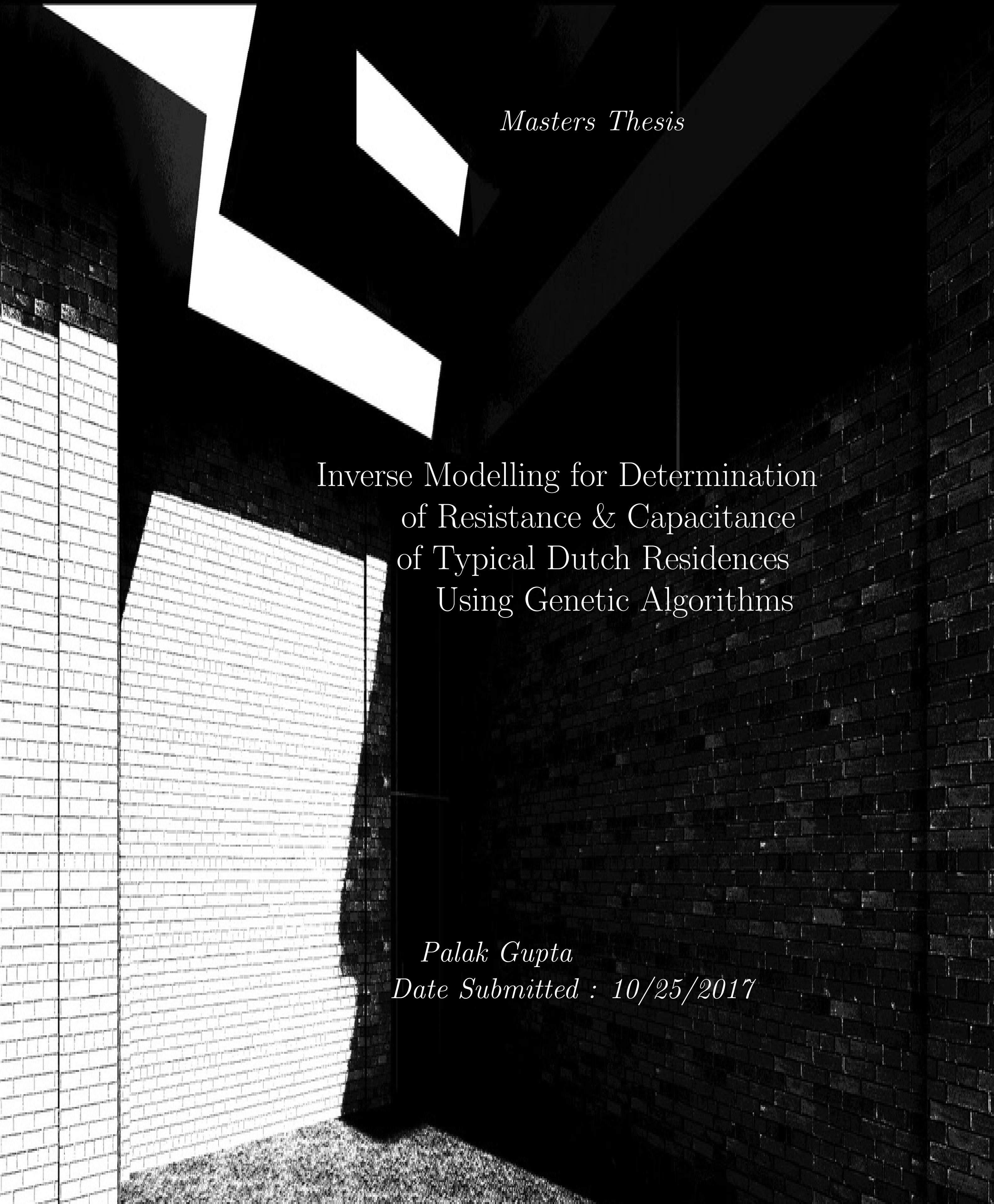
TU Delft

Masters Thesis



Inverse Modelling for Determination of Resistance & Capacitance of Typical Dutch Residences Using Genetic Algorithms

Palak Gupta

Date Submitted : 10/25/2017

Acknowledgements

I would like to thank my Thesis Supervisor Dr. Laure Itard for inspiring me to pursue my graduation thesis in Energy Modelling of Buildings and being a constant guide and mentor throughout. Her encouraging words and patience were indispensable in the last one year.

I would like to thank the staff and colleagues from the OTB for providing administrative help and ensuring I have a good working environment.

I would like to especially thank Ir. Vaidehi Parab, for providing me with a foundation for the thesis with her work on "Thermal Modelling of Existing Residential Buildings in North-Western Europe" and Mr. Pranav Chandramouli for enlightening discussions on the academic parts of the thesis.

I would like to thank my parents Mrs. Renu Gupta, Mr. R. P. Gupta and my brother Mr. Ankur Gupta for their confidence and support in me.

Summary

This thesis deals with determination of Resistance (R) value and Heat Capacity of building envelopes in the commonly found residential architecture in Holland i.e. pitched roof houses. The aim of this thesis is to find an easy to implement novel method which gives reasonably accurate results. In the absence of actual sensor data, the data-sets are obtained by modelling the buildings in EnergyPlus, with the help of Design Builder which provides the specifications of the buildings and its energy systems. Data-sets obtained via EnergyPlus provide the heat demand of the buildings. These data-sets, which are considered to be an approximation of actual data that would be available in future, is fed into MATLAB to perform the inverse modelling using Genetic Algorithms (GA) which estimates the unknown parameters by fitting the energy demand curve. Genetic Algorithms are known to converge to the global optimum unlike other regression techniques and parameter estimation methods. The objective function of GA is derived from the most accurate thermal network model. The equation is an energy balance of the room for the indoor temperature node which constitutes transmission losses, ventilation losses, solar gain and internal heat gain.The working and behaviour of Genetic Algorithms with varying optimisation parameters are thoroughly studied to make a comprehensive report on the scope of the Algorithm. Artificial Neural Networks and Hessian Matrix are studied in the course of thesis and are mentioned briefly in the appendix.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Built environment which comprises of any form of landscape created for human interaction, ranges from buildings to gardens. According to European Commission[1] humans spend 90% of their time indoors. Energy research for the built environment has a pivotal role to play in today's scenario.Energy Research for the Built Environment involves energy analyses of buildings populated by humans such as offices, classrooms, labs, common rooms and houses. Energy consumption in houses formed 25.4% of total energy consumption of Europe in 2015 according to EU commission[1]. The statistical figures for household energy consumption in Netherlands given by Santin *et al.* [2] is much higher and approximated to be around 42%. According to the study for building stock age in Netherlands by Laure Itard[3], the building stock percentage for post war construction accounts to 35%. As post war construction is poorly insulated, there is a high number of buildings consuming more energy than modern efficient houses and need to be retrofitted. It is essential to recognise the building envelope's thermal values to retrofit accordingly.

Building retrofitting can also introduce considerable amount of flexibility in energy control in existing building stock. Major proportion of building energy can be easily regulated using control mechanisms especially if designed taking into account specific building behaviour. For this reason, there is a constant need for research on understanding the behaviour of residence buildings. The work done in this project will focus on inverse modelling of the buildings and aim on extracting the important building parameters such that the thermal performance of the building can be understood and predicted. This study will be founded on studies of thermal networks, parameter estimation methods and pattern recognition. Hence, this work will be an amalgamation of heat transfer principles and statistical modelling with the aid of machine learning.

1.1 Research Objectives

The aim of the thesis is to study a generic reverse modelling method for existing buildings in the Netherlands. The study will be based on grey box modelling of the building discussed in Chapter 2 and determining the building parameters via machine learning algorithm which could provide us with information on sources of infiltration/ventilation. This study is builds on the thesis by Vaidehi Parab[4] and will take the work forward towards creating an algorithm for a typical dutch house with pitched roof. The intermediate objectives in this study are:

- Experiment with different thermal network models to identify the best suited RC model for different types of buildings in the building stock.
- Find the best technique for implementation of thermal networks using genetic algorithms as a parameter estimation technique to provide physically relevant values of R and Heat Capacity of the building while retaining ease of implementation.
- Draw up reverse engineering process for energy consumption profile created by DesignBuilder(DB) to test the technique before application on real time data collected from houses across Netherlands through smart meters for Opschaler project.

The report will start with discussing the basic concepts this study is based on and the associated literature in chapter 2. It will broadly comprise of mathematical modelling, thermal networks, heat dynamics of buildings and genetic algorithms(GA). This is followed by brief chapter defining the methodology employed in this work.

Chapter 2

Literature

2.1 Mathematical Modelling

Mathematical modelling refers to the depiction of a physical system in the form of mathematical equations to better study the behaviour of the system. Governing equations of the mathematical model define the system based on its parameters and their inter-dependency. It can be applied, in this case, for extracting information about the indoor dynamics of building systems. There are a number of ways to classify mathematical models, dynamic or static models, linear or non linear models, etc., which will be discussed in the next section. Another important classification is forward modelling or inverse modelling. The names give a clear indication of the model ideology: forward modelling is the conventional form where the model parameters are used to calculate the mathematical response and inverse modelling is when the response (building energy demand) is known which are correlated using mathematical models and processed to give the missing causal factors which produced the corresponding response. The knowledge gained would be in terms of R and C value of the fabric and in general building behaviour which could help identify anomalies in the measured output as compared to the predicted demand for the given building. In this study, the aim is to estimate the Resistance and the Capacitance of the building given the energy demand of space. Hence, for the purpose of this thesis inverse modelling is applied.

Inverse modelling has already been used for retrofit buildings in Zhang *et al.*[5] in order to make the indoor climate more suitable for the inhabitants, decrease the electricity bills, model control mechanisms, predict energy usage and increase energy efficiency of the building. The ideals of Zhang *et al.*. fall in the same general category as this study. The process for inverse modelling is shown in Figure 2.1.

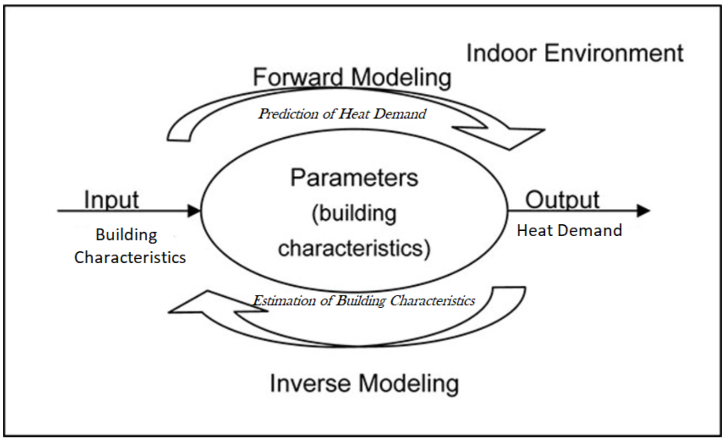


Figure 2.1: Algorithm for Inverse Modelling [6]

Another important factor to identify in data driven modelling is the selection of statistical process. ASHRAE classifies them as :

- Empirical or 'black box' approach: This method applies to cases where least information about the building is available. The information available is from the installed sensors and meters possibly about the temperature, occupancy, humidity and the loads but there is no insight into the thermal processes occurring inside. Once the approach is applied, the processes inside can be mapped out and analysed.

- Calibrated Simulation Approach: In this approach, the energy demand profile of the building is obtained using the energy software available like Trynsys, EnergyPlus, etc. This profile is plotted against the measured profile to calibrate the input variables such that the measured and predicted match as closely as they can. It is a necessary step as a lot of models require input data which is difficult to obtain and time consuming to estimate.

- Grey Box Approach: Grey Box modelling is a variation of black box model with the difference that there is some information available about the inner dynamics of the building. This approach is mostly used in buildings where the physical model is partially available in the form of sensor data or building fabric details etc.

2.2 Modelling Approach

Inverse modelling can take the form of steady state inverse models or dynamic inverse models. Steady state models are employed when the building at equilibrium is studied whereas the problem statement of the thesis involves analysis over time(hourly) hence dynamic inverse modelling will be carried out. Dynamic inverse modelling is one of the branches of inverse modelling which is based on differential equations of heat transfer mechanisms. Also, the aim of the project will be to work at the micro level which would include the transient changes within the space, which falls perfectly under the definition of dynamic inverse modelling. To further examine dynamic inverse models, we enumerate the different approaches available.

Zhang *et al.*.[5] compared four traditional identification methods namely change point model, Gaussian process model, Gaussian mixture model and artificial neural network. They found the Gaussian mixture to be the most accurate for limited data availability cases but change point model to be the optimum model as the lack of accuracy is compensated by the ease of application. Haberl *et al.*[7] proposed the hourly bin method which was able to capture non linear data changes. Wang and Xu[8] have shown good results with a simple thermal network model. Also Vaidehi Parab[4] has worked extensively on thermal networks and her work produced interesting results which could be studied further. Aim of this thesis is to take the work further by experimenting with thermal networks by trying different and more realistic thermal network models and studying the energy profile further to identify irregularities and their causes.

2.3 Parameter Estimation

Parameter estimation is the process of determining unknown parameters of a distribution with the help of pre-known distribution data. For a given equation, data becomes the known variables and the unknown variables are found by search and optimisation algorithms which converge to the most accurate values by minimising or maximising the objective function. Example of parameter estimation can be taken as a simple linear line equation y = mx+ c, where the left hand side(LHS) value of y is known for every right hand side(RHS) value of x. The unknown parameters of m and c are to be estimated and for that the objective function, $F = (mx + c) - y_{measured}$ needs to be minimised. For multiple data sets the objective function can be the sum of the error functions for each data point.

Parameter estimation in this thesis is used to determine the unknown coefficients of the mathematical model equation such that the modelled behaviour replicates the measured behaviour. Glenn[9] has discussed the most common parameter estimation techniques namely Root Mean Square, Maximum Likelihood Values using Markov Chain Monte-Carlo Algorithm and

Bayesian Approach. While root mean square involves a simple formula based on measuring the difference between the calculations and the observations, ML method gives the probability of model fitting to the data. Root Mean Square is a simple method and is used commonly in recent studies whereas likelihood method is slightly more complicated where the estimators can be achieved by various sub-methods like Markov Chain Monte Carlo(MCMC) or Bayes. It can also be implemented using Continuous Time Stochastic Modelling Using R(CTSM-R)[1] tool as covered in the thesis by Vaidehi Parab[4].

Genetic Algorithms have been used for parameter estimation by Wang *et al.*[8] and for curve fitting by Messa[10]. Genetic algorithms has been a subject of discussion for providing an edge over the other algorithms by showing an inherent ability to solve equations with large number of unknowns. Goldberg[13] states the strengths of GA which set it apart from other techniques like the Gradient Descent Methods which determine the parameter by a point-to-point transition. GA operates on parallel assessment of points in the database which leads to lesser likelihood of getting stuck at a local minima. Another feature of GA which aids in convergence to global optima is the mutation process. As GA is based on probabilistic approach, it randomly cruises through the search space to find the likely solution adding a random strength to the process of GA.Another reason for enhanced practicality of GA is its simplicity in execution compared to a gradient descent method which requires a differentiable objective function with respect to the parameter to be optimised. Its not always the case, especially for error functions, to have a first order derivative with respect to the parameter to be estimated providing a serious roadblock to application of these methods. Another major drawback of the gradient descent methodology is a tendency to get stuck in a local optima instead of identifying the global optima. Based on the above discussion, GAs might be helpful in solving under-determined mathematical systems. Under-determined mathematical systems are a set of equations which constitute more unknowns than the number of equations - this is the general case for thermal networks. Hence, Genetic Algorithms have been employed for parameter estimation in this thesis.

2.3.1 Introduction to Genetic Algorithms

Genetic Algorithms form a branch of evolutionary programming which is based on the principle of survival of the fittest. Genetic Algorithms were first

[1]CSTM-R is a software developed by Technical University of Denmark for estimating parameters in equations based on continuous time mathematical models

invented by John Holland in the 1960s. The reason that led to Holland's discovery of these algorithms was rooted in his scientific curiosity to better learn about natural adaptation in human evolution. Genetic algorithms later filled the space for computing algorithms which could abstract the concept of genetic evolution to other applications and better pave the way for computers to solve problems using the working knowledge of naturally occurring systems in the world.Since 1980s, scientists have been working on mimicking the working of natural processes occurring in the human brain (neural networks) as well as natural selection (evolutionary programming). The advancement of both algorithms introduced new and improved methods for problem solving. While neural networks are extensively used as black boxes for any problem solving, an associated drawback is their dependence on hit and trial method for adjusting the coefficients of input parameters depending on the change in output. Hence, for an ANN, change in output is the tracker and is thus a necessity which is not for the purpose of this thesis as it revolves around determination of resistances and capacitances which are constant for a building. Hence, rendering application of ANN for this case as null.

2.3.2 Natural Selection and Genetic Algorithms

Lets look into Charles Darwin's Theory of Evolution to better lay a foundation for the understanding of GA. Charles Darwin's Theory of Evolution states that individuals have a tendency through gene alterations to increase their fitness with every subsequent generation increasing their ability to reproduce, compete and survive. This selection process ensures that the subsequent generation will have higher number of fit individuals than the previous generation. Starting at the beginning of the human race, the first individuals would constitute the initial populationAssuming at the start there were two people, a male 'X' and a female 'Y'. The two people would constitute the zeroth generation. The children of X and Y will all be constituted from the genetic material of parents.Reproduction in humans is based on interaction of male and female gametes which respectively contain half the number of chromosomes needed for an offspring. After replication of gametes, they are crossed and recombined to form four unique half chromosome haploids. These haploids combine in pairs to generate the genetic makeup of the off-springs. This process is shown in the following Figure 2.2,

All the children of 'X' and 'Y' combined form the first generation. Theory of the fittest play a role in succeeding generation where in the words of Charles Darwin, the genetic makeup that provided the strongest offspring in the first generation will leave more copies of itself in the next generation than the

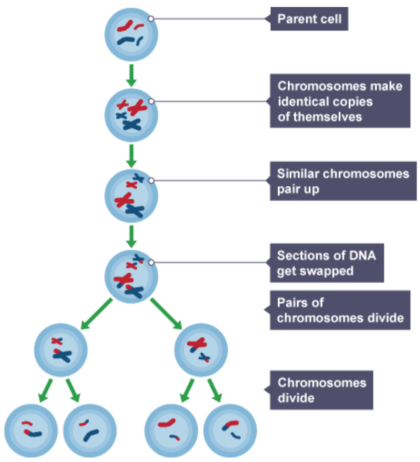


Figure 2.2: Process of Meiosis - Biological Reproduction of Crossover and Mutation

weaker offspring.

