des transferts de chaleur dans le sol et par l'utilisation de polynômes de Tchebychev pour accélérer l'intégration du modèle de la source linéique finie utilisé pour représenter le comportement des forages dans le sol. Ces travaux permettent de calculer une fonction de transfert en un temps variant de quelques dixièmes de secondes à moins de trois secondes sur une période de 40 ans en fonction de la complexité du champ de captage géothermique considéré.

La deuxième contribution originale de cette thèse a été de concevoir et d'entraîner un réseau neuronal pour accélérer davantage l'obtention des fonctions de transfert. Grâce à la rapidité de la méthode développée pour le premier objectif, une base de données de 500000 fonctions de

transfert a pu être constituée permettant l'entraînement du réseau de neurones développé. À ce stade, l'obtention des fonctions de transfert est devenue si rapide que le produit de convolution utilisé pour obtenir l'évolution horaire des températures aux parois des forages du champ consomme 19 fois plus de temps que l'obtention de la fonction de transfert, un gain qui était impensable il y a quelques années à peine. Lorsque l'on met ce résultat en perspective avec l'efficacité reconnue des produits de convolution dans le domaine spectral, on comprend l'étendue des gains qui ont été réalisés dans le domaine de l'obtention des fonctions de transfert.

Pour finir, un algorithme permettant le dimensionnement de systèmes géothermiques hybrides en boucles fermées verticales sous incertitude a été développé. Il utilise conjointement un algorithme d'optimisation non linéaire et une méthode de programmation en nombre entier, le *branch and bound*, pour éviter qu'il ne propose des designs où les forages et les thermopompes sont fractionnaires. De plus, et c'est là la principale contribution de ce travail, il utilise une fonction objectif stochastique appelée la valeur actuelle nette à risque (VANaR), inspirée de travaux réalisés en finance et en génie civil, qui permet de concevoir un système en considérant simultanément les revenus ou les pertes attendues lors des pires scénarios envisageables. Cette façon de procéder possède l'avantage d'incorporer directement lors de la phase de design les distributions de probabilités des paramètres incertains pour permettre leur mitigation durant le processus d'optimisation. Cette approche contraste avec les méthodes habituelles qui se contentent généralement de procéder à un dimensionnement puis à évaluer après coup l'impact potentiel des incertitudes sur l'optimum en procédant à des analyses de sensibilité. Ces travaux ont démontré que la VANaR permet un dimensionnement sécuritaire des systèmes hybrides en réduisant grandement les pertes potentiellement encourues lors de conditions adverses tout en assurant un bon retour sur investissement lors de conditions favorables.

8.2 Améliorations futures

Cette thèse, au mieux de la connaissance de l'auteur, constitue le premier travail visant à intégrer et à mitiger à la source le risque dans le domaine de la géothermie. Par conséquent, il reste encore beaucoup de travail pour que cette philosophie de design se démocratise et rejoigne les ingénieurs dans leur pratique professionnelle. À cause du temps considérable que requiert le dimensionnement par optimisation stochastique d'un système géothermique hybride, plusieurs étapes devront être franchies pour qu'il soit réduit à des niveaux raisonnables pour l'industrie. Dans un premier temps, les conditions dans lesquelles un réseau neuronal peut être utilisé pour l'approximation de fonctions de transfert devront être éten-

dues pour couvrir les variations de davantage de paramètres et un plus grand nombre de forages. Deuxièmement, l'algorithme d'optimisation devra être revisité pour permettre de converger plus rapidement vers un optimum. Pour finir, il faudrait que l'algorithme puisse permettre de considérer un système de chauffage utilisant les énergies fossiles comme système de référence. Sans ces étapes franchies, il est peu probable que les travaux effectués dans cette thèse sortent un jour d'un cadre purement académique.

Bibliographie

- Adnot J., Marchio D. et Rivière P. (2012). Cycles de vie des systèmes énergétiques. *Presses des MINES, collection les cours édition*.
- Alavy, M., Nguyen, H.V., Leong, W.H., et Dworkin, S.B. (2013). A methodology and computerized approach for optimizing hybrid ground source heat pump system design. *Renewable Energy*, 57, 404–412.
- Aljundi, R. (2019). Continual Learning in Neural Networks. Ithaca, NY, USA : Cornell University.
- Arat, H., et Arslan, O. (2017). Optimization of district heating system aided by geothermal heat pump : A novel multistage with multilevel ANN modelling. *Applied Thermal Engineering*, 111, 608–623.
- ASHRAE. (2011). Chapter 34 : Geothermal Energy. In ASHRAE (dir.), *ASHRAE Handbook - HVAC Applications*, (pp. 34.31-34.20). Atlanta : American Society of Heating, Refrigerating and Air-Conditioning Engineers.
- Atam, E., et Helsen, L. (2016a). Ground-coupled heat pumps : Part 1 — Literature review and research challenges in modeling and optimal control. *Renewable and Sustainable Energy Reviews*, 54, 1653-1667.
- Atam, E., et Helsen, L. (2016b). Ground-coupled heat pumps : Part 2—Literature review and research challenges in optimal design. *Renewable and Sustainable Energy Reviews*, 54, 1668–1684.
- Ayala, H.V.H., Keller, P., Morais, M.F., Mariani, V.C., Coelho, L.S., et Rao, R.V. (2016). Design of heat exchangers using a novel multiobjective free search differential evolution paradigm. *Applied Thermal Engineering*, 94, 170–177.
- Bandos, T.V., Montero, A., Fernández, E., Santander, J.L.G., Isidro, J.M., Pérez, J., Fernández de Córdoba, P.J., et Urchueguía, J.F. (2009). Finite line-source model for borehole heat exchangers : Effect of vertical temperature variations. *Geothermics*, 38(2), 263–270.
- Banks, D. (2012). An Introduction to Thermogeology : Ground Source Heating and Cooling, 2nd Edition. *John Wiley & Sons*.