To understand how GA is modelled, let's replace the generation of individuals with generation of parameter values. Each generation is an iteration. Each generation has a certain number of people, GA has a certain number of guesses for all the parameters. Hence, GA is nothing but iterations of sets of possible solutions until the fittest solution is found. GA goes from one generation to the next by mutating and crossing the parameters from the previous generation. For this purpose, GA employs the Schema Theorem. A 'schema' is a string of bits, (zeros and ones) which can contain a zero, one or an asterisk as shown in figure 2.3. The asterisk in the schema is an unknown value which lays the way for chromosome sets. The sets of chromosomes are better known as instances of a schema or a 'schemata'. As from conventional bit system, the physical value can be determined by decoding it with a base of 2. Schemata with their decoding is shown in figure 2.4.

0	1	*	1

Figure 2.3: Schemata Representation

GA works by experimenting with schemata to achieve the optimal solution. In natural selection, the chromosomes which make future genes are mutated and crossed to make the subsequent fitter generation. This process continues until the strongest offspring is created. Similarly, when GA is applied to parameter estimation, the parameters act like the chromosomes which are varied within the constraints to obtain the optimum result. The algorithm runs through different parameter values whilst dropping the instances less likely to contribute to the needed output.

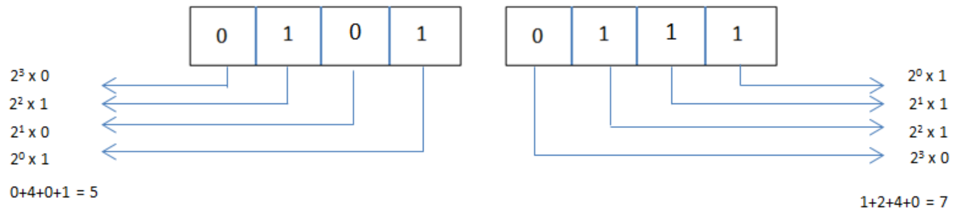
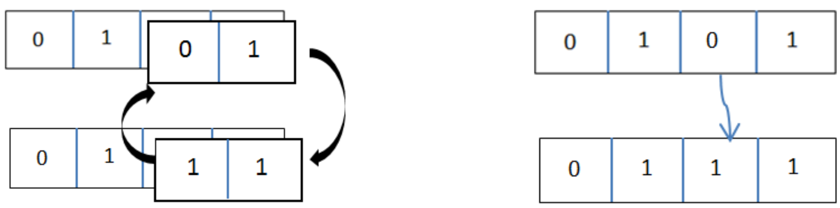


Figure 2.4: Schemata Decoding

The Schema theorem proposed by John Holland states that fitness of the schemata through the generations can be assessed by the number of instances present in the particular generation. Also, the probability of the schemata being present in the next generation can be calculated based on their fitness ratio as GA selection is based on a general principle of allotting more children to the fitter schema. An easy example of that will be roulette selection is which the most fit individual is correspondingly allotted a higher proportion on the roulette wheel which means that there will be more children of the most fit individual in the next generation than the relatively less fit parents.



(a) Crossover Representation (b) Mutation Representation

Figure 2.5: Mutation and Crossover in GA

Initialisation of GA requires 'seeds' which make the first generation for the algorithm. These seeds form the initial population coded into schemata. Once coded, the value of the objective function of each entry of the population is evaluated to find the fittest solution. After dropping the weaker population, the healthier individual genes are mutated and crossed to generate the next generation. Process of mutation and crossover is shown in figure 2.5. The process continues till the next generation has the same number of individuals as the initial generation. Process of mutation can be described as introducing random alterations in the schema. This population is then mated in a way that if a certain point is chosen for mating then a replication of that point is first created in the mating pool for the next mating out of the original pool. Mutation gives GA the unique quality of obtaining global minima as mutation by definition does not let the population saturate when its close to the local minima. As mutation introduces random changes in variables, when GA reaches a local minima, the algorithm does not consider only the data points near it but also random points which may be far away from the current optima hence possibly giving a more probable global optima. Crossover can be described as swapping the bits of the 2 parents at the point of 'break' to produce an offspring.

Genetic Algorithm can be depicted by the algorithm shown in figure 2.6,

2.3.3 GA terminology

- Fitness function handle : It is the name of the function file as well as the function name in the file. This is the objective function that is calculated for every individual and the aim of the genetic algorithm is to decrease the objective function hence increasing the fitness of subsequent generations.

- Constraints : Optimisation problems can be constrained with the help of linear relations or bounds.

- Population Type : Type can be either bit string (in the form of ones and zeros) or double vector. A bit is a value of 1 or 0 whereas a combination of bits can produce more values e.g. 00 = 0, 01 = 1, 10 = 2, 11 = 3 and so on. However, a limiting quality of bit-string is its inability to define decimal numbers. Double vector is a form of mixed integer programming where the variables can be decimals.

- Population Size : This can be specified by the programmer. Population size signifies the number of chromosomes created for each generation. Population size is an important criteria for achieving the desired result. There is a general understanding that a bigger population will have higher odds of achieving the precise solution whilst taking longer time to converge as compared to a small population GA. Population size also needs to be increased for a higher mutation rate for optimal results.

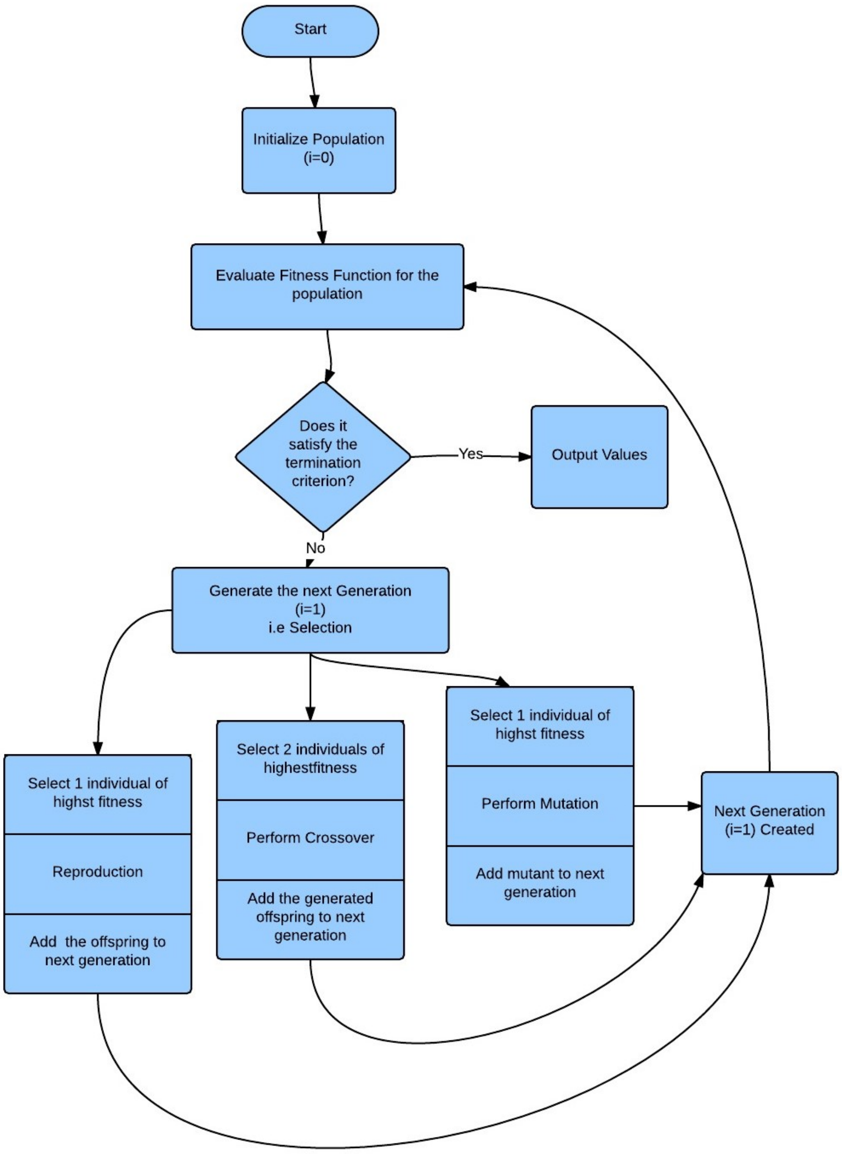


Figure 2.6: Genetic Algorithm Layout

- Creation Function : Creation function determines how the population is created. It can be 'Uniform' which creates a uniformly distributed populations in between the bounds. It can be 'feasible population' which creates population in bounds and pays attention only to linear constraints.

- Scaling Function : Scaling function is used to scale the individual scores for easier analysis of fitness values. It can either be 'rank' which ranks individuals based on their scores. Least score gets rank one, the next smallest score will get rank two and so on. There is 'top scale' which gives the rank 1 to top 40% of the fit individuals and 0 to the rest. Top linear and proportional both make the scaled value proportional to the fitness score.

- Selection Function : Selection function determines how the parents of the present generation are chosen to create the next generation. Principally, all selection methods follow the rule of creating more off-springs from the fitter individuals. It can be in the form of roulette where the schema with higher fitness covers a larger part of the roulette wheel hence creating more offsprings from the fitter individuals. Stochastic Uniform uses a line gauge instead of a roulette wheel.

- Reproduction : With the reproduction parameter, the user can specify the elite count which defines the percentage of individuals which are guaranteed to last through the next generation i.e. some of the fittest children that can be produced by the current generation are directly incorporated in the next generation. Also, we can determine the ratio of children produced by crossover and mutation.

- Mutation Function : Mutation function determines how the individuals are mutated. We can choose between Gaussian which alters all the parameters by a Gaussian distribution value centred around zero, uniform reduces the value of the chosen gene by a uniform value which is in range of the bounds, and adaptive feasible which chooses the value and direction of mutation variation depending on the last successful generation with in the bounds specified. Gaussian does not take into account the bounds for the parameters and hence has increased probability of resulting in non feasible solutions. Mutation fraction is mostly set to 0.01 to make sure that the optimisation problem does not suffer from the harmful effects of mutation in which the algorithm might be distracted from its global minima. However, for some cases, a low fraction does not provide satisfactory results. Hence in such a case, the mutation value can be increased, keeping in mind that the population needs to be increased as the mutation fraction increases. The variation of mutation rates is also a test for the consistency of the results.

- Crossover Function :Crossover function can be chosen from a variety of options - It can be a ratio which depends on weighted average of the parents; it can be scattered where a crossover vector chooses which genes are chosen from which parent; single point where the given value determines which fraction of genes from the second parent are incorporated; two point uses two splits in the genes mix of the parents; arithmetic creates children which lie at the arithmetic mean of the parents; heuristic creates children which lie a small distance away from the healthier parent.

- Stall Generations : Stall generations are the number of generations GA will run once the fitness function values change is less than the set function tolerance value. For more parameters, stall generations need to be increased to make sure it is not a local minima.

- Objective function : Objective function is the minimisation equation which seeks to reduce the difference between the measured and calculated values of the parameters to be optimised/estimated. To determine building parameters, genetic algorithm is applied with an objective function representing the error between the measured indoor temperature and the estimated indoor temperature.It is to be noted that for a genetic algorithm implementation, objective function and fitness function are synonymous.

- Average Distance Between Individuals : The parameter of distance in reference to GA pertains to the variation in input datasets.For multiple parameters, this distance refers to the summation of distances for each parameter.For example the distance between individuals for a constrained problem with bounds[0,0] to [10,10] will be max 20.To ensure an accurate solution, the distance at the time of convergence should be small signifying narrowing to the optimal solution. Through the process of finding the minima, the distance should not be too large or too small to ensure thorough screening of the input space.

- Fitness Value : Fitness Value is the value of objective function for each input parameter set.They are also known as scores.Fitness Values are displayed in terms of best fitness value and mean fitness value at the time of convergence.A correct measure of a successful result would be when the mean and best fitness value are close to each other signifying the converged solution as the best possible solution for the problem.This is because the closer the mean value is to best fitness value, the more replicated the population of GA is at the time of convergence hence there is no expected improvement in results.

2.3.4 Application of Genetic Algorithms

To introduce the mechanics of GA, consider the simple example of a linear line equation, $y = mx + c$. GA can be applied to estimate the two constants m and c given 10 data points for x and y. Objective function used for this example can be formulated as,

$$A = mx + c - y \tag{2.1}$$

A data-set of (x, y) corresponding to m=4 and c=9.6 was used as input for GA to identify m and c with the initial bounds [1,1] and final bounds [10,10]. The results are shown in Figure 2.7.

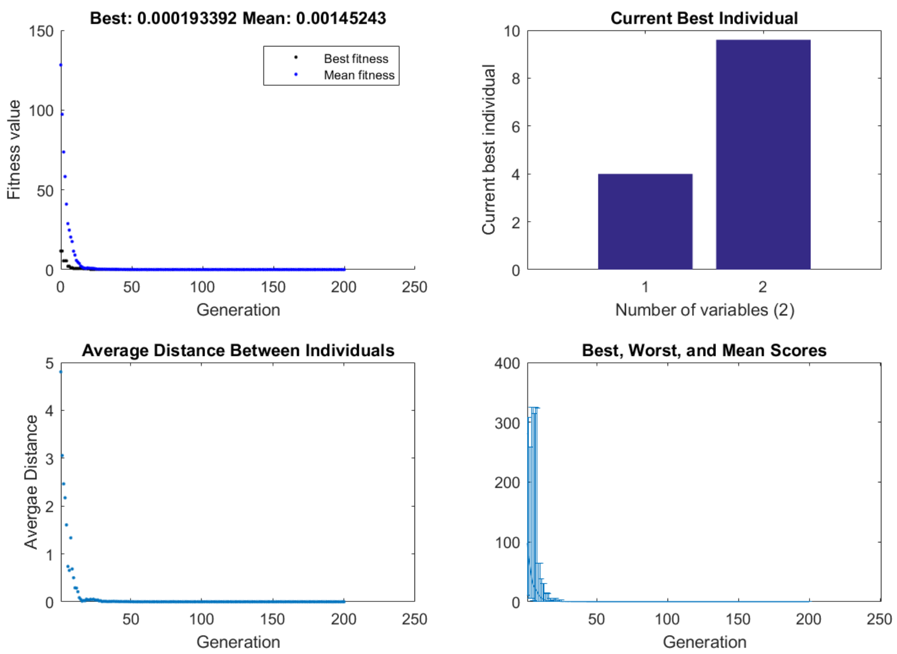


Figure 2.7: GA computed parameters for equation y=mx+c

The above figure shows that GA successfully predicted the value of m and c in 200 generations. As seen from the first plot, best fitness value is negligible demonstrating an accurate result. At the same time the mean distance is low as well which shows that the convergence is not premature. Second Plot is a bar chart of the best individual according to GA. Average Distance Between Individuals is shown in the third plot above.It is seen that the average distance decreased rapidly in the beginning and gradually later as it became harder to find a better solution.It can be seen that after generation

25, average distance is almost negligible hence GA has found roughly the optimal solution area and is now finding the exact solution.Here the stall generations are set to 50 hence GA will stop only in the case when fitness value improves lesser than 10^-6 for 50 generations. The fourth plot depicts the worst score/fitness value of the generation in addition to the first plot information. It is important to note here that the converged generation should not have the worst score far from the best score.

2.4 Heat Transfer Fundamentals

Having established the principles of GA, it is necessary now to formulate the accurate objective function for GA to utilise. This requires an understanding of the thermal behaviour of a building which can be achieved by the study of heat transfer principles which are enumerated in this section. The main energy processes taking place in a generic building are; solar influx, ventilation losses, transmission losses and internal gains. Energy balance is created based on the interplay between these four processes as shown in Figure 2.8.

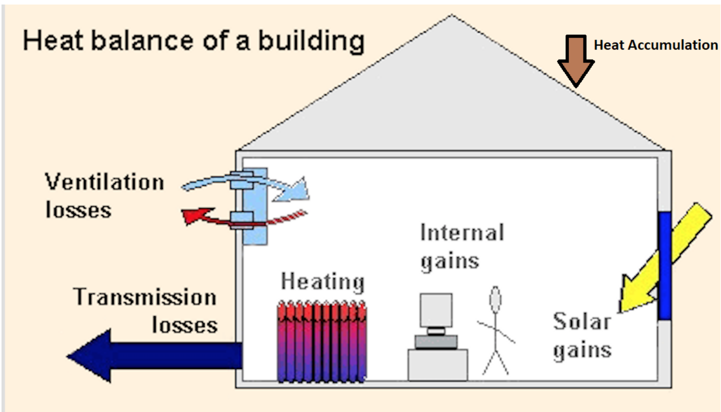


Figure 2.8: Heat Balance of a Building[11]

Solar heat gain is the measure of incoming heat from solar radiation through the building fabric and windows. This phenomena is responsible for the highest heat input during sunny weather and also leads to stored heat in the walls resulting in the peculiar dynamics of thermal response which play a significant role in the overall behaviour of the building.

Transmission losses refer to the heat loss from the indoor environments to the outer environment due to temperature differences. The mode of heat transfer is through conduction via wall and convection (high in case of fast winds). They become significant in case of extreme temperatures differences.

Ventilation losses occur via convective heat transfer. As there is constant inflow and outflow of fresh air into the building, there is heat loss between the inner space and environment based on the number of air changes per hour between the two - the ventilation system installed plays a distinct role in defining the number of air changes.

Internal gains are the heat gains in the indoor environments originating from the occupants, equipment, lighting, etc., which radiate heat due to internal processes. These gains are significant contributors to the energy balance during high occupancy or high equipment usage.

For a detailed heat transfer balance, heater output and heat stored in furniture are included in the dynamics of internal gains. The following processes determine the overall dynamics of a building (see figure 2.9),

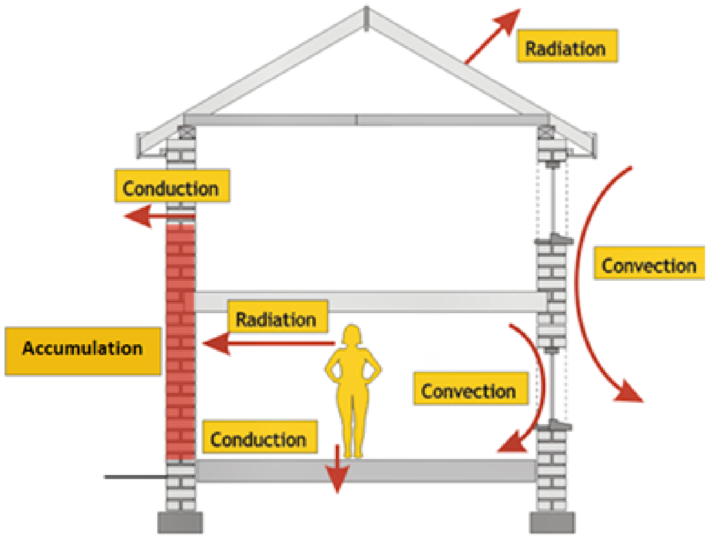


Figure 2.9: Detailed Thermal Processes In the Room[12]

- Conduction through the outer envelope (External Walls, Roof, External Floor, Windows). This heat transfer will be high for extreme temperature difference between indoors and outdoors. Also, for an un-insulated building, the conduction losses will strongly affect the inner dynamics.

- Convection from ambient air to the outer envelope. In the presence of high winds, convection effects will be higher.

- Convection from the internal envelope to indoor air. In case of high mechanical ventilation/infiltration, convection from the inner walls will be high.
- Convection from the furniture to indoor air.
- Convection from the heater to indoor air. The convective effect from the heaters will be considerable for a convective heating system. For a fully convective system, there is no radiation fraction.
- Radiation from the occupants, appliances, lights. There is a radiative and convective effect of the heat sources which is decided by the temperatures of the inner surfaces. In case of a heavy built fabric, the temperature of the wall will be high which will reduce the radiative effect of the heat sources thereby increasing the convective effect.
- Radiation influx from the sun. The transmitted solar radiation is absorbed by the inner surfaces of the building and absorbed radiated fully into the space gradually. The radiation falling on the envelope outer surface might play an additional role for walls with no insulation. The solar radiation falling on the outer surface will get stored in the wall possibly increasing the temperature of the inner surfaces hence decreasing the radiative effect of the transmitted solar radiation and the inner heat sources. This will in turn increase the convective fraction for the solar gains
- Radiation between the wall, heater and furniture
- Heat transfer through infiltration. This heat transfer is a direct measure of the temperature difference between indoors and outdoor.

The interplay of all these processes determines indoor climate.

2.5 Thermal Networks

Thermal network is a term coined for an electrical network representation of the various thermal processes described in the previous section. Heat transfer is based on a fundamental concept that heat flows due to a temperature difference between thermal boundaries. Heat transfer processes are dictated by heat transfer properties and geometric framing of the system elements like heat transfer coefficients, cross sectional areas and specific heat capacities. Heat transfer processes when mapped to an electrical network will present heat transfer coefficients as resistance and heat capacities as capacitance. Heat sources are formulated as current sources while the temperature difference which drives the entire process as voltage which drives current in an electrical circuit.

To explain the element conversions, the principles of a thermal process and an electrical process are compared. Temperature difference is the acting force for thermal exchange, which is equivalent to a voltage source which is the driving force for current. Every node in a thermal network is a component of the thermal system with a corresponding temperature just like every node in an electrical circuit has its corresponding voltage. Hence, for thermal network modelling the thermal behaviour of building, ambient, envelope, indoor air, and furniture each become nodes of the thermal network. Heat capacity of a component is given in terms of capacitance as they are both a measure of energy storage in their respective networks. Thermal heat capacity of envelope or furniture are parallel processes of heat transfer and hence are incorporated in parallel with in the thermal network. Sources of heat flux in a building are depicted as currents in an electrical circuit as heat flux is a source of heat into the temperature node. Resistance of the walls is the property which determines the amount of heat allowed indoors through a given surface equivalent to resistance in an electrical circuit. So, a simple thermal network can be formulated for heat transfer through the external wall as shown in figure 2.10

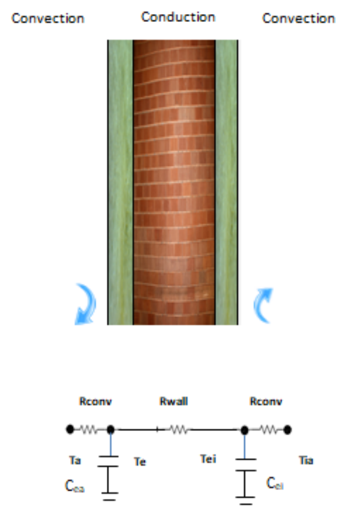


Figure 2.10: Thermal Network-Electric Model Analogy

Nomenclature of a thermal network is given in the form of either xRyC

(Harb [14]) where x denotes the number of resistances in the thermal network model and y corresponds to the number of capacitances or TiTjAi (Bacher [15]) where i, j denote the nodes in the model while Ai represents the areas of the node.

Harb[14] drew up the most common thermal models of the xRyC type, namely 1R1C, 3R2C, 4R2C and 8R3C with increasing model complexity - higher R and C leads to more parameters and hence a more detailed draw up of the building behaviour. The various models of Harb are shown in figure 2.11.

Subscript	Component
In	Interior(Including Walls,Floors,Furniture)
E	Exterior(External Surface of Envelope)
Ia	Indoor Air
A	Ambience

Table 2.1: Harb Model Parameter Nomenclature

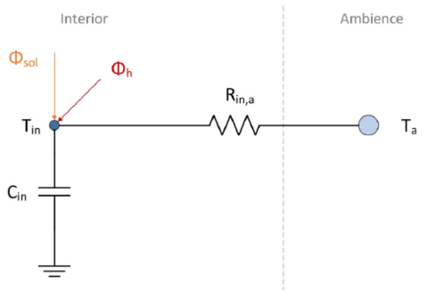
Similarly, Bacher formulated models of the TiTjAi type ranging from Ti model to TiTmTeThTsAe model. These models are shown in Figure 2.13.

Subscript	Component
M	Medium(Including Internal Walls,Floors,Furniture)
S	Sensor
H	Heater
E	Envelope(External Surface)
I	Interior(Indoor Air)
A	Ambient

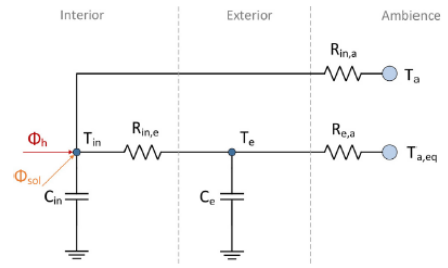
Table 2.2: Bacher Model Parameter Nomenclature

1R1C model is the basic model of Harb. In this model,T_{in} refers to the internal thermal mass node. The capacity of all the building components is represented by one aggregated capacitance. The equivalent resistance also takes into account all the resistances inside the building. From an understanding of thermal processes, it can be concluded that the equivalent capacitance would not be a linear summation of all capacitance but an indirect combination of them.

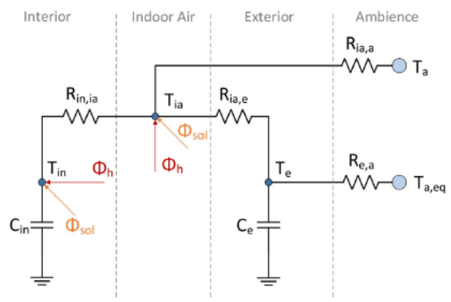
Equivalent resistance will represent the resistance of the wall in series with the convective resistances of the surfaces whilst internal surface resistance of furniture will be parallel combined with the convective resistance of the



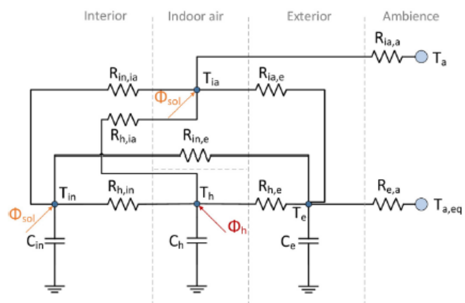
(a) 1R1C Model by Harb



(b) 3R2C model discussed by Harb



(c) 4R2C model discussed by Harb



(d) 8R2C model discussed by Harb

Figure 2.11: Thermal Network Models By Harb

internal thermal mass. The series resistance value will be in parallel with the infiltration resistance. The schematic is shown in Figure 2.12

A point of note is that parameters such as resistances and capacitances in thermal networks might represent different things for each network as they develop. Solar heat flux is the heat absorbed by the interior components.

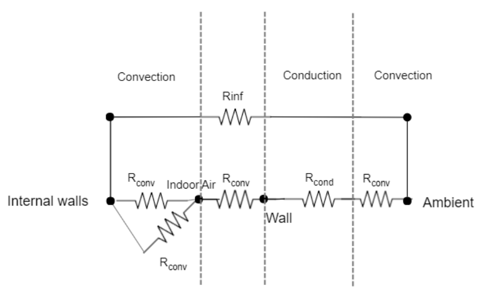


Figure 2.12: Schematic of Resistances

3R2C model splits the internal thermal mass node represented by T_{in} in the previous network to two nodes representing the temperature of external surface of envelope,$T_e$and internal thermal mass temperature,T_{in}. This model divides the equivalent capacitance into envelope capacitance and capacitance of internal thermal mass which is effectively the capacitance of the internal surface namely walls, floors and furniture. In this model, resistance of infiltration is presented separately and resistances between indoor air and ambient are split in to resistances of the inner, envelope and outer surfaces of the envelope. $R_{in,e}$ is the radiative exchange between the interior components.The effect of solar radiation on the outer surface of envelope are not taken into account.

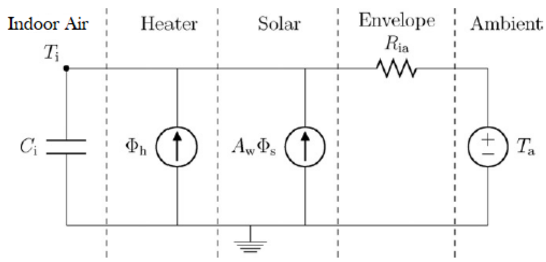
4R2C model shown in Figure 2.11c splits the T_{in} node further to include a separate node for indoor air, T_{ia}. This leads to separation of resistances $R_{in,e}$ to $R_{ia,e}$ and $R_{in,ia}$ which represent the convection processes from the internal surfaces to the indoor air which further convects to/from the wall. In this model, the heat flux from the sun and the heater are applied to the indoor air and the internal surfaces but not to the external envelope.Indoor air receives the convective fractions of the heat flux and the radiative fractions are considered into the envelope node. The radiative effect between the envelope and the interior components is neglected while incorporating the convective parts.

8R3C is the most complex model presented by Harb and is shown in Figure 2.11d. It includes heater as an extra node further complicating the model and introduces the thermal processes occurring between the heater node and previously present nodes. Convection and radiation resistances from the heater are added to the indoor air and all surfaces (external and internal) respectively. Solar flux is added to the indoor air and internal node while heat flux is added to the heater node. Capacitance of the heater is also included in this model while radiative resistances are added between the external envelope and the internal surfaces.

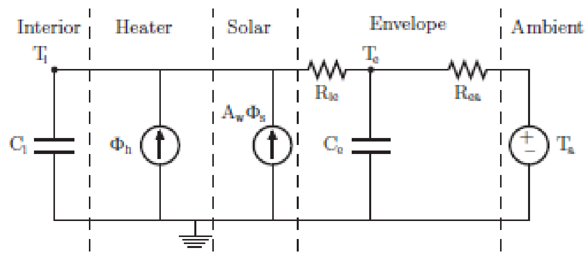
Ti model by Bacher shown in Figure 2.13a which is his most basic model with two nodes: T_i representing temperature of the indoor air and T_a the ambient temperature. This model utilises a combined resistance and capacitance similar to the 1R1C model of Harb. The solar flux is represented as a multiplication of effective window area with the measured global irradiance.

TiTe model is shown in Figure 2.13b. It splits the nodes into indoor air and envelope node. The convective resistance between the indoor air and ambient air is shown as a series combination of R_{ie} and R_{ea}. Capacitance of the envelope is also included with C_i representing the capacitance of the internal thermal mass.

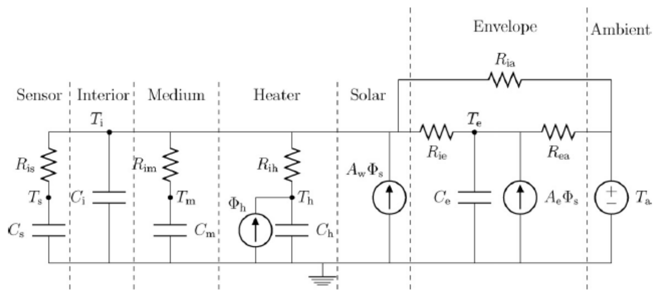
The models shown in Figure 2.13c and 2.13d split the nodes further to include a node for furniture and sensor. With the addition of these



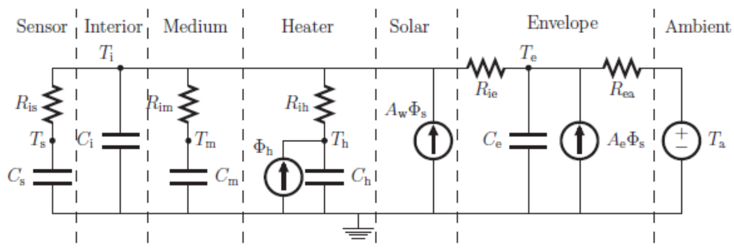
(a) Ti Model by Bacher



(b) TiTe Model by Bacher



(c) TiTmTeThTsAeRia Model by Bacher



(d) TiTmTeThTsAe Model by Bacher

Figure 2.13: Thermal Network Models By Bacher

nodes, corresponding convective resistances and capacitances are added. The difference between the above two models arises from the presence of R_{ia}, the infiltration resistance, which is ignored in the TiTmTeThTsAe model. TiTmTeThTsAe model shown in Figure 2.13d is concluded by Bacher as the best suited model.

The above circuits can be reduced to a set of differential equations which describe the thermal process occurring at each node. The equations pertaining to the models will be discussed in latter topics.

Thermal Networks have been gathering attention lately in the field of energy modelling of buildings. One extensive work on thermal networks was by Bacher[15] whose study identifies thermal models in the form of nodes unlike the former nomenclature elucidated in this section. These models are variations to the above models ranging from the simplest Ti model to the most complex model $T_iT_mT_eT_hT_sA_eR_{ia}$. The difference between Bacher's and Harb's most complex models is that in 8R3C, the radiation effect of the heater and the interior furniture is incorporated whereas in $T_iT_mT_eT_hT_sA_eR_{ia}$ model, only the convective heat transfer is taken into account. Also in Bacher's model, heat capacities of indoor air are taken into account whereas it is ignored in Harb's work. It should be noted that the inclusion of indoor air heat capacity might not be important as it is very low compared to the heat capacity of the envelope. Also, indoor air flow in and out of the building is constant and high for purposes of mechanical ventilation. One more significant difference is that in Harb's model, the sensor node is not accounted for unlike Bacher's.

Bacher identifies $T_iT_mT_eT_hT_sA_e$ to be the most suitable model as the house Bacher applied the model to was well insulated hence infiltration did not play a big role. However, infiltration can affect the performance of a typical building built in post war era hence for our study inclusion of infiltration is necessary. Harb concludes 4R2C to be the most suitable model. The common denominator between the two studies is their conclusion that the most complex model is not suitable due to introduction of errors to compensate for determining more parameters. However, Harb's study is based on buildings with insulation on the external surface of the building hence the effect of solar radiation on the external envelope was not included. However, for houses with less insulation, this factor cannot be ignored.

It can be concluded from the above discussion on thermal networks that various versions are possible to model a building. It is also possible to improve the models by using the knowledge on energy building simulation. It is known from literature that heat capacity of indoor air can be neglected as it does not play a relatively big role compared to the other capacities in the network. Also, the effect of solar radiation on the external envelope surface can be neglected for an insulated building as the transmitted solar gains are the main contributors to the indoor climate. In addition, the temperature of envelope cannot be considered uniform as it varies to a high extent in the presence of large temperature differences. Heat Capacity of the heater can be neglected unless the mode of heating is via floor heating.

2.6 Formulation of Differential Models

To derive thermal networks, mathematical equations governing the thermal processes occurring in buildings are compared to equations governing an electrical network.

Thermal Process	Electrical Process
Rate of Transmission Losses : Q = U_{wall}AdT	$I = V/R$
Rate of Ventilation Losses : Q = m$C_p dT$	Q=CV where $I = dQ/dt$
Heat Flux : Q	Q=CV where $I = dQ/dt$
Rate of Convective Losses :$Q = UAdT$	$I = V/R$

Table 2.3: Analogy of Thermal and Electrical Processes

It is clear from table 2.3, R (resistance of the thermal network) is the inverse of heat transfer coefficient multiplied by the area of the wall ($R = 1/UA$) and C (Capacitance of a thermal circuit) is the multiplication of mass of the material with its specific heat capacity($C = mC_p$).

As thermal networks are drawn in principle parallel to electrical networks, their equations are derived based on Kirchhof's Current Law which states that the sum of currents on each node is zero. Hence for a thermal node, all the processes occurring with respect to the node can be equated to zero hence obeying the law of energy conservation.