- Bayer, P., de Paly, M., et Beck, M. (2014). Strategic optimization of borehole heat exchanger field for seasonal geothermal heating and cooling. *Applied Energy*, 136, 445–453.
- Belotti, P., Kirches, C., Leyffer, S., et Linderoth, J. (2013). Mixed-integer nonlinear optimization. *Acta Numer.*, 22, 1–131.
- Bennet, J., Claesson, J., et Hellstrom, G. (1987). *Multipole method to compute the conductive heat flows to and between pipes in a composite cylinder*, Tiré de Lund Institute of Technology.
- Bernard, J.T., Bolduc, D., et Yameogo N.D. (2011). A pseudo-panel data model of household electricity demand. *Resource and Energy Economics*, 33(1), 315–325.
- Bernier, M. (2000). A Review of the Cylindrical Heat Source Method for the Design and Analysis of Vertical Ground-Coupled Heat Pump Systems. *4th International Conference on Heat Pumps in Cold Climates*.
- Bernier, M. (2001). Ground-coupled Heat pump system simulation. *ASHRAE Transactions*, 107(1), 605–616.
- Bernier, M. (2006). Closed-loop ground-coupled heat pump systems. *ASHRAE Journal*, 48(9), 12–25.
- Bernier, M., et Salim Shirazi, A. (2007). Solar Heat Injection into Boreholes : a Preliminary Analysis. *2nd Canadian Solar Buildings Conference*, Calgary.
- Bernier, M., Pinel, P., Labib, R., et Paillot, R. (2004). A Multiple Load Aggregation Algorithm for Annual Hourly Simulations of GCHP Systems. *HVAC and R Res*, 10 :4, 471–487.
- Bernier, M., Chahla, A., et Pinel, P. (2008). Long-Term Ground-Temperature Changes in Geo-Exchange Systems. *ASHRAE Transactions*, 114(2), 342–350.
- Bourry, F., Juban, J., Costa, L., et Kariniotakis, G. (2008). Advanced strategies for wind power trading in short-term electricity markets. *European Wind Energy Conference and Exhibition EWEC 2008*, Brussels, Belgium.
- Broyden, C.G. (1967). Quasi-Newton Methods and their Application to Function Minimisation. *Mathematics of Computation*, 21(99), 368–381.
- Cannavó, F. (2012). Sensitivity analysis for volcanic source modeling quality assessment and model selection. *Computers and Geosciences*, 44, 52–59.

- Caputo, A.C., Pelagagge, P.M., et Salini, P. (2008). Heat exchanger design based on economic optimisation. *Applied Thermal Engineering*, 28(10), 1151–1159.
- Carslaw, H.S., et Jaeger, J.C. (1946). Chapter 2, Linear flow of heat : the infinite and semi-infinite solid. In O. U. Press (dir.), *Conduction of Heat in Solids*. (2nde éd., pp. 50-91). Oxford : Oxford University.
- CCEG. (2012). État de l'industrie canadienne de la géothermie 2011. Analyse du marché et enquête de l'industrie. *Rapport technique, Coalition canadienne de l'énergie géothermique*.
- Chapuis S. (2009). Stockage thermique saisonnier dans un champ de puits géothermiques verticaux en boucle fermée. *Mémoire de maîtrise*, École Polytechnique de Montréal.
- Charitopoulos, V.M., et Dua, V. (2017). A unified framework for model-based multi-objective linear process and energy optimisation under uncertainty. *Applied Energy*, 186(3), 539–548.
- Cho, W.J., Kwon, S., et Choi, J.W. (2009). The thermal conductivity for granite with various water contents. *Engineering Geology*, 107 (3 - 4), 167-171.
- Clausen, J. (1999). Branch and Bound Algorithms — Principles and Examples (Technical report). *Department of Computer Science* : University of Copenhagen.
- Clauser, C., et Huenges, E. (1995). Thermal conductivity of rocks and minerals. *Rock Physics and Phase Relations : A Handbook of Physical Constants. T. J. Ahrens Ed.*, 105 - 126.
- Chen, Y., Pan, B., Zhang, X., et Du, C. (2019). Thermal response factors for fast parameterized design and long-term performance simulation of vertical GCHP systems. *Renewable Energy*, 136, 793-804.
- Chikhaoui, M. (2015). Gestion de risque de portefeuille : estimation de la VaR et CVaR. *Probabilités* [math.PR]. HAL Id : hal-01246153.
- 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. (2018). 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(2), 1-14.
- Cimmino, M. (2019). g -Functions of bore fields with series- and parallel-connected boreholes. *Science and Technology for the Built Environment*, 25(2), 1-40.