Lets take an example of the internal envelope node discussed in latter chapters. The heat transfer processes for this node include the convection to the ambient node combined with the conduction through the envelope and the convection to the indoor air node. It also includes the solar gain and heat flux absorbed by the wall.

$$\begin{aligned} 0 &= Q_{tran} - Q_{flux} \\ &= UAdT - Q_{flux} \\ &= \frac{(T_i - T_e)}{R_{ie}} + \frac{(T_a - T_e)}{R_{ae}} - CdT_e \end{aligned} \tag{2.2}$$

$$\frac{\partial T_e}{\partial t} = \frac{(T_i - T_e)}{R_{ie}C_e} + \frac{(T_a - T_e)}{R_{ae}C_e}, \tag{2.3}$$

Every element in the circuit above has a sub circuit containing a capacitance, resistance and a driving temperature difference. For every node, the above equation will be modified to correspond to the node and in addition, heat flux if available will also be considered.

$$dT_i = \frac{1}{R_{ie}C_i}\left(T_e - T_i\right)dt + \frac{1}{C_i}\phi_h dt + \frac{1}{C_i}\phi_s A_w dt \tag{2.4}$$

Similarly, the above equation is an example of the equation for the temperature of internal components inclusive of indoor air and internal surfaces. Hence, it consists of solar radiation, heater radiation, and envelope heat transfer which form the major components of heat transfer.

For the purpose of this study, the input will be the energy profile and the aim will be to compute the building envelope characteristics. For these results, the parameters needed will be T_a, T_i, T_e,Heater Demand and Solar Radiation.

2.7 List of Symbols for Harb Models

For 1R1C Model,
$R_{in,a}$ - Equivalent Resistance between Internal Surfaces and Ambient
C_{in} - Equivalent Capacitance of the Building
T_{in} -Equivalent Temperature of Internal Surfaces(Internal Walls, Floors, Furniture)
T_a -Temperature of Ambient

For 3R2C Model,
$R_{in,a}$ - Infiltration /Ventilation Resistance
$R_{e,a}$ - Resistance between Envelope and Ambient
$R_{in,e}$ - Radiative Resistance between Internal Surfaces and Envelope
C_e - Capacitance of External Envelope(External Walls, Floors, Roof
T_e - Uniform Temperature of Envelope(External Walls, Floors, Roof
C_{in} - Capacitance of Internal Surfaces(Internal Walls, Floors, Furniture
T_{in} - Temperature of Internal Surfaces(Internal Walls, Floors, Furniture

For 4R2C Model,
$R_{ia,a}$ - Infiltration/Ventilation Resistance
$R_{ia,e}$ - Convective Resistance between Envelope and Indoor air
$R_{e,a}$ - Convective Resistance between External Surfaces and Ambient
$R_{in,ia}$ -Convective Resistance between Internal Surfaces and Indoor Air
C_e - Capacitance of Envelope(External Walls, Floors, Roof

T_e - Uniform Temperature of Envelope(External Walls, Floors, Roof
T_{ia} - Temperature of Indoor Air
C_{in} - Capacitance of Internal Surfaces(Internal Walls, Floors, Furniture
T_{in} - Temperature of Internal Surfaces(Internal Walls, Floors, Furniture

For 8R2C Model,
$R_{ia,a}$ - Infiltration/Ventilation Resistance
$R_{ia,e}$ - Convective Resistance between Envelope and Indoor air
$R_{e,a}$ - Convective Resistance between Envelope and Ambient
$R_{in,ia}$ -Convective Resistance between Internal Surfaces and Indoor Air
$R_{in,e}$ - Radiative Resistance between Internal Surfaces and Envelope
$R_{h,in}$ -Radiative Resistance between Heater and Internal Surfaces
$R_{h,e}$ -Radiative Resistance between Heater and External Envelope
$R_{h,ia}$ -Convective Resistance between Heater and Indoor Air
C_e - Capacitance of Envelope(External Walls, Floors, Roof
T_e - Uniform Temperature of Envelope(External Walls, Floors, Roof
T_{ia} - Temperature of Indoor Air
C_{in} - Capacitance of Internal Surfaces(Internal Walls, Floors, Furniture
T_{in} - Temperature of Internal Surfaces(Internal Walls, Floors, Furniture
C_h - Capacitance of Heater
T_h - Temperature of Heater Surface

2.8 List of Symbols for Bacher Models

For Ti Model,
R_{ia} - Equivalent Resistance between Indoor Air and Ambient
C_i - Equivalent Capacitance(Internal Walls, Floors, Furniture,Envelope)
T_i -Temperature of Indoor Air
T_a -Temperature of Ambient
A_w -Area of External Windows
ϕs -Global Solar Radiation
ϕh -Heater Radiation

For TiTe Model,
R_{ie} - Convective Resistance between Indoor Air and Envelope
R_{ea} - Convective Resistance between Envelope and Ambient
C_i - Capacitance of Internal Surfaces(Internal Walls, Floors, Furniture)
C_e - Capacitance of Envelope
T_i -Temperature of Indoor Air

T_a -Temperature of Ambient
T_e -Temperature of Envelope
A_w -Area of External Windows
ϕs -Global Solar Radiation
ϕh -Heater Radiation

For TiTmTeThTsAe Model,
R_{ie} - Convective Resistance between Indoor Air and Envelope
R_{ea} - Convective Resistance between Envelope and Ambient
R_{ih} - Convective Resistance between Indoor Air and Heater
R_{im} - Convective Resistance between Indoor Air and Internal Surfaces
R_{is} - Convective Resistance between Indoor Air and Sensor
C_m - Capacitance of Internal Surfaces(Internal Walls, Floors, Furniture)
C_e - Capacitance of Envelope
C_h - Capacitance of Heater
C_i - Capacitance of Indoor Air
C_s - Capacitance of Sensor
T_i -Temperature of Indoor Air
T_a -Temperature of Ambient
T_e -Temperature of Envelope
T_h -Temperature of Heater
T_m -Equivalent Temperature of Internal Surfaces
T_s -Temperature of Sensor
A_w -Area of External Windows
A_e -Area of Envelope
ϕs -Global Solar Radiation
ϕh -Heater Radiation

For TiTmTeThTsAeRia Model,
R_{ie} - Convective Resistance between Indoor Air and Envelope
R_{ea} - Convective Resistance between Envelope and Ambient
R_{ia} - Infiltration Resistance
R_{ih} - Convective Resistance between Indoor Air and Heater
R_{im} - Convective Resistance between Indoor Air and Internal Surfaces
R_{is} - Convective Resistance between Indoor Air and Sensor
C_m - Capacitance of Internal Surfaces(Internal Walls, Floors, Furniture)
C_e - Capacitance of Envelope
C_h - Capacitance of Heater
C_i - Capacitance of Indoor Air
C_s - Capacitance of Sensor
T_i -Temperature of Indoor Air

T_a -Temperature of Ambient
T_e -Temperature of Envelope
T_h -Temperature of Heater
T_m -Equivalent Temperature of Internal Surfaces
T_s -Temperature of Sensor
A_w -Area of External Windows
A_e -Area of Envelope
ϕs -Global Solar Radiation
ϕh -Heater Radiation

Chapter 3

Methodology

The thesis now delves into discussing the modes of implementation for realising the theoretical concepts discussed in the previous chapter. In the next chapter, working of emulation software, DesignBuilder, is described. DesignBuilder is employed to model the common architectures of residences seen in Holland. Building models are displayed and their specifications are described in detail including their construction, installed systems and the building function. Data emulator has been employed due to lack of availability of sensor data for the duration of this thesis. The two major case studies introduced are further examined in chapter 7.

In chapter 5, relevant thermal networks are modelled and explained with the help of mathematical equations and the underlying heat transfer fundamentals. These models are discussed in increasing level of complexity and later applied to the case studies in chapter 7.

Chapter 6 forms the foundation for understanding the mechanisms of GA. It demonstrates the process of GA with the help of linear polynomial objective functions increasing in complexity culminating in thermal network like function. In this chapter, effects of varying input parameters on the performance of GA are also observed.

The thermal network concepts developed in chapter 5 are implemented in GA using data-sets obtained from DesignBuilder for the cases defined in chapter 4 to estimate the unknown building parameters. In the subsequent chapter, GA is applied in a similar fashion as in chapter 6 to the case studies in an attempt to understand the possible strengths and limitations of GA implementation to the building data.

As mentioned in research objectives, this study is a possible stepping stone for Opschaler project where GA will be implemented to smart meter data collected from houses all over the Netherlands. The real time data obtained from smart meters could be noisy and this provides the motivation for chapter

9 where the effect of noise on GA performance is studied.

In the conclusion, the learning acquired from the thesis is discussed with respect to the research objectives. It is followed by a section of recommendations to further the research of the thesis. Possible branches of study are listed which could improve the results obtained from coupling GA with thermal networks.

Details of DesignBuilder programming and extraneous results are shown in Appendix A.

Chapter 4

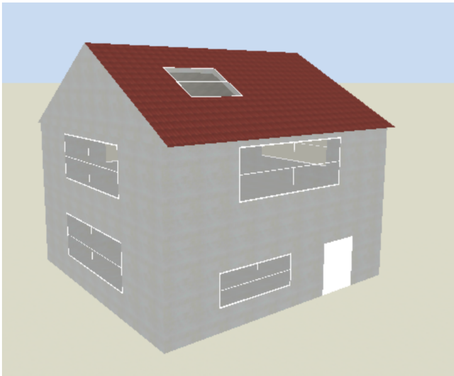
Data Emulation - Design Builder

Design Builder is an accessory software to EnergyPlus[16] (EnergyPlus is a simulation software for "building energy" modelling). Design Builder makes building modelling graphically convenient by providing a pick -and- choose tool for modelling the building components. It simulates the building taking into account the construction characteristics, installed equipment and services, activity, and occupation - controlled via schedules. It is used as an emulator, producing surrogates for actual data obtained via smart meter. The energy profile received from EnergyPlus will be used as output and the grey box approach will be employed on it to determine the building characteristics using statistical approaches. These values will then be compared to the initial values given in the software to validate the findings.

This thesis will mainly study a typical Pitched Dutch residential house along with one study on Twin Houses. They are shown in Figure 4.1.

4.0.1 Layout

Layout of the pitched roof house is shown in Figure 4.1a. The building comprises of 2 floors and an attic floor. There is a constant 30% glazing ratio i.e the windows form 30% of the envelope area. The attic has a small window installed in the tilted roof and the door of the house is facing south. Table 4.1 tabulates the areas of the facades with their orientation. The building is collectively modelled as one zone. The location is set to Amsterdam and the ground is modelled as an exterior surface hence it is exposed to the ambient conditions like the rest of the envelope.This is done to avoid modelling ground as a separate node. The programmed specifications for both case studies are shown in Table 4.2, and 4.4.



(a) A typical Pitched Roof House Architecture seen in the Netherlands

(b) A typical Twin House Architecture seen in the Netherlands

Figure 4.1: Residential Architectures in Holland

4.0.2 Heavy Built Uninsulated Building

The modelled building for this case study has a heavy built i.e the materials used for the construction have high thermal capacity and have the ability to store high amounts of heat. This also means that the temperature of the construction components will change slowly increasing the thermal constant for the building. A high thermal constant indicates lesser temperature fluctuations throughout the day.

The house is modelled with an Uninsulated template. Though there is an air gap modelled in each construction component as it can be seen in the figures 4.2-4.5.

East	West	South	North	Roof	Windows	Floor	OFA	OV
43.2	43.2	35.7	35.7	91.46	44.9	73.9	132	336

Table 4.1: Area of Envelope with Orientation of the Facades for a Single Pitched House (m^2). (OFA - Occupied Floor Area; OV - Occupied Volume;)

Parameters	Specification
Occupation	0.01 people/m^2
Wall	Heavy Weight Uninsulated Wall [$Rc = 0.663m^2K/W$]
Roof	Light Weight Uninsulated Roof [$Rc = 0.341m^2K/W$]
Internal Walls	115mm Single Leaf Brick [$Rc = 0.510m^2K/W$]
Internal Floor	300mm Concrete Block [$Rc = 0.484m^2K/W$]
External Floor	Heavy Weight Uninsulated[$Rc = 0.506m^2K/W$]
Infiltration	0.3 ach(Infiltration Case)
Glazing Type	Double Glazing 3mm with 3mm spacing[$Rc = 0.316m^2K/W$]
Heating	Boiler with Radiator(Central Heating- Air)
Ventilation	Mechanical Ventilation with no Heat Recovery(Ventilation Case)
Schedule	8am to 6pm(Monday to Saturday)

Table 4.2: Experimental Parameters of Heavy Uninsulated House

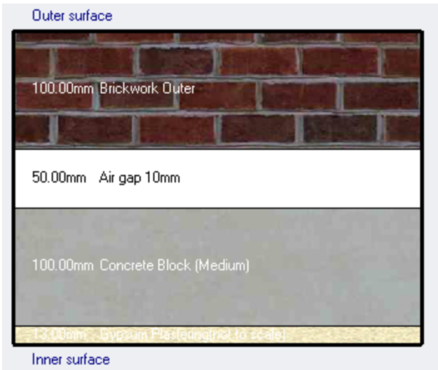


Figure 4.2: Composition of External Walls - Heavy Uninsulated

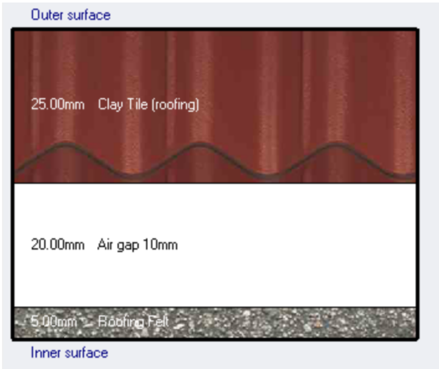


Figure 4.3: Composition of Pitched Roof- Heavy Uninsulated

Figure 4.4: Composition of External Floor- Heavy Uninsulated

Figure 4.5: Composition of Internal Floor- Heavy Uninsulated

	R_e (m^2K/W)	$R_{e,ia}$ (m^2K/W)	$R_{e,a}$ (m^2K/W)	C_{ei} (kJ/m^2K)	C_{ea} (kJ/m^2K)	C_{eqv} (kJ/m^2K)
Ext. Walls	0.49	0.13	0.04	13	136	134.8
Roof	0.20	0.1	0.04	4.017	40	4.01
Ext. Floor	0.23	0.17	0.1	32.5	3.	3.9
Int. Floors	0.21	0.1	–	176	176	176
Window	0.31	–	–	–	–	
Door	0.17	0.13	0.04	–	–	
Rf Window	0.31	–	–	–	–	

Table 4.3: Heat Transfer Properties of a Pitched Roof Heavy House (Ext -External, Int -Internal, Rf -Roof)

4.0.3 Light Built and Best Practice Insulation Building

Calculations were also made for a single pitched roof dutch house modelled as light built and best practice insulation without infiltration.

The modelled building for this case study has a light built hence the construction does not have a high ability to store heat which leads to quick changes in the building temperatures. The characteristic of low thermal constant indicates higher level of temperature fluctuations of the building.

The house is modelled with best practice insulation and contains additional insulation in the envelope which reduces the effect of outdoor conditions on the indoor space. The house designed has the specifications shown in Table 4.4. The house is modelled with no infiltration or mechanical ventilation.

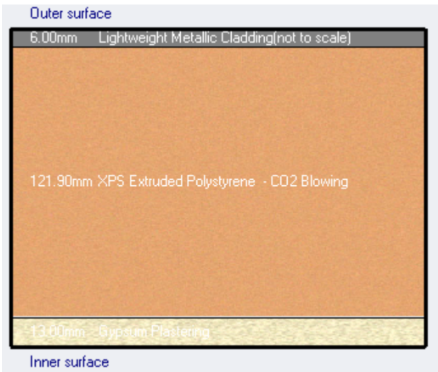


Figure 4.6: Composition of External Walls - Light Best Practice

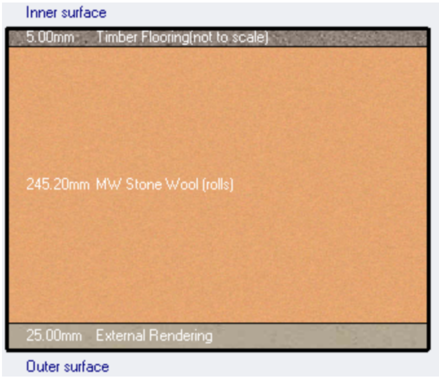


Figure 4.7: Composition of External Floor - Light Best Practice

Figure 4.8: Composition of Internal Floor - Light Best Practice

Figure 4.9: Composition of Roof - Light Best Practice

Parameters	Specification
Occupation	0.01 people/m^2
Wall	State of the Art Light Weight Wall [$Rc = 3.808 m^2K/W$]
Roof	Uninsulated Light Weight Roof [$Rc = 0.341 m^2K/W$]
Internal Floor	100mm Concrete Slab [$Rc = 0.484 m^2K/W$]
External Floor	State of the Art Light Weight Floor[$Rc = 6.4863 m^2K/W$]
Infiltration	0.3 ach(Infiltration Case)
Glazing Type	Double Glazing 3mm with 3mm spacing[$Rc = 0.316 m^2K/W$]
Heating	Boiler with Radiator
Ventilation	Mechanical Ventilation with no Heat Recovery(Ventilation Case)
Schedule	8 am to 6 pm (Monday to Saturday)

Table 4.4: Experimental Parameters of Single Pitched Roof -Light House

	R_e (m^2K/W)	$R_{e,ia}$ (m^2K/W)	$R_{e,a}$ (m^2K/W)	C_{ei} (kJ/m^2K)	C_{ea} (kJ/m^2K)	C_{eqv} (kJ/m^2K)
Ext. Walls	3.63	0.13	0.04	13	7.5	17.26
Roof	0.20	0.13	0.04	4.017	40	4.017
Ext. Floor	6.21	0.17	0.1	0.0039	32.5	6.29
Int. Floors	0.07	0.1	–	176	176	88.2
Window	0.31	–	–	–	–	
Door	0.17	0.13	0.04	–	–	
Rf. Window	0.31	–	–	–	–	

Table 4.5: Heat Transfer Properties of a Pitched Roof Light House (Ext -External, Int -Internal, Rf -Roof)

The EnergyPlus output data-set consists of the variation of indoor temperature, the transmitted solar radiation through windows, and instantaneous heater usage through the winter months (Oct 1 to March 31).

The two case studies are programmed such that the heating is switched on in instances when there is more heat loss from the building than there is heat gain.As the heating schedule is set to 8 am to 6pm, solar gains are present throughout the HVAC schedule. For heavy un-insulated case, losses from the building are higher than the light best practice building with the same solar gain. Hence, there are more instances of heating in heavy case. After 6pm, when there is no solar gain or HVAC, heavy building radiates heat slowly and steadily in the next hours whereas the light building radiates more heat in the hours right after sunset and less heat in the later hours.

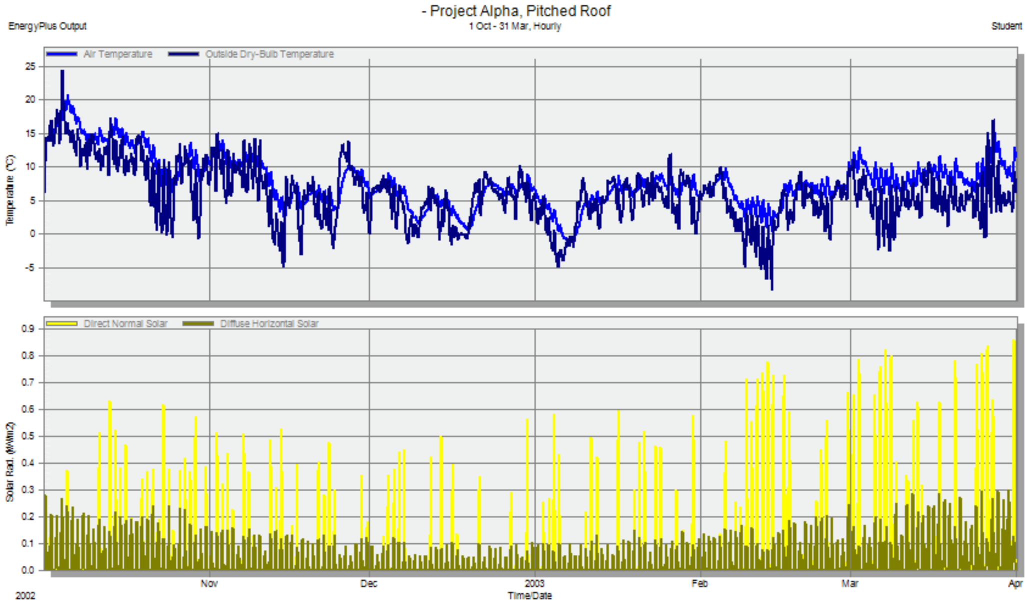


Figure 4.10: Hourly Outdoor Temperature and Solar Radiation For Winter Months

The variation of outdoor temperature and solar radiation for the location of Amsterdam is shown in figure 4.10. The weather file is obtained for the city of Amsterdam from the EnergyPlus directory. The weather files of EnergyPlus are available only in an hourly format and this restricts this study to dealing with hourly analysis of the energy in the building. Design builder can be used to obtain results in sub hourly format by means of interpolating the weather file according to the formulae given in the documentation of EnergyPlus. However, this is a cumbersome process and in addition, the non-linearity of the weather parameters reduces the accuracy of the interpolation methodology. Hence, this study focuses only on the available data-set to perform hourly analyses.

Figure 4.10 shows the output given by Design Builder. Indoor temperature profile shows that the temperature varies from ~ 0 °C to ~ 22 °C which is appropriate for the outdoor temperature profile of Amsterdam given that there are no set-points provided. Correspondingly, heater usage shown in Figure 4.11 is present only during the programmed day schedule and is highest for the lowest indoor temperature observed in mid February.

Figure 4.12 and 4.13 demonstrate the temperature of internal and external

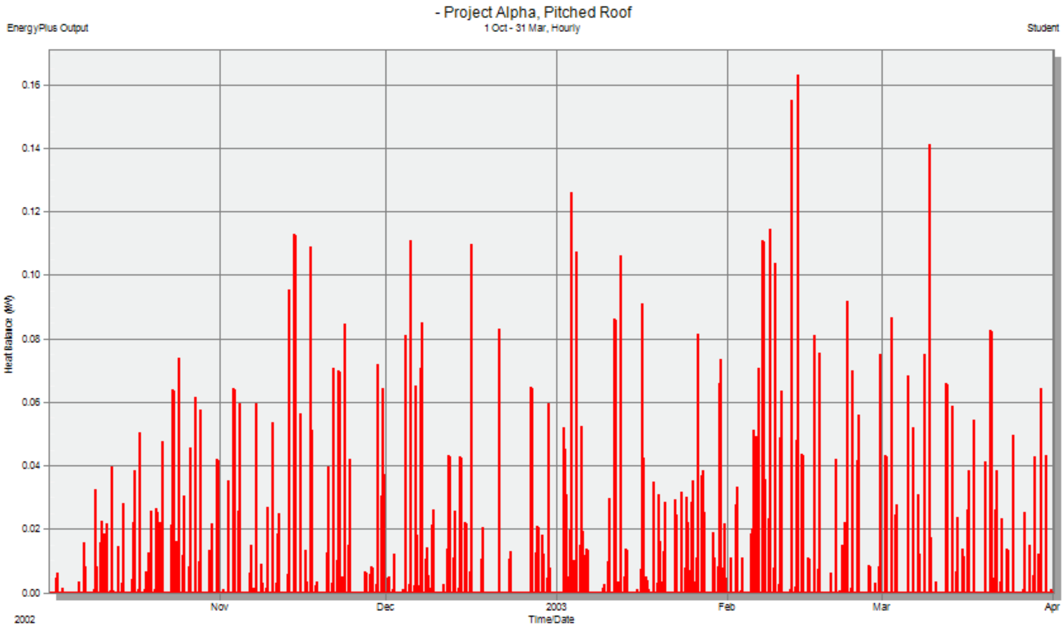


Figure 4.11: Hourly Sensible Heating For Winter Months of Heavy Uninsulated Case

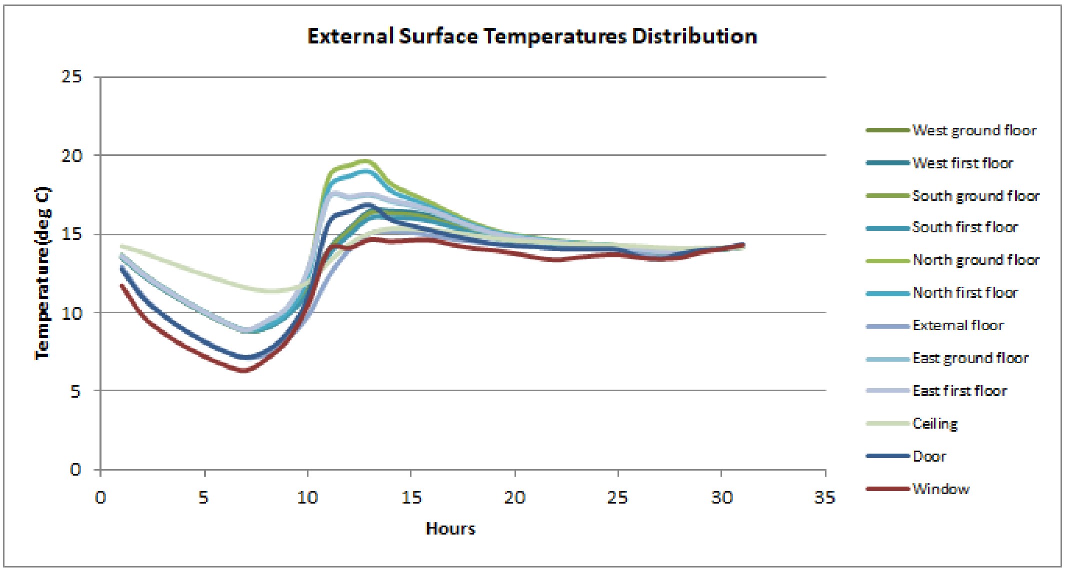


Figure 4.12: Hourly External Surface Temperatures of All Facades of Heavy Uninsulated Case

surfaces of the house. As seen, the surface temperatures of window, door, and external ground floor is low at nights and hence if the three components

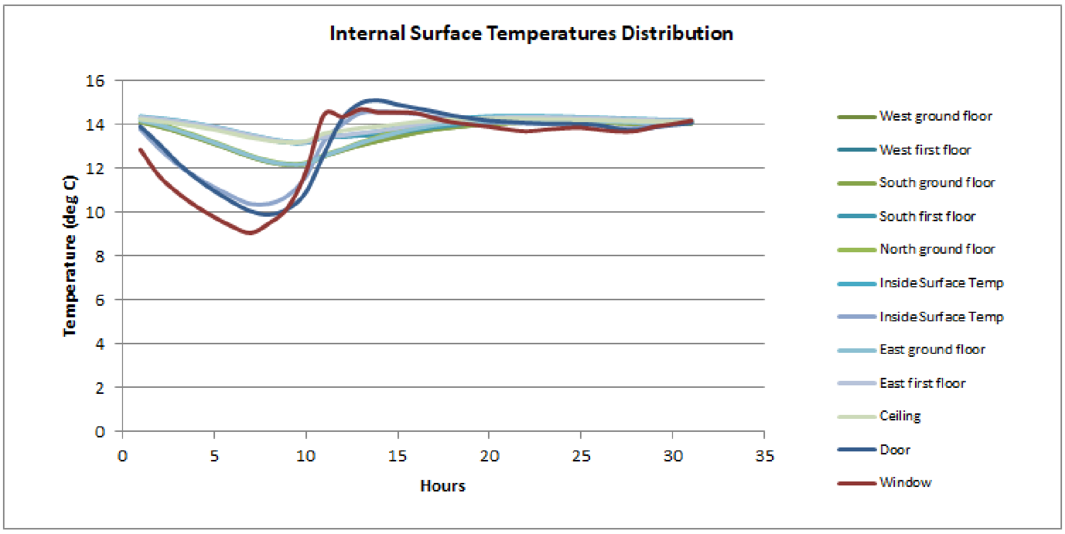


Figure 4.13: Hourly Internal Surface Temperatures of All Facades of Heavy Uninsulated Case

are not considered separately, errors could ensue in the two and three node networks.

Chapter 5

Sensible Thermal Networks

This chapter will deal with the formulation of thermal networks to model the buildings analysed in this thesis. Clarke[17] represents all possible energy flow-paths for a building in Figure 5.1. The processes shown in the schematic form the basis for modelling relevant thermal networks in the chapter.

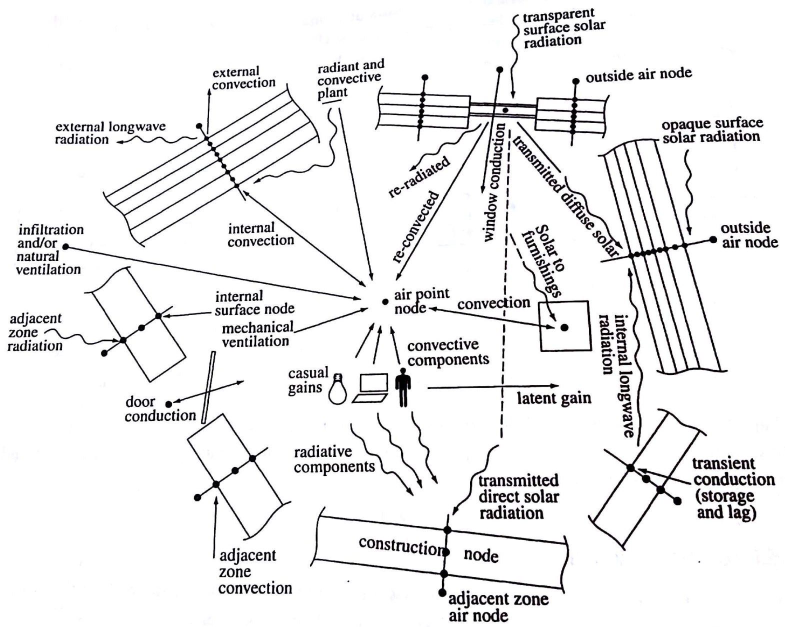


Figure 5.1: Building Energy Flow-paths by Clarke

Thermal networks are formulated by reducing the energy flow-paths to strictly the major contributors to the dynamics of the building. Hence, furniture is neglected for the purpose of this study and the entire building is modelled as one zone to eliminate the energy flows introduced by adjacent zones.In an attempt to simplify the network, the nodes of windows, doors and floor are included in the envelope node. The floor is assumed to be exposed to the ambient conditions to aid the integration into the envelope. The transparent facades are considered to be thin and do not store heat hence there is no time lag involved for windows.

For the thermal networks modelled in the report, it is assumed that the indoor air has negligible capacitance because of constant airflow due to ventilation. It is also assumed that the temperature of envelope is uniformly distributed in each model. Indoor Air is assumed to be perfectly mixed hence the indoor conditions are uniform everywhere inside the building. Extra thermal mass such as the internal walls are accounted for in the heat capacity and resistance of the envelope. For a three node model, this capacitance and resistance is combined along with the internal envelope node. The internal gains by appliances and lighting are neglected for all models.

It is known from literature that the processes occurring between ambient and indoor air can be roughly displayed in the form of thermal networks as can be seen in Figure 5.2. The thermal network is based on the lecture by Laure Itard[18].

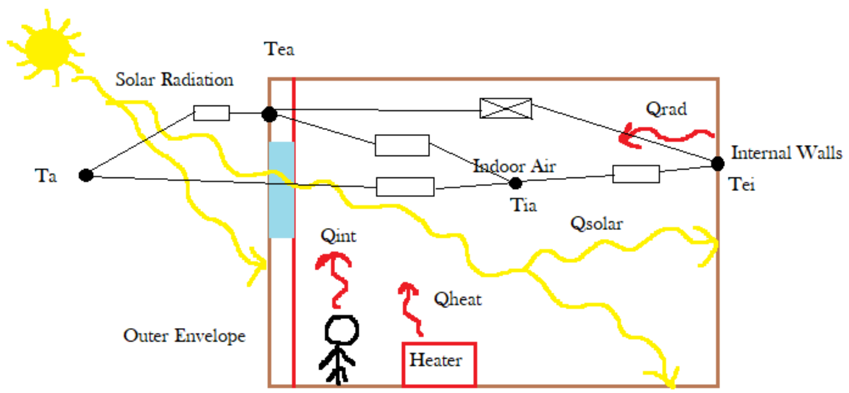


Figure 5.2: Thermal Network Presentation of Thermal Processes

As can be seen from figure 5.2, there are four nodes named Ambient Temperature node(node T_a), Envelope External Surface node (node T_{ea}), Indoor Air node (node T_{ia}), and Internal Surface of the External Envelope (node T_{ei}). This model is a more accurate representation owing to the fact that large temperature differences exist between the temperatures of outer and inner surfaces of the envelope as the outer surface is at a temperature close to the ambient temperature and internal surface has a temperature close to the indoor air temperature. The heat flux received by all these nodes are also split into their respective elements. Hence, only the convective gains of the heat flux are directed to the indoor node while the internal surfaces of the envelope receive the radiant fraction of the same heat flux. External surfaces get heat flux from the amount of solar radiation falling on the surface. Solar flux for the interiors is the transmitted solar gains through the windows. Nomenclature of the models is based on the amount of thermal nodes excluding the ambient node as it is the thermal reference node.

The thermal network is redrawn in a more comprehensive format in Figure 5.3

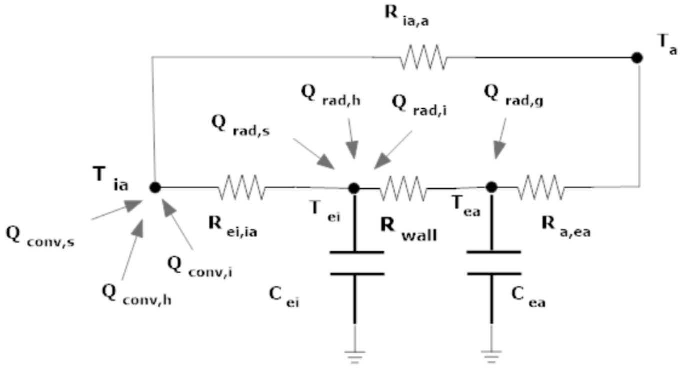
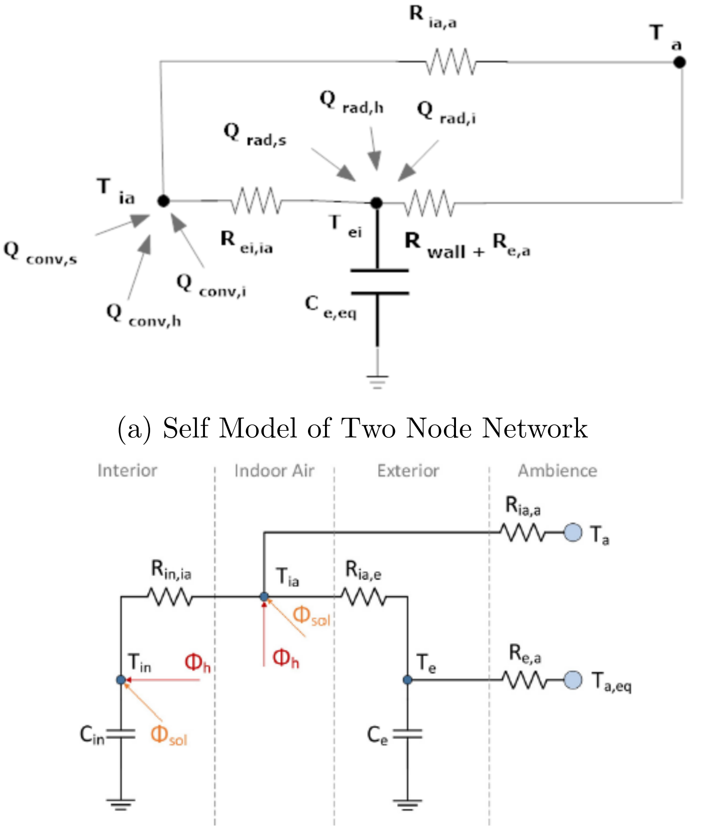


Figure 5.3: Three Node Thermal Network

However, the focus of this thesis is to apply the model to real life data to determine building parameters. In view of this, obtaining the temperature data for three nodes might not be practical.Also, measuring the temperature of the outdoor envelope may create practical problems especially seeing that the outdoor radiation may not affect the indoor dynamics to the extent of the transmitted solar radiation. For that reason, the thermal network is minimised to a two nodes thermal network shown in Figure 5.4a. This thermal network has aggregated the internal and external surfaces of the envelope into one node being the internal surface of the external envelope. Due to the aggregation, capacitance of the surfaces is now shown by a new equivalent capacitance which will be some form of average of the capacitance of the outer and inner surface. Also, the resistance between the ambient and envelope node is now a summation of the convective resistance and the resistance of the wall.The radiation from the sun entering the building is considered to be absorbed by the inner surfaces through radiation and convection increasing the temperature of the wall. This heat is radiated to the indoor air and other components. There is also heat radiated and convected from the heater. Indoor air is assumed to be circulating which convects the heat from the walls. Finding the correct resistance is more important than capacitance for the purpose of this thesis as for a yearly data analysis, effects of capacitances are



(a) Self Model of Two Node Network

(b) 4R2C Model By Harb

Figure 5.4: Comparison of Self made Model with 4R2C Model by Harb

nullified but resistance effects on the indoor conditions remain. This model could be ideal towards achieving this goal. However, the preceding three node model will also be tested to characterise model type and associated accuracy of results.