- Cimmino, M., et Bernier, M. (2014). A semi-analytical method to generate g-functions for geothermal bore fields. *International Journal of Heat and Mass Transfer*, 70, 641-650.
- Cimmino, M., et Bernier, M. (2015). Experimental determination of the g-functions of a small-scale geothermal borehole. *Geothermics*. 56, 60-71.
- Cimmino, M., Bernier, M., et Pasquier, P. (2012). Utilisation des g-functions de Eskilson pour la simulation de systèmes géothermiques. *Proceedings of aSim 2012, Halifax*, NS, 282-295.
- Cimmino, M., Bernier, M., et Adams, F. (2013) A contribution towards the determination of g-functions using the finite line source. *Applied Thermal Engineering*, 51(1-2), 401-412.
- Claesson, J., et Bennet, J. (1987). *Multipole method to compute the conductive heat flows to and between pipes in a cylinder*, Tiré de Lund Institute of Technology.
- Claesson. J., et Hellstrom, G. (2011). Multipole method to calculate borehole thermal resistances in a borehole heat exchanger. *HVAC&R Research*, 17(6), 895-911.
- Claesson, J., et Javed, S. (2011). An Analytical Method to Calculate Borehole Fluid Temperatures for Time-scales from Minutes to Decades. *ASHRAE Transactions*, 117 :2, 279-288.
- Claesson, J., et Javed, S. (2012). A load-aggregation method to calculate extraction temperatures of borehole heat exchangers. *ASHRAE Transactions*, 118(1), 530-539.
- Coleman, T.F., et Li, Y. (1996). An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal on Optimization*, 6(2), 418-445.
- Cooper, L.Y. (1976). Heating of a cylindrical cavity. *International Journal of Heat and Mass Transfer*, 19, 575-577.
- Courchesne-Tardif, A. (2011). Conception et optimisation de systèmes énergétiques hybrides pour communautés durables. Mémoire de maîtrise, École Polytechnique de Montréal.
- Cukier, R.I., Fortuin, C.M., et Shuler, K.E. (1973). Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *The Journal of Chemical Physics*, 59 : 3873.
- Cui, P., Yang, H., et Fang, Z. (2006). Heat transfer analysis of ground heat exchangers with inclined boreholes. *Applied Thermal Engineering*, 26(11-12), 1169-1175.
- Cui, W., Zhou, S., et Liu, X. (2015). Optimization of design and operation parameters for hybrid ground-source heat pump assisted with cooling tower. *Energy and Buildings*, 99, 253-262.

- Cullin, J.R., Spitler, J.D., Montagud, C., Ruiz-Calvo, F., Rees, S.J., Naicker, S.S., Konecný, P., et Southard, L.E. (2015). Validation of vertical ground heat exchanger design methodologies. *Science and Technology for the Built Environment*, 21(2), 137–149.
- Curry, H.B. (1944). The Method of Steepest Descent for Non-Linear Minimization Problems. *Quart. Appl. Math.*, 2, 258-261.
- Cybenko, G. (1989). Approximation by Superpositions of Sigmoidal Functions. Budapest : Eötvös Loránd University.
- Daniilidis, A., Alpsoy, B., et Herber, R. (2017). Impact of technical and economic uncertainties on the economic performance of a deep geothermal heat system. *Renewable Energy*, 114 (Part B), 805 - 816.
- Das, S., et Suganthan, P.N. (2011). Differential evolution : a survey of the state-of-the-art. *IEEE Transactions on evolutionary computation*, 15(1).
- De Rosa, M., Ruiz-Calvo, F., Corberán, J.M., et Montagud, C. (2015). A novel TRNSYS type for short-term borehole heat exchange rsimulation : B2G model. *Energy Conversion and Management*, 100, 347–357.
- Diao, N.R., Zeng, H.Y., Fang, Z.H. (2004). Improvement in modeling of heat transfer in vertical ground heat exchangers. *HVAC&R Research*, 10(4), 459–470.
- Driscoll, T.A., Hale, N., et Trefethen, L.N. (2014). Chebfun Guide. *Pafnuty Publications*, Oxford.
- Dusseault, B., Pasquier, P. et Marcotte, M. (2017). Assessment of the gradient of an objective function by analytical derivation for optimization-based design of ground-coupled heat pump systems. *Proceedings of the IGSHPA Technical/Research Conference and Expo 2017*, Denver, CO.
- Dusseault, B., et Pasquier, P. (2018). Near-instant g-function construction with artificial neural networks. *Proceedings of the IGSHPA Research Track 2018*, Stockholm, Sweden, September 18-20.
- Dusseault, B., et Pasquier, P. (2019). Efficient g-function approximation with artificial neural networks for a varying number of boreholes on a regular or irregular layout. *Science and Technology for the Built Environment*, 25(8), 1023-1035.