It is interesting to note that the three node model is a reduced version of the 4R2C model. If the interior node of furniture and internal walls is removed and the capacitance is incorporated in the capacitance of the envelope while convective resistance is incorporated in the convective resistance between envelope and indoor air, the models are equivalent.

As enumerated in Section 2.5, the most complex thermal networks do not work well due to over complications within the model. Similarly, the simplest model is oversimplified hence lacking the intricacies needed to produce accurate results. Thus, it is expected that one of the two node or three node network will produce reasonable results.

For all thermal networks in this thesis, indoor air node is assumed to be massless hence it cannot store any heat within and has zero capacitance. Heater is not considered as a node, with the radiation from the heater being split into indoor air and envelope respectively. Also the T_{eq} is replaced by T_{ea} in three node model. A further reduced version of the three node model will also be tested to check whether the parameters can be obtained with only data from the indoor air and ambient.

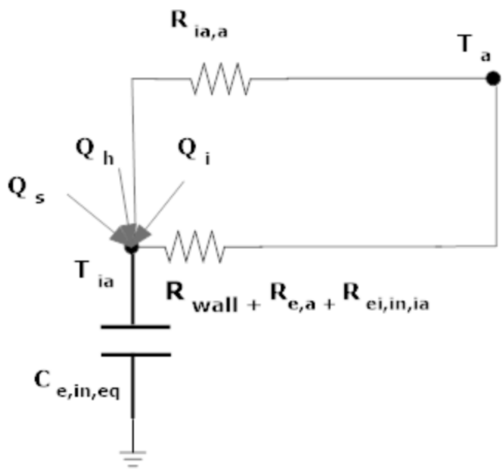


Figure 5.5: Self Made Simplistic Model

The network in figure 5.6 consists of only two nodes, ambient and indoor air. In this model, the heat flux does not need to be split into its components as all the effects are aggregated directly on the indoor air node.

5.1 Global RC Network

Now we will discuss the model shown in Fig 5.6.

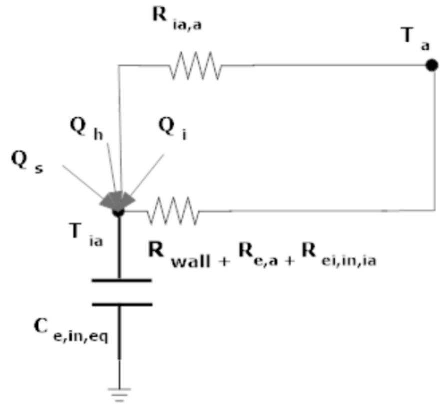


Figure 5.6: Self Model of Three Node Network

This network raises the question, what is the physical significance of equivalent resistance and capacitance? Equivalent resistance is the effective resistance the building displays to heat transfer between indoors and outdoors. Hence, a higher resistance would mean there is less exchange of heat and a better insulation. Also as the effective resistance incorporates the resistance of all the components of a building carrying their heat transfer in parallel, the effective resistance is n times smaller (assuming that the n components in the system have the same resistances). The resistance of the house will then be the combined resistance of the convective resistances on the inner and outer surfaces of the external walls along with the heat conductivity of the envelope. Envelope here will include every component that is responsible for the heat flux flow to the indoor air, which would be the external walls(exposed to outdoor air), roof(exposed to outdoor air), windows and doors(exposed to outdoor air) and ground(exposed to ground temperature). The resistances of the internal walls and floors will not be a part of equivalent resistance as these processes are occurring within the system.

Equivalent Capacitance is every surface between the indoors and outdoors capable of storing heat coming off the solar radiation, heater and internal gains. Hence, for this particular calculation each of the internal and external components storing the heat will be included. If in an experiment, it is assumed that all the heat gains are only interacting only with the floors (which is a common assumption in classical simulations and a choice in EnergyPlus Simulations) then the heat capacity of the house will reflect only the capacitance of the floors. This is because capacitance is a measure of stored heat. For application of a more detailed thermal network where the capacitances are dealt with in different equations, C_e would be negligible as it is not participating in storing heat.

The node for indoor air includes the equivalent capacitance of the building. $R_{a,ia}$ includes the convective resistance of the envelope surfaces and internal walls in parallel, combined with the resistance of wall in series. Infiltration resistance is considered separately. There is no need to separate the heat flux because all of them combine into one node.$C_{e,in,eq}$ in this case can be reduced to $C_{e,eq}$ as there is no furniture or internal thermal mass present for our study. The governing equations for this model are,

Balance on T_{ia} Node,

$$A = (T_{ia,n+1} + T_{ia,n}) + \frac{1}{R_{eqv}C_{e,in,eq}}(T_a - T_{ia})dt + \frac{1}{C_{e,in,eq}}(Q_{sol} + Q_h + Q_i)dt \tag{5.1}$$

Equations for R_{eqv} and $C_{e,in,eq}$ are ,

$$\begin{aligned} R_{eqv} &= \frac{(R_{wall} + R_{ea,a} + R_{ei,in,ia}) * R_{ia,a}}{(R_{wall} + R_{ea,a} + R_{ei,ia} + R_{ia,a})} \\ C_{e,in,eq} &= (C_{ei} + C_{int,floor} + C_{roof} + C_{ext,floor}) \end{aligned} \tag{5.2}$$

where,

T_{ia} - Temperature of Indoor Air(K)
T_a - Temperature of Ambient Air(K)
$R_{ei,ia}$ - Convective Resistance between indoor air and inner surfaces of envelope (K/kW)
R_{wall} -Resistance of envelope(K/kW)
$R_{ea,a}$ - Convective Resistance between outdoor air and envelope(K/kW)
Q_i - The heat gains because of occupants (kW)
Q_s - The solar gains transmitted to the indoor air through the window as well as through envelope radiation(kW)
Q_h - The heat supplied by the heating system(kW)
$C_{e,in,eq}$ - Equivalent Heat Capacity of Interior and Exterior Envelope(kWh/K)
R_{eqv} - Equivalent series resistance of the building(K/kW)

5.2 Two Node Network

In this section, a description of the chosen model (see figure 5.7) is given.

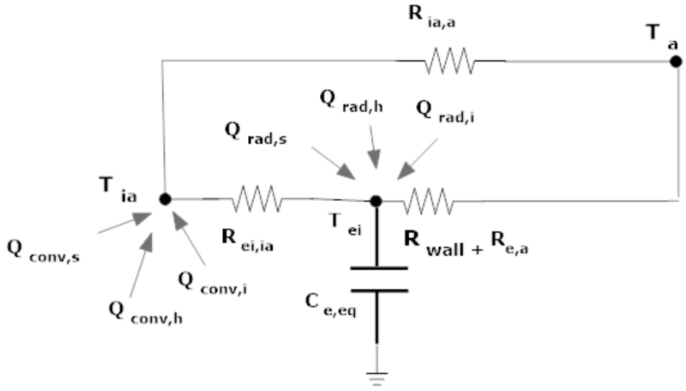


Figure 5.7: Self Model of Two Node Network

Envelope refers to the external walls, roof and ground floor. Hence, it exchanges heat by convection with the ambient air and the indoor air. It absorbs the solar radiation falling on the outer surface and the transmitted solar gains on the internal surfaces. Other sources of heat like the heater, internal gains are also major components interacting with the envelope. Coefficients can be added to all the internal and solar gains to assess the amount of gains interacting with the envelope.

The governing equations for this model are,

Balance on T_{ia} Node,

$$0 = \frac{1}{R_{ei,ia}}(T_{ei}-T_{ia})dt+\frac{1}{R_{ia,a}}(T_a-T_{ia})dt+(f_{conv,s}\phi_{sol}+f_{conv,h}\phi_h+f_{conv,i}Q_i)dt. \tag{5.3}$$

The node for internal envelope surface, T_{ei}, includes the capacitance of the external surface of envelope. Balance on T_{ei} Node,

$$\begin{aligned} dT_{ei} = \frac{1}{R_{ei,ia}C_{e,eq}}(T_{ia}-T_{ei})dt + \frac{1}{(R_{ea,a}+R_{wall})C_{e,eq}}(T_a - T_{ei})dt+ \\ \frac{1}{C_{e,eq}}(f_{rad,s}\phi_{sol}+f_{rad,h}\phi_h+f_{rad,i}Q_i)dt, \end{aligned} \tag{5.4}$$

$$\begin{aligned} Q_{rad,s} &= f_{rad,s}\phi_{sol}; \\ Q_{rad,h} &= f_{rad,h}\phi_h; \\ Q_{rad,i} &= f_{rad,i}Q_i; \\ Q_{conv,s} &= f_{conv,s}\phi_{sol}; \\ Q_{conv,h} &= f_{conv,h}\phi_h; \\ Q_{conv,i} &= f_{conv,i}Q_i; \end{aligned} \tag{5.5}$$

where,

T_{ei} - Temperature of Internal Surfaces(K)
$f_{conv,s}$ - convective contribution of incoming solar radiation to the indoor air.
$f_{conv,h}$ - convective contribution of heater flux to the indoor air.
$f_{conv,i}$ - convective contribution of internal gains to the indoor air.
$f_{rad,h}$ - Fraction of total heat given by the heater radiated
$f_{rad,s}$ - Fraction of total solar gains radiated
$f_{rad,i}$ - Fraction of total internal gains radiated
ϕ_s - Incoming solar radiation absorbed by internal components(kW)
Q_i - The heat gains because of occupants, lighting, catering etc (kW)
ϕ_h - The heat supplied by the heating system(kW)
$Q_{conv,s}$ - Convective Fraction of incoming solar radiation absorbed by internal components(kW)
$Q_{conv,i}$ - Convective Fraction of The heat gains because of occupants, lighting, catering etc (kW)
$Q_{conv,h}$ - Convective Fraction of The heat supplied by the heating system(kW)
$Q_{rad,s}$ - Radiative Fraction of incoming solar radiation absorbed by internal components(kW)
$Q_{rad,i}$ - Radiative Fraction of The heat gains because of occupants, lighting, catering etc (kW)
$Q_{rad,h}$ -Radiative Fraction of The heat supplied by the heating system(kW)

5.3 Three Node Network

Now we will discuss the model shown in figure 5.8.

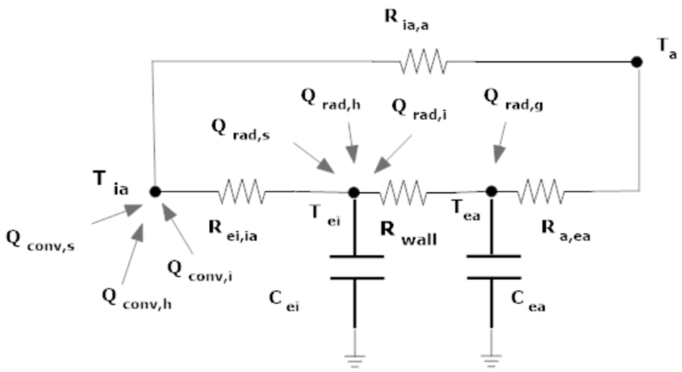


Figure 5.8: Self Model of Three Node Network

The additional node T_{ea} refers to the outer surface of the envelope. Hence ,heat processes for this node include the solar radiation falling on the surface and the convective heat exchange between the ambient and the outer surface. There is also conduction in between the two envelope surfaces. The same term for conduction is also added in the T_{ei} node.

The governing equations for this model are,

Balance on T_{ia} Node,

$$0 = \frac{1}{R_{ei,ia}}(T_{ei} - T_{ia})dt + \frac{1}{R_{ia,a}}(T_a - T_{ia})dt + (f_{conv,s}\phi_{sol} + f_{conv,h}\phi_h + f_{conv,i}Q_i)dt \tag{5.6}$$

Balance on T_{ei} Node,

$$\begin{aligned} dT_{ei} = \frac{1}{R_{ei,ia}C_{ei}}(T_{ia} - T_{ei})dt + \frac{1}{R_{wall}C_{ei}}(T_{ea} - T_{ei})dt \\ + \frac{1}{C_{ei}}(f_{rad,s}\phi_{sol} + f_{rad,h}\phi_h + f_{rad,i}Q_i)dt \end{aligned} \tag{5.7}$$

Balance on T_{ea} Node,

$$dT_{ea} = \frac{1}{R_{a,ea}C_{ea}}(T_a - T_{ea})dt + \frac{1}{R_{wall}C_{ea}}(T_{ei} - T_{ea})dt + \frac{1}{C_{ea}}(f_{rad,g}Q_{sol})dt \tag{5.8}$$

$$Q_{rad,g} = f_{rad,g}Q_{sol} \tag{5.9}$$

where,

T_{ea} - Temperature of External Envelope(K)
$f_{rad,g}$ - Fraction of absorbed solar radiation falling on external envelope radiated
$Q_{rad,g}$ -Radiative Fraction of the solar radiation falling on the external envelope(kW)

5.4 Summary of Thermal Networks

	Global Network	Two Node Network	Three Node Network
N_{eq}	1	2	3
$N_{unknowns}$	2	6	9
Acc_{Res}	Low	High	Low

Table 5.1: Summary of Thermal Networks (N_{eq} - Number of equations; $N_{unknowns}$ - Number of unknowns; Acc_{Res} - Estimated accuracy of results).

We know from table 5.1 that the number of equations and unknowns reflects on the difficulty in solving the equations. The third row of expected accuracy demonstrates the interplay of thermal network models with parameter estimation limitations using GA.

Global Network would theoretically be the easiest to solve, however, global model might not be the best representation of real values of resistance and capacitance. This leads to low expected accuracy. Two Node Model has six unknowns which could be reasonably well predicted by GA and the thermal model also includes more intricacies of the heat dynamics. The expected accuracy is hence high. Three Node Model has nine unknowns which could be difficult to determine for GA especially without narrow bounds, which would be difficult to provide. The thermal model does include more intricacies but it is not expected to overcome the estimation challenges.

Chapter 6

Familiarising with GA

Like every algorithm, GA, despite claims of flexibility to varied problem statements, has its limitations. One example is an objective function with an unknown variable without a corresponding coefficient. An example of that would be a three variable linear equation with a constant. The mathematical representation is :

$$A = ax + by + cz + d \tag{6.1}$$

where, a, b, c and d are the unknown coefficients to be determined by GA.

Another possible limitation of GA to be explored is the number of variables GA can predict accurately. It is expected to break down at some point if the number of unknowns are too large - this occurs due to decreasing accuracy as an error in one prediction can be compensated by a corresponding negative error in another variable/variables. Methods to improve the behaviour of GA under such circumstances will also be tested. The amount of data points needed by GA for an accurate prediction will also be studied.

The parameters set of GA is as below,
Bounds : Adjusted in Each Case
Population Size : 50
Creation Function : Uniform
Scaling Function : Rank
Selection Function : Stochastic Uniform
Elite Count : 0.05 * Population Size
Crossover Fraction : 0.8
Crossover Function : Scattered
Mutation Function : Adaptive Feasible
Stall Generations : number of variables * 100
Function Tolerance : 1E-06

6.1 Linear Equation with a Constant

First case study is a linear equation with a constant to analyse GA's capabilities to estimate parameters without coefficients (w). In this case, we go a step further and consider 4 unknowns instead of 2 as taken previously. The equation to be modelled is given as:

$$a = xb + yc + zd + w \tag{6.2}$$

Artificial data corresponding to (w, x, y, z) = (9.6, 4, 5, 6) for the variables a, b, c, and d is created for 25 points. Data created for a, b, c and d is created by using a random integer generator function. Final results shown in figure 6.1 are given bounds of [0,0,0,0] and [10,10,10,10].