- Dusseault, B., et Pasquier, P. (2020). Usage of the Net Present Value-at-Risk to Design Ground-Coupled Heat Pump Systems Under Uncertain Scenarios. *Submitted to Applied Energy*.
- Dusseault B., Pasquier P., et Marcotte D. (2018). A block matrix formulation for efficient g-function construction. *Renewable Energy*, 121, 249 - 260.
- Ehrgott, M. (2005). Multicriteria Optimization. Germany, Berlin : Springer.
- Erdeweghe, S.V., Van Bael, J., Laenen, B., et D'haeseleer, W. (2018). Feasibility study of a low-temperature geothermal power plant for multiple economic scenarios. *Energy*, 155, 1004-1012.
- Erol. S., et François, B. (2017). Multilayer analytical model for vertical ground heat exchanger with groundwater flow. *Geothermics*, 71.
- Escalante Soberanis, M.A., Mithrushi, T., Bassam, A., et Mérida, W. (2018). A sensitivity analysis to determine technical and economic feasibility of energy storage systems implementation : A flow battery case study. *Renewable Energy*, 115, 547 - 557.
- Esen, H., Inalli, M., Sengur, A., et Esen, M. (2008). Performance prediction of a ground-coupled heat pump system using artificial neural networks. *Expert Systems with Applications*, 35(4), 1940–1948.
- Eskilson, P. (1987). Thermal Analysis of Heat Extraction Boreholes. Lund, Sweden : University of Lund.
- Fannou, J.-L.C., Rousseau, C., Lamarche, L., et Kajl, S. (2014). Modeling of a direct expansion geothermal heat pump using artificial neural networks. *Energy and Buildings*, 81, 381–390.
- Fernández, M., Eguía, P., Granada, E., et Febrero, L. (2017). Sensitivity analysis of a vertical geothermal heat exchanger dynamic simulation : Calibration and error determination. *Geothermics*, 70, 249-259.
- Gang, W., Wang, J., et Wang, S. (2014). Performance analysis of hybrid ground source heat pump systems based on ANN predictive control. *Applied Energy*, 136, 1138–1144.
- Ghanei, A., Assareh, E., Biglari, M., Ghanbarzadeh, A., et Noghrehabadi, A.R. (2014). Thermal-economic multi-objective optimization of shell and tube heat exchanger using particle swarm optimization (PSO). *Heat and Mass Transfer*, 50(10), 1375–1384.

- Ghaffari, S., Aghajani, G., Noruzi, A., et Mehr, H.H. (2016). Optimal economic load dispatch based on wind energy and risk constrains through an intelligent algorithm. *Complexity*, 21(S2), 494–506.
- Gehlin, S. (2002). Thermal response test method development and evaluation. Lulea, Suède : Lulea University of Technology.
- Gosselin, L., Tye-Gingras, M., et Mathieu-Potvin, F. (2009). Review of utilization of genetic algorithms in heat transfer problems. *International Journal of Heat and Mass Transfer*, 52(9–10), 2169–2188.
- Grossmann, I.E., Viswanathan, J., Vecchietti, A., Raman, R., et Kalvelagen, E. (2003). DICOPT. Logiciel distribué par GAMS.
- Guo, H., et Meggers, F. (2017). Geothermal distribution network modeled as Heat Exchanger Network to be optimized with Mixed Integer Nonlinear Programming. *Energy Procedia*, 122, 1105–1110.
- Hackel, S., Nellis, G., et Klein, S. (2008). Optimization of hybrid geothermal heat pump systems. *9th International IEA Heat Pump Conference*, 20–22 May 2008. Zürich, Switzerland.
- Hackel, S., Nellis, G., et Klein, S. (2009). Optimization of cooling-dominated hybrid ground-coupled heat pump systems. *ASHRAE transactions*, 115(1), 565–580.
- Hadidi, A. (2015). A robust approach for optimal design of plate fin heat exchangers using biogeography based optimization (BBO) algorithm. *Applied Energy*, 150, 196–210.
- Hagan, M.T., Demuth, H.B., Beale, M.H., et De Jesus, O. (1996). Neural Network Design. 2nd Ed. Stillwater : Oklahoma State University.
- Halton, J.H. (1964). Algorithm 247 : Radical-inverse quasi-random point sequence. *Communications of the ACM*, 7(12), 701–702.
- Hecht-Méndez J., Molina-Giraldo N., Blum P., et Bayer P. (2010). Evaluating MT3DMS for heat transport simulation of closed geothermal systems. *Ground Water*, 48(5), 741–756.
- Hellström, G. (1989). Duct ground heat storage model : Manual for computer code. Tiré de University of Lund, Department of Mathematical Physics.
- Hellström, G. (1991). Ground heat storage : Thermal analysis of duct storage systems. PhD, University of Lund, Lund, Sweden.