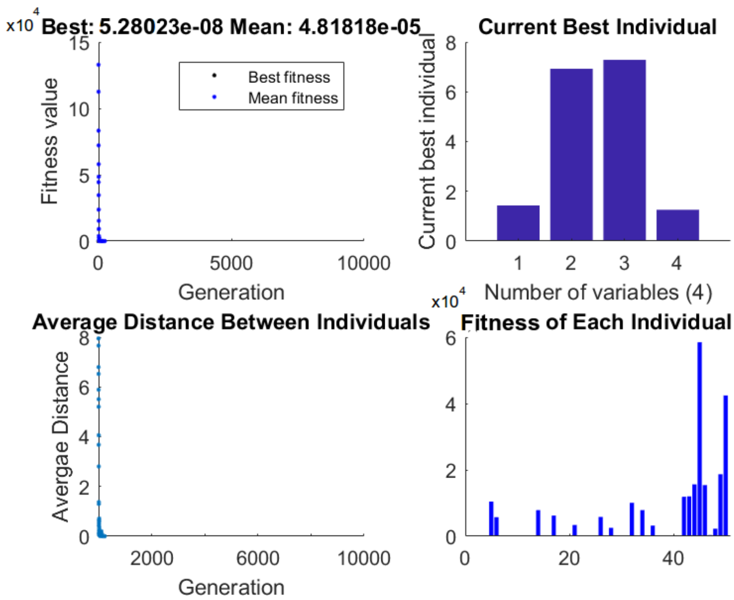


Figure 6.1: GA computed parameters for Linear Line with Constant

As it can be seen GA could not converge to the right result.The possible reasons for the failure is that the parameters x, y and z are adjustable to any value with the presence of a free variable w. The possible solution could be to provide more data points and to definitely tighten the bounds whilst increasing the constraint tolerance to 10^-12.

	w	x	y	z
Expected Results	9.6	4	5	6
GA Results	1.681	6.33	6.84	1.432

Table 6.1: Values of GA computed parameters for a Line With Constant

6.2 Linear Equation With 4 Unknowns

This subsection focuses on identifying the limits of GA with respect to the number of unknowns using a linear equation with 4 unknowns. The equation considered is similar to the previous case study but for the lack of a variable without a coefficient. This change allows to better study the behaviour of GA towards prediction a variable with and without a coefficient. The modelled equation is given below:

$$A = aw + bx + cy + dz \tag{6.3}$$

Artificial data for the variables a, b, c, d is created for 25 points using Microsoft Excel. The unknowns GA has to determine are w, x, y, z, the values for which are respectively set to 29, 5.1, 2.34, 8.5.

These parameters can be taken as default for all further studies unless specified otherwise.

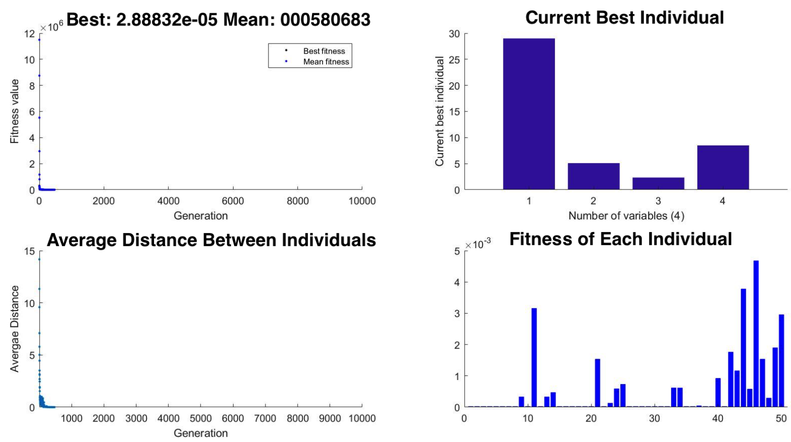


Figure 6.2: GA computed parameters for Linear Line without Constant

GA converged to perfect results with no final bounds provided. Initial bounds are given as [0,0,0,0]. The above results show that the algorithm converged after proper exploration of space as the average distance between individuals varied reasonably. The best fitness value is quite low and so is the mean fitness value. This is also shown in the score plot. Hence, the results are perfect without any adjustments to GA.

w	x	y	z
29	5.1	2.34	8.5

Table 6.2: Values of GA computed parameters for a Line With No Constant

6.3 Thermal Network Equation With 3 Unknowns Using Synthetic Data

This case is a thermal network look-alike equation modelled as:

$$A = \frac{a}{xy} + \frac{b}{yz} + \frac{c}{y} + \frac{d}{y} \tag{6.4}$$

This equation is considered to better understand how the thermal network equation will behave when modelled with GA. As a thermal network has unknowns R_i and C_i in the denominator similarly this equation has x, y and z as unknowns in the denominator. Like the thermal network equations with temperature differences known, for this case a, b, c and d are known.

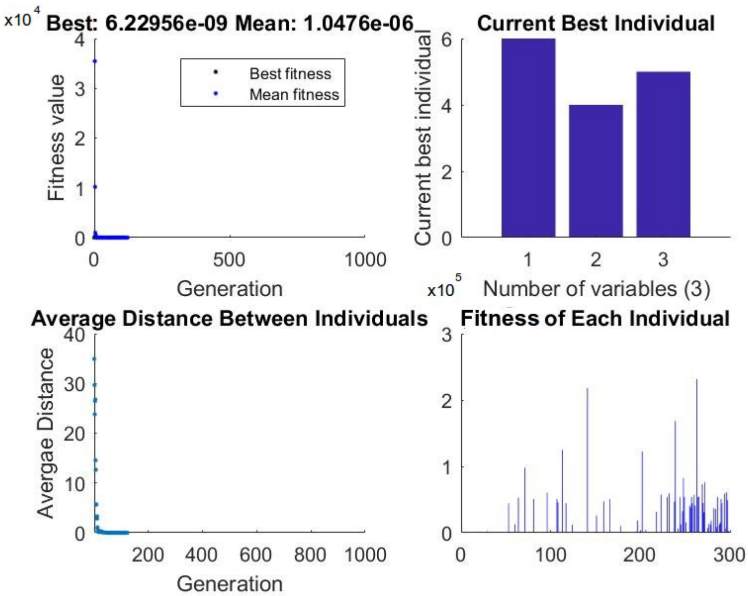


Figure 6.3: GA computed parameters for Thermal Network Equation With Three Unknowns

Similar to previous case studies, artificial data for the variables a, b, c and d is created for 25 points and the values for x, y and z are respectively fixed to 6, 4, 5. Results are shown in Figure 6.3. The results are accurate with no adjustment and same initial bounds and no final bounds as the last case.

x	y	z
6	4	5

Table 6.3: Values of GA computed parameters for a Thermal Network Like Equation With Three Unknowns

6.4 Thermal Network Equation With 6 Unknowns Using Synthetic Data

To understand whether the higher node models will behave well with GA, the below equation is modelled with 6 unknowns.

$$A = \frac{a}{uv} + \frac{b}{vw} + \frac{c}{wx} + \frac{d}{xy} + \frac{e}{vz} + \frac{f}{v} + \frac{g}{v} \tag{6.5}$$

The values of u, v, w, x, y and z are given in Table 6.4. Bounds are adjusted to [0,0,0,0,0,0] and [100,100,100,100,100,100]

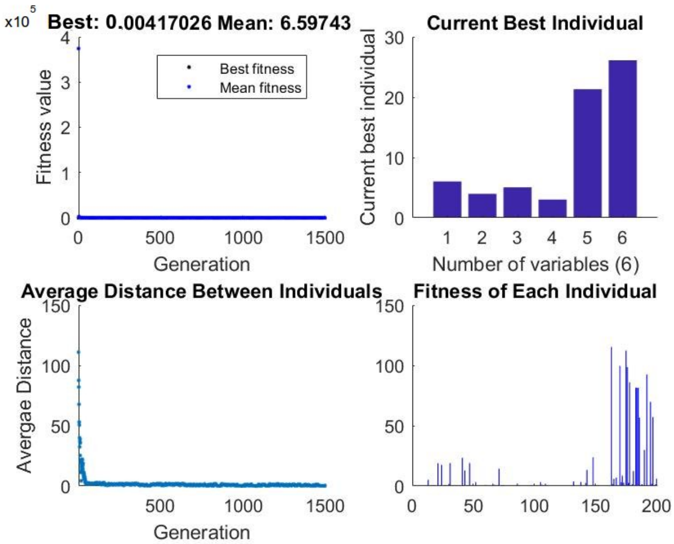


Figure 6.4: GA computed parameters for Thermal Network Equation (Six Unknowns) with 2000 Generations

The computed values for the unknowns are u = 6.043, v = 3.955, w = 5.118, x = 3.014, y = 49.577, z = 26.711. It is also observed that the algorithm gave premature results as the default number of iterations ran to the maximum and the algorithm did not converge. A solution to this would be to increase max generations to 1000 - the algorithm was found to still stop prematurely. Results for 2000 stall generations are shown in Figure 6.4

GA prediction with increased generations are more accurate with the following estimated values:

	u	v	w	x	y	z
Expected Results	6	4	5	3	16	25
Default Stall Gen (600)	6.043	3.955	5.118	3.014	49.577	26.711
2000 Stall Gen	6.008	3.993	5.019	3.009	16.924	25.527

Table 6.4: Values of GA computed parameters for Thermal Network Equation (Six Unknowns) with Increase in Stall Generations

Hence, after adjustment of stall generations ,it can be concluded that GA can predict nearly perfect values for all 6 variables. To study the effect of

Population Size, we increase the population to 100, 200, and 300 and compare the results to identify any improvement due to varied population sizes (see Figure 6.5 and Table 6.5. Results for 200 and 300 population are shown in Appendix A.3.2.

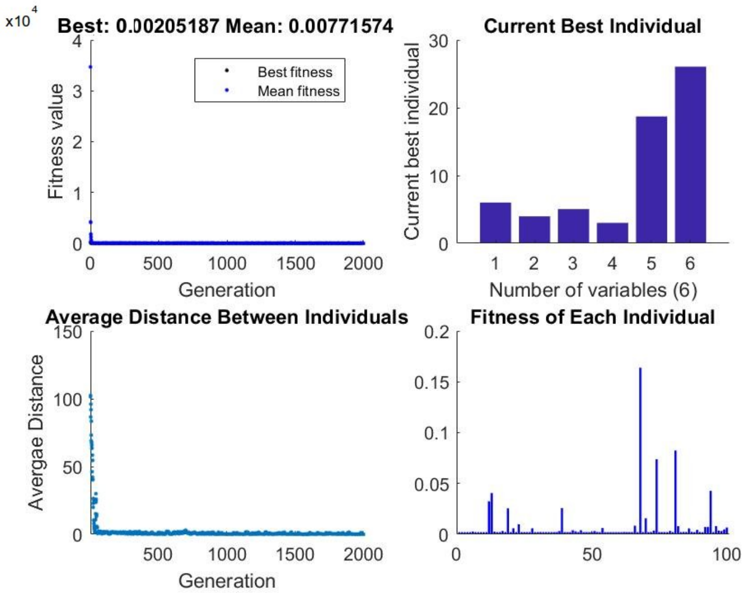


Figure 6.5: GA computed parameters for Thermal Network Equation (Six Unknowns) for Population Size of 100

	u	v	w	x	y	z
Expected Results	6	4	5	3	16	25
100 Population	6.022	3.984	5.043	3.016	18.722	26.06
200 Population	5.99	3.995	5.017	3.006	16.757	25.354
300 Population	6.003	3.996	5.009	3.005	16.437	25.231

Table 6.5: Comparison of Results with Change in Population Size

It can be seen that an increase the population results in a reduced number of iterations and an improved prediction accuracy.Another important parameter of GA to study is the number of data points. All previous studies have been performed with 25 data points. This was augmented with further studies with 250 data points (see figure 6.6 and 2500 data points, the results of which are shown in table 6.6. The remaining results for more than 250 data points are shown in Appendix A.3.3.

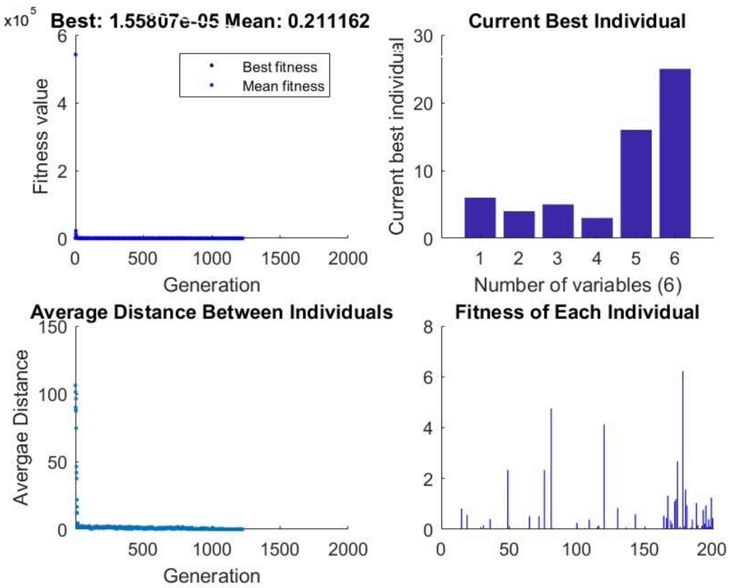


Figure 6.6: GA computed parameters for Thermal Network Equation (Six Unknowns) with 250 Data Points

	u	v	w	x	y	z
Expected Results	6	4	5	3	16	25
250 Data Points	6.0003	3.905	5.001	3.000	15.999	25.007
2500 Data Points	6.0004	3.999	5.000	3.0002	16.012	25.011

Table 6.6: Comparison of Results with Increase in Data Points

A convergence is observed at 250 data points itself, and any further increase in data points can be deemed irrelevant for this case.

6.5 Effect of Noise in Convergence

One of the most important measures of compatibility of an algorithm for parameter estimation is its accuracy to a noisy input data.This section is discussed along with the results in Chapter 9.

Chapter 7

Implementation of thermal network in GA

As elucidated in Section 2.3.4, GA estimated the parameters correctly and provided an insight into the capabilities of GA towards parameter estimation. Objective function for the determination of thermal parameters of building is based on the equations shown in Chapter 5.

Genetic Algorithms are modelled in MATLAB by providing data points which reflect the hourly readings of the outdoor temperature, solar radiation, sensible heating demand and indoor temperature.
The objective function of GA is the error between the calculated and the measured indoor temperature of the room. To identify the exact coefficients of the equation, the measured indoor temperature and the heater demand are needed which are provided by EnergyPlus results of the building model. Once the inputs are added into the equation, the unknown coefficients are determined by GA given the upper and lower bounds of both the variables.

7.1 Overview of Case Studies

7.1.1 Heavy Uninsulated

House modelled in EnergyPlus shown in Figure 4.1a is specified to have a heavy built and uninsulated pre-templates for construction. The house is modelled with no infiltration or mechanical ventilation.

Expectations from the behaviour of the building lie in its design. Heavy built suggests a big heat store in the walls which would imply a slower increase in wall temperature. As the building is uninsulated, conduction through the fabric will be high. For the case studies discussed below, there is no infiltration or mechanical ventilation but the internal and external coefficients are fixed to 3 W/m2K and 11.95 W/m2K respectively hence the inner resistance is expected to be 1/4th of the outer resistance. The profile for heavy uninsulated building will be such that the indoor temperature will drop quickly after heating is switched off as the heat conducts through the fabric uninhibited. High thermal mass helps increase the temperature when it falls. Also , due to a lack of insulation, the house heats up very quickly during day time when outdoor temperature rises but walls take longer time to heat. This phenomena would mean that the radiative effect will be considerable because of temperature difference of the internal surfaces.

Medium Typical Reference Case Study is shown in Appendix A.4.1.

7.1.2 Light Best Practice

House modelled in EnergyPlus is now specified to have a light built and best practice insulation pre-templates for construction. The house is modelled with no infiltration or mechanical ventilation.

Expectations from the behaviour of the building lie in its light built.It suggests a small heat capacity which would imply a sharp increase in wall temperature in the presence of heat sources.Similarly, there is a sudden decrease in temperature at night. This phenomenon is heightened by the high insulation as heat will not be easily allowed to pass through creating hot spots. It will further heat up the surfaces leading to higher rise and drop in the space. For this case study, there is the same inner and outer convective resistances.

All the results in the chapter are based on data from November to April because of the phenomena of mid-October summer in Netherlands. Hence, in order to include October, the equation would need to include cooling

7.2 Calculation from DB

This section deals with giving insight on the process of determination of resistances and capacitance required for analysis of GA results.

It is known from Chapter 5, the resistances required are the internal surface resistance ($R_{ei,ia}$), external envelope resistance ($R_{ea,a}$) and the resistance of the envelope (R_e).

Equivalent Resistance can be calculated by using the relation below,

$$U_{eqv} = \frac{\sum_{i=1}^{n} U_i A_i}{\sum_{i=1}^{n} A_i} \tag{7.1}$$

The components considered are the external walls, roof, windows, door, and ground floor. Internal floors are not included in equivalent calculation as the building is modelled as one zone. U_i includes the convective, conductive and radiative heat transfer coefficient on ith surface.

Internal and external resistances can be determined by using the following equations,

$$U_{ei} = \frac{\sum_{i=1}^{n} U_{i,in} A_i}{\sum_{i=1}^{n} A_i} \tag{7.2}$$

$U_{i,in}$ is the convective heat transfer coefficient of the inner surface of component i given by DB.

$$U_{ea} = \frac{\sum_{i=1}^{n} U_{i,out} A_i}{\sum_{i=1}^{n} A_i} \tag{7.3}$$

$U_{i,out}$ is the convective heat transfer coefficient of the outer surface of component i given by DB.

$U_{e,ia}$ and $U_{ea,a}$ are calculated depending on the surfaces they interact with. Hence, for internal heat coefficient, additional surface of internal floors is incorporated in the calculation.

Resistance can be calculated by inverting the heat transfer coefficient ,

$$R = \frac{1}{U} \tag{7.4}$$

Now, determination of capacitance from DB is discussed. Calculation of Equivalent Capacitance can be carried out by multiplication of internal heat capacity obtained from Design Builder by the surface area. It should be noted here that DB provides equivalent internal heat capacity for each component (walls, floor, roof) in $\frac{kJ}{m^2K}$. Hence multiplication of this value with the area of the envelope would include the capacitance on both the surfaces.

$$C_{eqv} = C_{wall}A_{wall} + C_{roof}A_{roof} + C_{gr,floor}A_{gr,floor} + C_{int,floor}A_{int,floor} \quad (7.5)$$

C_i is the equivalent capacitance of the component i given by DB.