- Hénault, B. (2015). Éléments de conception d'un système géothermique hybride par optimisation financière. Mémoire de maîtrise, École Polytechnique de Montréal.
- Hénault, B., Pasquier, P., et Kummert, M. (2016). Financial optimization and design of hybrid ground-coupled heat pump systems. *Applied Thermal Engineering*, 93, 72-82.
- Hestenes, M.R., et Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of research of the national bureau of standards*, 49(6), 409-436.
- Holl, M., et Pelz, P.F. (2016). Multi-pole system analysis (MPSA) – A systematic method towards techno-economic optimal system design. *Applied Energy*, 169, 937–949.
- Hornik, K., Stinchcombe, M., et White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366.
- Hu, J. (2017). An improved analytical model for vertical borehole ground heat exchanger with multiple-layer substrates and groundwater flow. *Applied Energy*, 202, 537-549.
- Huang, S., Ma, Z., et Wang, F. (2015). A multi-objective design optimization strategy for vertical ground heat exchangers. *Energy and Buildings*, 87, 233-242.
- Ingersoll L.R., et Plass H.J. (1948). Theory of the ground pipe source for the heat pump. *ASHRAE Transaction*, 54, 339–348.
- Ingersoll, L.R., Adler, F.T., Plass, H.J., et Ingersoll, A.C. (1950). Theory of earth heat exchangers for the heat pump. Heating. *Piping & Air Conditioning*, 22, 113-122.
- Ingersoll L., Zobel O., et Ingersoll A. (1954). Heat Conduction with Engineering geological and other applications. *Édition McGraw-Hill*, NewYork.
- Jansen, M.J.W. (1999). Analysis of variance designs for model output. *Comput. Phys. Commun.*, 117, 35 - 43.
- Jinggang, W., Xiaoxia, G., Zhenjiang, Y., et Fang, L. (2009). Study of Operating Control Strategies for Hybrid Ground Source Heat Pump System with Supplemental Cooling Tower. *IEEE*, Zhangjiajie, China, 507–510.
- Johansen, O. (1977). Thermal Conductivity of Soils, *Univ. of Trondheim*, Trondheim, Norway.
- Kavanaugh, S.P., et Rafferty, K. (1997). Ground-Source Heat Pumps-Design of Geothermal Systems for Commercial and Institutional Buildings. *ASHRAE Book*. 167 pages.

- Khodabakhsh, R., et Sirouspour, S. (2016). Optimal Control of Energy Storage in a Microgrid by Minimizing Conditional Value-at-Risk. *IEEE Transactions on Sustainable Energy*, 7(3), 1264-1273.
- Klein K., Huchtemann K. et Müller D. (2014). Numerical study on hybrid heat pump systems in existing buildings. *Energy and Buildings*, 69, 193–201.
- Lamarche, L. (2009). A fast algorithm for the hourly simulations of ground-source heat pumps using arbitrary response factors. *Renewable Energy*, 34(10), 2252-2258.
- Lamarche, L. (2011). Analytical g-function for inclined boreholes in ground-source heat pump systems. *Geothermics*, 40(4), 241-249.
- Lamarche, L. (2017). Mixed arrangement of multiple input-output borehole systems. *Applied Thermal Engineering*, 124, 466-476.
- Lamarche, L., et Beauchamp, B. (2007). A new contribution to the finite line-source model for geothermal boreholes. *Energy and Buildings*, 39(2), 188-198.
- Lamarche, L., et Beauchamp, B. (2007). New solutions for the short-time analysis of geothermal vertical boreholes. *International Journal of Heat and Mass Transfer*, 50(7-8), 1408-1419.
- Lamarche, L. et Pasquier, P. (2019). Higher-order temporal scheme for ground heat exchanger analytical models. *Geothermics*, 78, 111-117.
- Law, Y.L.E., et Dworkin S.B. (2016). Characterization of the effects of borehole configuration and interference with long term ground temperature modelling of ground source heat pumps. *Applied Energy*, 179, 1032-1047.
- Lawler, E.L., et Wood, D.E. (1966). Branch-and-Bound Methods : A Survey. *Operations Research*, 14 (4), 699 - 719.
- 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, 1380-1393.
- Lazzarotto, A. (2016). Thermal modeling of bore fields with arbitrarily oriented boreholes. *EGU General Assembly 2016*, held 17-22 April, 2016 in Vienna Austria.

- Lazzarotto, A., et Björk, F. (2016). A methodology for the calculation of response functions for geothermal fields with arbitrarily oriented boreholes : Part 2. *Renewable Energy*, 86, 1353-1361.
- Li, B., Hu, P., Zhu, N., Lei, F., et Xing, L. (2019). Performance analysis and optimization of a CCHP-GSHP coupling system based on quantum genetic algorithm. *Sustainable Cities and Society*, 46, Article 101408.
- Liu, X. (2005). Development and Experimental Validation of Simulation of Hydronic Snow Melting Systems for Bridges. Stillwater : Oklahoma State University.
- Malayappan, V., et Spitler, J.D. (2013). Limitations of using uniform heat flux assumptions in sizing vertical borehole heat exchanger fields. *Proceedings of CLIMA 2013*, Prague, Czech Republic.
- Marcotte, D., et Pasquier, P. (2008). Fast fluid and ground temperature computation for geothermal ground-loop heat exchanger systems. *Geothermics*, 37(6), 651-665.
- Marcotte, D., et Pasquier, P. (2009). The effect of borehole inclination on fluid and ground temperature for GLHE systems. *Geothermics*, 38(4), 392-398.
- Marcotte, D., et Pasquier, P. (2014). Unit-response function for ground heat exchanger with parallel, series or mixed borehole arrangement. *Renewable Energy*, 68, 14-24.
- Marcotte, D., Pasquier, P., Sheriff, F., et Bernier, M. (2010). The importance of axial effects for borehole design of geothermal heat-pump systems. *Renewable Energy*, 35(4), 763-770.
- MathWorks. 2015. *Neural Network Toolbox : User's Guide (R2015b)*. <https://www.mathworks.com/help/releases/R2015b/nnet/index.html>.
- Quebec's Ministry of Energy and Natural Resources. (2016), <https://mern.gouv.qc.ca/energie/statistiques/statistiques-energie-prix-electricite.jsp>
- Philippe, M., Bernier, M., et Marchio, D. (2010). Sizing calculation spreadsheet - vertical geothermal borefields. *ASHRAE Journal*, 52(7), 20-28.
- Millette, J., Sansregret, S., et Daoud, A. (2011). Simeb : Simplified interface to DOE2 and EnergyPlus - A User's perspective - Case study of an existing building. *Proceedings of Building Simulation : 12th Conference of International Building Performance Simulation Association*, Sydney.

- Mizutani, F.T., Pessoa, F.L.P., et Queiroz, E.M. (2003). Mathematical Programming Model for Heat-Exchanger Network Synthesis Including Detailed Heat-Exchanger Designs. 2. Network Synthesis. *Ind. Eng. Chem. Res.*, 42(17), 4019–4027.
- Molina-Giraldo, N., Blum, P., Zhu, K., Bayer, P., et Fang, Z. (2011). A moving finite line source model to simulate borehole heat exchangers with groundwater advection. *International Journal of Thermal Sciences*, 50(12), 2506–2513.
- Montmarquette, C., et Scott, I. (2007). Discount rate for the evaluation of public spending in Quebec. Research report. *Centre for Interuniversity Research and Analysis on Organizations* (CIRANO), Montreal.
- Monzó, P., Acuña, J., Fossa, M., et Palm, B. (2013). Numerical generation of the temperature response factors for a borehole heat exchanger field. *European Geothermal Congress 2013*, Pisa, Italy : 1-8.
- Monzó P., Mogensen P., et Acuña J. (2014). A novel numerical model for the thermal response of borehole heat exchanger fields. *11th IEA Heat pump conference*, Montréal.
- Monzó, P., Puttige, A.R., Acuña, J., Mogensen, P., Cazorla, A., Rodríguez, J., Montagud, C., et Cerdeira, F. (2018). Numerical modeling of ground thermal response with borehole heat exchangers connected in parallel. *Energy Build*, 172, 371-384.
- Morris, M.D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33, 161 - 174.
- Najib, A., Zarrella, A., Narayanan, V., Grant, P., et Harrington, C. (2019). A revised capacitance resistance model for large diameter shallow bore ground heat exchanger. *Applied Thermal Engineering*, 162, 114305.
- Naldi, C., et Zanchini, E. (2018). A new numerical method to determine isothermal g-functions of borehole heat exchanger fields. *Geothermics*, 77. 278-287.
- Newton, I. (1969). Mathematical Papers Vol. III, D.T. Whiteside (Ed.). *Cambridge Univ. Press, London*.
- Ni, L., Song, W., Zeng, F., et Yao, Y. (2011). Energy saving and economic analyses of design heating load ratio of ground source heat pump with gas boiler as auxiliary heat source. *2011 International Conference on Electric Technology and Civil Engineering (ICETCE)*.

- Nguyen, H.V., Law, Y.L.E., Alavy, M., Walsh, P.R., Leong, W.H., et Dworkin, S.B. (2014). An analysis of the factors affecting hybrid ground-source heat pump installation potential in North America. *Applied Energy*, 125, 28 - 38.
- Nguyen, A., Pasquier, P., et Marcotte, D. (2017). Borehole thermal energy storage systems under the influence of groundwater flow and time-varying surface temperature. *Geothermics*, 66, 110–118.
- O'Hagan, A. (2006). Bayesian analysis of computer code outputs : a tutorial. *Reliability Engineering & System Safety*, 91, 1290 - 1300.
- Omer, A.M. (2008). Ground-source heat pump systems and applications. *Renewable and Sustainable Energy Reviews*, 12(2), 344-371.
- Onishi, V.C., Ravagnani, M.A.S.S., et Caballero, J.A. (2013). Mathematical programming model for heat exchanger design through optimization of partial objectives. *Energy Conversion and Management*, 74, 60–69.
- Ozyurt, O., et Ekinici, D.A. (2011). Experimental study of vertical ground-source heat pump performance evaluation for cold climate in Turkey. *Applied Energy*, 88 (4), 1257 - 1265.
- Pandey, N., Murugesan, K., et Thomas, R.H. (2017). Optimization of ground heat exchangers for space heating and cooling applications using Taguchi method and utility concept. *Applied Energy*, 190, 421-438.
- Pasquier, P., et Marcotte, D. (2012). Short-term simulation of ground heat exchanger with an improved TRCM. *Renewable Energy*, 46, 92–99.
- Pasquier, P. et Marcotte, D. (2013). Efficient computation of heat flux signals to ensure the reproduction of prescribed temperatures at several interacting heat sources. *Applied Thermal Engineering*, 59(1-2), 515-526.
- Pasquier, P., et Marcotte, D. (2014). Joint use of quasi-3D response model and spectral method to simulate borehole heat exchanger. *Geothermics*, 51, 281-299.
- Pasquier, P., et Marcotte, D. (2020). Robust identification of volumetric heat capacity and analysis of thermal response tests by Bayesian inference with correlated residuals. *Applied Energy*, 261, 114394.
- Pasquier, P., Marcotte, D., Bernier, M., et Kummert, M. (2013). Simulation of ground-coupled heat pump systems using a spectral approach. *Proceedings of BS 2013. 13th Conference of International Building Performance Simulation Association*, Chambéry, France, 2691-2698.

- Pasquier, P., Zarrella, A., et Labib, R. (2018). Near-instant construction of short-term g-functions by artificial neural network. *Applied Thermal Engineering*, 143, 910–921.
- Pearl, J. (1984). Heuristics : intelligent search strategies for computer problem solving. United States : Addison-Wesley Pub. Co., Inc., Reading, MA., p. 3. OSTI 5127296.
- Penner, E. (1970). Thermal conductivity of frozen soils. *Canadian Journal of Earth Sciences*, 7(3), 982 - 987.
- Pianosi, F., Sarrazin, F., et Wagener, T. (2015). A Matlab toolbox for Global Sensitivity Analysis. *Environmental Modelling and Software*, 70, 80 - 85.
- Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., et Wagener, T. (2016). Sensitivity analysis of environmental models : A systematic review with practical workflow. *Environmental Modelling and Software*, 79, 214–232.
- Priarone, A., et Fossa, M. (2016). Temperature response factors at different boundary conditions for modelling the single borehole heat exchanger. *Applied Thermal Engineering*, 103(25), 934–944.
- Pu, L., Qi, D., Xu, L., et Li, Y. (2017). Optimization on the performance of ground heat exchangers for GSHP using Kriging model based on MOGA. *Applied Thermal Engineering*, 118(25), 480–489.
- Rao, R.V., et Saroj, A. (2018). Multi-objective design optimization of heat exchangers using elitist-Jaya algorithm. *Energy Systems*, 9(2), 305-341.
- Rees S.J., et He M. (2013). A three-dimensional numerical model of borehole heat exchanger heat transfer and fluid flow. *Geothermics*, 46, 1-13.
- Retkowski, W., et Thöming, J. (2014). Thermoeconomic optimization of vertical ground-source heat pump systems through nonlinear integer programming. *Applied Energy*, 114, 492-503.
- Rivera J.A., Blum, P., et Bayer, P. (2016). A finite line source model with Cauchy-type top boundary conditions for simulating near surface effects on borehole heat exchangers. *Energy*, 98, 50-63.
- Robert, F., et Gosselin, L. (2014). New methodology to design ground coupled heat pump systems based on total cost minimization. *Applied Thermal Engineering*, 62(2), 481-491.

- Rockafellar, R.T., et Uryasev, S. (1999). Optimization of conditional value-at-risk. *Journal of Risk*, 2, 21–41.
- Rosenblatt, F. (1958). The perceptron : A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–407.
- Rui, Y., Garber, D., et Yin, M. (2018). Modelling ground source heat pump system by an integrated simulation programme. *Applied Thermal Engineering*, 134, 450–459.
- Ruiz-Calvo, F., De Rosa, M., Acuña, J., Corberán, J.M., et Montagud, C. (2015). Experimental validation of a short-term Borehole-to-Ground (B2G) dynamic model. *Applied Energy*, 140, 210–223.
- Samson, M., Dallaire, J., et Gosselin, L. (2018). Influence of groundwater flow on cost minimization of ground coupled heat pump systems. *Geothermics*, 73, 100–110.
- Sanaye, S., et Hajabdollahi, H. (2010). Multi-objective optimization of shell and tube heat exchangers. *Applied Thermal Engineering*, 30(14–15), 1937–1945.
- Sayyadi, H., et Nejatollahi, M. (2011). Thermodynamic and thermoeconomic optimization of a cooling tower-assisted ground source heat pump. *Geothermics*, 40(3), 221–232.
- Sayyadi, H., Amlashi, E.H., et Amidpour, M. (2009). Multi-objective optimization of a vertical ground source heat pump using evolutionary algorithm. *Energy Conversion and Management*, 50(8), 2035–2046.
- Schelenz, S., Vienken, T., Shao, H., Firmbach, L., et Dietrich, P. (2017). On the importance of a coordinated site characterization for the sustainable intensive thermal use of the shallow subsurface in urban areas : a case study. *Environmental Earth Sciences*, 76 (2), 73.
- Schulte, D.O., Rühaak, W., Welsch, B., et Sass, I. (2016). BASIMO–Borehole Heat Exchanger Array Simulation and Optimization Tool. *Energy Procedia*, 97, 210–217.
- Self S.J., Reddy B.V., et Rosen M.A. (2013). Geothermal heat pump systems : Status review and comparison with 530 other heating options. *Applied Energy*, 101, 341–348.
- Sheriff, F. (2007). Génération de facteurs de réponse pour champs de puits géothermiques verticaux. Mémoire de maîtrise Montréal, Canada : École polytechnique.
- Shonder, J.A., Baxter, V.D., Thornton, J., et Hughes, P. (1999). A new comparison of vertical ground heat exchanger design methods for residential applications. *ASHRAE Transactions*, 105(2).

- Sivasakthivel, T., Murugesan, K., et Sahoo, P.K. (2014). Optimization of ground heat exchanger parameters of ground source heat pump system for space heating applications. *Energy*, 78, 573–586.
- Sobol, I.M. (1993). Sensitivity estimates for non linear mathematical models. *Math. Modell. Comput.*, 14, 407–414.
- Sun, W., Hu, P., Lei, F., Zhu, N., et Jiang, Z. (2015). Case study of performance evaluation of ground source heat pump system based on ANN and ANFIS models. *Applied Thermal Engineering*, 87, 586–594.
- Trefethen, L.N. (2007). Computing numerically with functions instead of numbers. *Mathematics in Computer Science*, 1, 9-19.
- Tsagarakis, K.P. (2019). Shallow geothermal energy under the microscope : Social, economic, and institutional aspects. *Renewable Energy*, In Press, Corrected Proof.
- TRNSYS - University of Wisconsin. (2000). TRNSYS, a transient system simulation program.
- Turgut, O.E. (2016). Hybrid Chaotic Quantum behaved Particle Swarm Optimization algorithm for thermal design of plate fin heat exchangers. *Applied Mathematical Modelling*, 40(1), 50–69.
- Tye-Gingras, M., et Gosselin, L. (2014). Generic ground response functions for ground exchangers in the presence of groundwater flow. *Renewable Energy*, 72, 354-366.
- Wen, J., Yang, H., Jian, G., Tong, X., Li, K., et Wang, S. (2016). Energy and cost optimization of shell and tube heat exchanger with helical baffles using Kriging metamodel based on MOGA. *International Journal of Heat and Mass Transfer*, 98, 29–39.
- Xia, L., Ma, Z., Kokogiannakis, G., Wang, Z. et Wang, S. (2018). A model-based design optimization strategy for ground source heat pump systems with integrated photovoltaic thermal collectors. *Applied Energy*, 214, 178–190.
- Yabanova, I., et Keeba, A. (2013). Development of ANN model for geothermal district heating system and a novel PID-based control strategy. *Applied Thermal Engineering*, 51(1–2), 908–916.
- Yamout, G.M. (2005), Applications of single party and multiple party decision making under risk and uncertainty to water resources allocation problems, Ph.D. diss., Civ. Eng. Dept., Univ. of Florida, Gainesville, Fla.

- Yang, H., Cui, P., et Fang, Z. (2010). Vertical-borehole ground-coupled heat pumps : A review of models and systems. *Applied Energy*, 87 (1), 16 - 27.
- Yavuzturk, C., et Spitler, J. D. (1999). A short time step response factor model for vertical ground loop heat exchangers. *ASHRAE Transactions*, 105(2), 475-485.
- Yavuzturk, C., Spitler, J.D., et Rees, S.J. (1999). A transient two-dimensional finite volume model for the simulation of vertical U-tube ground heat exchangers. *ASHRAE Transactions*, 105(2), 465-474.
- Ye, S, et Tiong, R.L.K. (2000) NPV-at-Risk Method in Infrastructure Project Investment Evaluation. *Journal of Construction Engineering and Management*, 126(3).
- Yousefi, M., Enayatifar, R., Darus, A.N., et Abdullah, A.H. (2013). Optimization of plate-fin heat exchangers by an improved harmony search algorithm. *Applied Thermal Engineering*, 50 (1), 877-885.
- Yousefi, M., Darus, A.N., Yousefi, M., et Hoorshyar, D. (2015). Multi-stage thermal-economical optimization of compact heat exchangers : A new evolutionary-based design approach for real-world problems. *Applied Thermal Engineering*, 83 : 71-80.
- Zanchini, E., et Lazzari, S. (2014). New g-functions for the hourly simulation of double U-tube borehole heat exchanger fields. *Energy*, 70, 444-455.
- Zeng, H.Y., Diao, N.R. et Fang, Z.H. (2002). A finite line-source model for boreholes in geothermal heat exchangers. *Heat Transfer-Asian Research*, 31(7), 558-567.
- Zeng, R., Li, H., Liu, L., Zhang, X., et Zhang, G. (2015). A novel method based on multi-population genetic algorithm for CCHP-GSHP coupling system optimization. *Energy Conversion and Management*, 105, 1138-1148.
- Zhang, C., Hu, S., Liu, Y., et Wang, Q. (2016). Optimal design of borehole heat exchangers based on hourly load simulation. *Energy*, 116(1), 1180-1190.
- Zurmuhl, D.P., Lukawski, M.Z., Aguirre, G.A., Schnaars, G.P., Beckers, K.F., Anderson, C.L., et Tester, J.W. (2018). Hybrid geothermal heat pumps for cooling telecommunications data centers Proceedings. *43rd Workshop on Geothermal Reservoir Engineering*, Stanford University, Stanford, California, February 12-14, 2018.