Calculating the internal and external capacitance is carried out by assuming that the thermal penetration depth of the first construction layer.

$$C_{ei} = C_{wall,in}A_{wall,in} + C_{gr,floor,in}A_{gr,floor,in} + 3C_{int,floor}A_{int,floor} \quad (7.6)$$

$C_{i,in}$ is the capacitance of the innermost layer of component i given by DB. Internal Floor capacitance is added thrice as there are three surfaces exposed to the indoor space including the ceiling.

$$C_{ea} = C_{wall,outl}A_{wall,out} + C_{roof,out}A_{roof,out} + C_{gr,floor}A_{gr,floor} \quad (7.7)$$

$C_{i,out}$ is the capacitance of the outermost layer of component i given by DB.

7.3 Global Model

The objective function for this model from Global Network will have the following equation after removal of infiltration resistance,

$$A = (-T_{ia,n+1} + T_{ia,n}) + \frac{1}{X(1)X(2)}(T_a - T_{ia})dt + \frac{1}{X(2)}(Q_{sol} + Q_h + Q_i)dt \quad (7.8)$$

The global model will now be applied first to the heavy built un-insulated building followed by the light built best practice building and the corresponding building parameters will be estimated using GA. The order of model application to the different buildings remains constant for all models.

7.3.1 Heavy Built Uninsulated Building

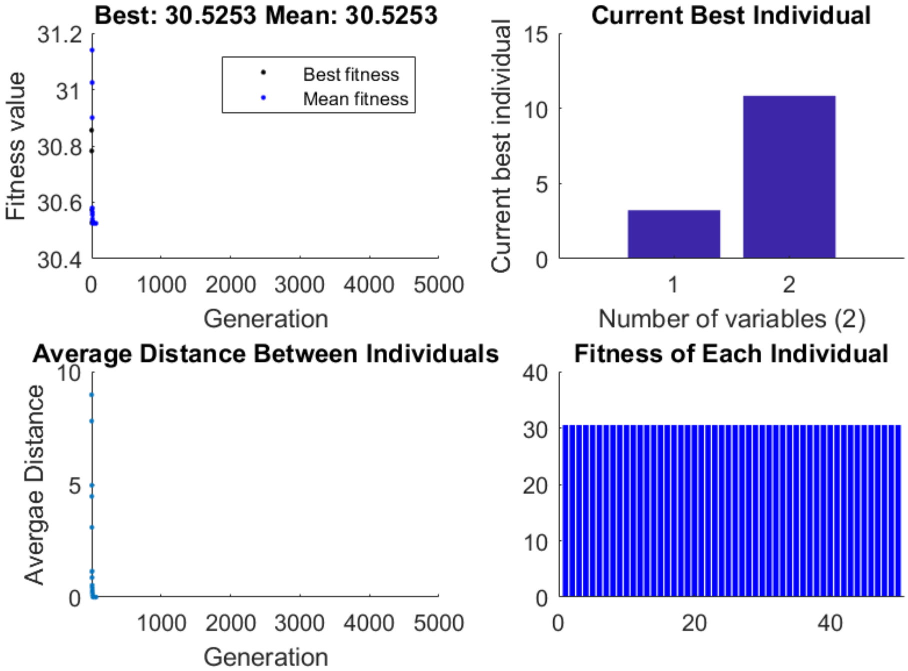


Figure 7.1: GA computed parameters for Single Pitched Roof House - Heavy Weight Uninsulated Global Network

	R_eA (K/W)	$R_{ei,ia}A$ (K/W)	$R_{ea,a}A$ (K/W)	$C_{ei}A$ (kWh/K)	$C_{ea}A$ (kWh/K)	$C_{eqv}A$ (kWh/K)
Exterior Walls	3.76	0.98	0.303	0.642	6.724	4.94
Roof	2.20	1.09	0.43	0.102	1.016	0.10
Ground Floor	3.18	0.32	0.15	0.569	0.06	0.08
Internal Floors	2.89	1.35	–	6.16	–	5.43
Window	7.04	2.89	–	–	–	–
Door	76.52	56.52	17.913	17.39	–	–
Roof Window	98.0049	–	–	–	–	–
	0.83	0.32	0.156	8.47	6.08	10.55

Table 7.1: Calculated Equivalent Heat Transfer Properties for the Pitched Roof Heavy House

GA parameter	Thermal Parameter	GA Value	Theoretical Value	Error
X(1)	R_{eqv}	3.22 K/kW	1.306 K/kW	146%
X(2)	C_{eqv}	10.82 kWh/K	10.55 kWh/K	2.5%

Table 7.2: Analysis of GA computed results for Heavy Uninsulated Model - Global Network

The fitness value of 30 as shown in Figure 7.1, is for approximately 3000 hours of data. It suggests that error between measured and calculated temperature every hour is very low which implies that the result found by GA for the global network equation is a good fit for the data. Though as it can be seen the results are not correct hence global network equation is not an accurate representation of the thermal behaviour of the building.It is also observed that capacitance was accurately estimated whereas the equivalent resistance has a large error.This can be due to the objective equation format as it contains two instances in the denominator for optimising the capacitance but only one for resistance. In addition, the term including the resistance is in the form of multiplication hence the resistance parameter has more freedom for optimisation than the capacitance parameter. It is also seen that the average distance dropped quickly in this case which can be explained by GA easily finding the global optima. This can be due to limited number of unknowns. Plot for Fitness of each individual is displayed to prove convergence of results. As the fitness of all the individuals in the last generation is equal, this result cannot be improved further.

7.3.2 Light Built Best Practice Building

Results of GA are obtained with the help of 4100 hours of raw data is shown in figure 7.2,

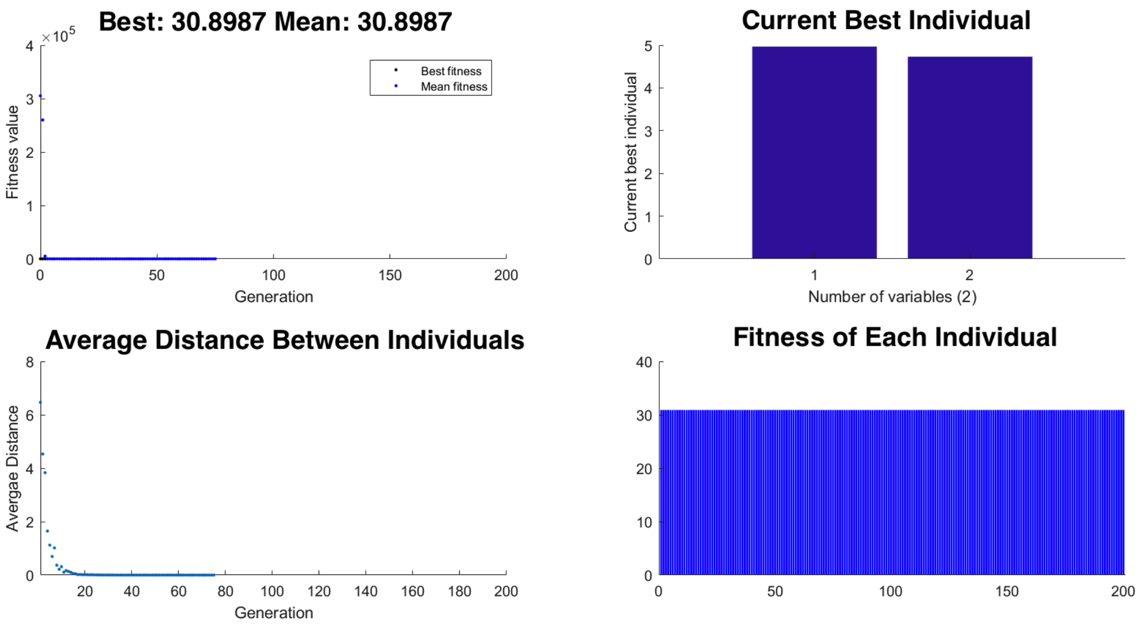


Figure 7.2: GA computed parameters for Single Pitched Roof House - Light Weight Best Practice Insulation Global Network

	R_eA (K/W)	$R_{ei,ia}A$ (K/W)	$R_{ea,a}A$ (K/W)	$C_{ei}A$ (kWh/K)	$C_{ea}A$ (kWh/K)	$C_{eqv}A$ (kWh/K)
Exterior Walls	27.5	0.98	0.303	0.642	0.37	4.94
Roof	2.205	1.09	0.43	0.102	1.016	0.10
External Ground Floor	84.02	2.29	1.35	0.068	0.56	0.08
Internal Floors	0.96	1.35	–	6.1818	–	5.43
Window	7.04	–	–	–	–	–
Door	76.522	0.13	0.04	–	–	–
Roof Window	98	–	–	–	–	–
	1.49	0.32	0.156	3.83	1.95	3.57

Table 7.3: Calculated Equivalent Heat Transfer Properties for the Pitched Roof Light House

GA parameter	Thermal Parameter	GA Value	Theoretical Value	Error
X(1)	R_{eqv}	4.96 K/kW	1.96 K/kW	153%
X(2)	C_{eqv}	4.72 kWh/K	3.57 kWh/K	33%

Table 7.4: Analysis of GA computed results for Light Best Practice Model

Similar trend is seen for the results of this case study. The fitness values are similar to the last case and error between expected and GA results is considerably high. It can also be seen that results for light case study are worse than the heavy case study.This can be explained by the fact that there are higher instances of heating in heavy building because of delayed time constant. GA similarly gave precedence to capacitance over resistance to optimise the objective function as it did for the heavy case study.To improve the results, two node network will be applied which is expected to better model the thermal behaviour.

7.4 Two Node Model

Implementation of Single Pitched Roof House in Two Node Network :

$$A = (-T_{ei,n+1} + T_{ei,n}) + \frac{(T_{ia} - T_{ei})}{X(1)X(2)} + \frac{(T_a - T_{ei})}{X(3)X(2)} + \frac{(X(4)Q_{sol} + X(5)Q_h + X(6)Q_i)}{X(2)} + \frac{(T_{ei} - T_{ia})}{X(1)} + ((1 - X(4))Q_{sol} + (1 - X(5))Q_h + (1 - X(6))Q_i) \tag{7.9}$$

7.4.1 Heavy Built Uninsulated Building

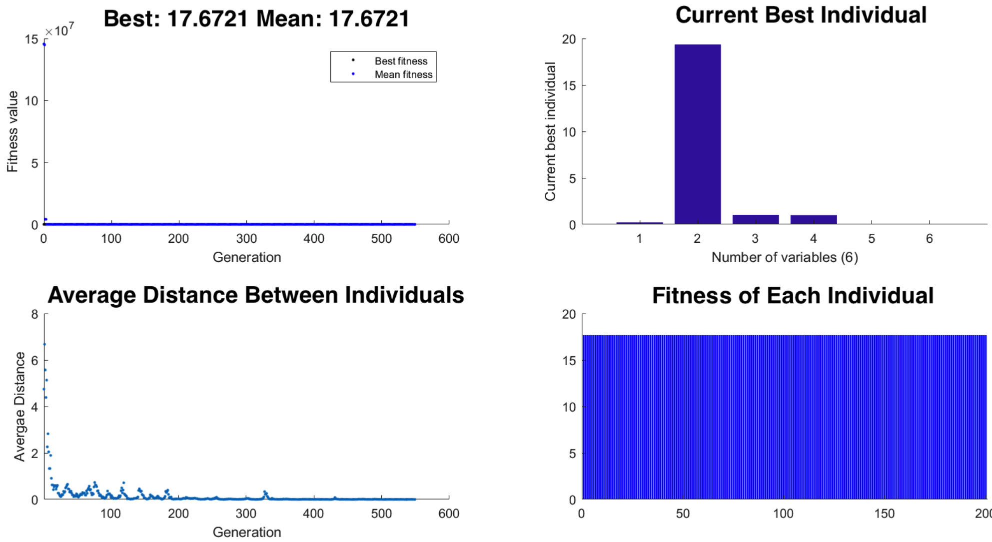


Figure 7.3: GA computed parameters for Single Pitched Roof House - Heavy Weight Uninsulated Insulation Two Node Network

GA parameter	Thermal Parameter	GA Value	Theoretical Value	Error
X(1)	$R_{ia,ei}$	0.22K/kW	0.32	31%
X(2)	C_{eqv}	19.35 kWh/K	10.55	84%
X(3)	$R_e + R_{ea,a}$	1.02K/kW	0.98	4%
X(4)	f_{sol}	0.99	0.99	0%
X(5)	f_h	1.21E-7	0	
X(6)	f_i	9.3E-9	0	

Table 7.5: Analysis of GA computed results for Heavy Uninsulated Model - Two Node Network

The results display the value of 17 for the best fitness value which is lower than the fitness value for the global network case. This suggests that there is an improvement in data fitting. Improvement in results can also be seen in the decreasing error in the predicted values.

f_{sol} value of 0.99 is justified by the fact that the solar radiation is absorbed by the internal thermal mass and the solar convective effect is very low as expected.According to Harb[14], convective effect of solar radiation can be assumed to be 0.09 which is in line with the results. Low values for f_h and f_i

can be explained by the fact that Heating is modelled as fully convective in the DesignBuilder file and internal gains are reported for zero radiant fraction.

There is a notable error in internal resistance which is due to collective modelling of windows and floor with the heat storing surfaces.It creates errors as windows and floor are examined with respect to the calculated Mean Radiant Temperature (MRT) instead of the measured internal temperature of the surfaces respectively.

The error in the results can also be explained by the fact that indoor air's capacity cannot be neglected for our case studies as there is no infiltration and ventilation considered for the models.

Apart from the above reasons, it is important to understand that in the objective equation for the two node network, the term containing $(T_{ei} - T_a)$ has the largest contribution because of larger temperature differences .This leads to a higher need for GA to estimate $R_e + R_{ea,a}$ accurately. In addition, the term $(T_{ei} - T_{ia})$ can produce a bigger bias compared to the heat flux terms limiting the flexibility of $R_{ei,ia}$. The remaining parameter of capacitance is then adjusted more freely to lower the objective function value.

7.4.2 Light Built Best Practice Building

For this case, the node equation for envelope is run for the optimisation and Matlab Results were found to be reasonably accurate.

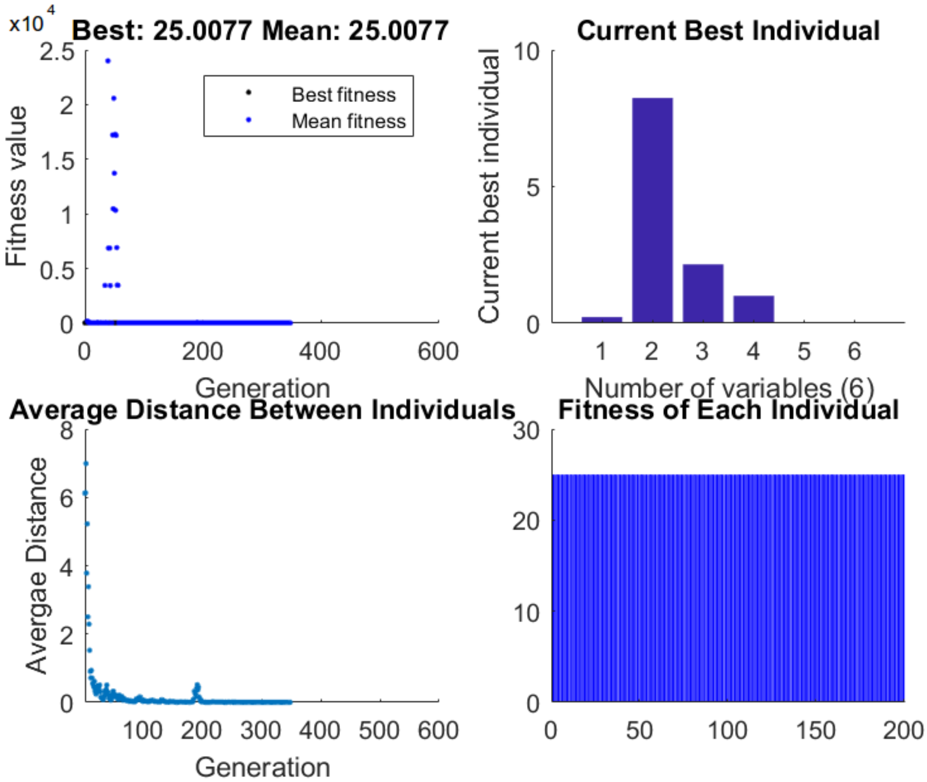


Figure 7.4: GA computed parameters for Single Pitched Roof House Without Infiltration - Light Weight Best Practice Insulation Two Node Network

GA parameter	Thermal Parameter	GA Value	Theoretical Value	Error
X(1)	$R_{ia,ei}$	0.225K/kW	0.32	30%
X(2)	C_{eqv}	8.24 kWh/K	3.57	130%
X(3)	$R_e + R_{ea,a}$	2.15K/kW	1.65	30%
X(4)	f_{sol}	0.99	0.99	0%
X(5)	f_h	2.5E-7	0	
X(6)	f_i	4.6E-9	0	

Table 7.6: Analysis of GA computed results for Light Best Practice Model - Two Node Network

The value of fitness function for this case is also observed to be lower than the global network case suggesting two node network is better suited for all kinds of buildings. The results given by GA are in accordance with the reasoning given for the heavy case study.

7.5 Three Node Model

In the third and final case, we study the three node model.

$$\begin{aligned}A =&(-T_{ei,n+1}+T_{ei,n})+\frac{T_{ia}-T_{ei}}{X(1)X(2)}+\frac{T_{ea}-T_{ei}}{X(3)X(2)}+\frac{X(7)Q_{sol}+X(8)Q_h+X(9)Q_{int}}{X(2)}\\&+\frac{T_{ei}-T_{ia}}{X(1)}+(1-X(7))Q_{sol}+(1-X(8))Q_h+(1-X(9))Q_{int}-\\&(-T_{ea,n+1}+T_{ea,n})+\frac{T_a-T_{ea}}{X(4)X(5)}+\frac{T_{ei}-T_{ea}}{X(3)X(5)}+\frac{X(6)Q_{sol}}{X(5)}\end{aligned} \tag{7.10}$$

7.5.1 Heavy Built Uninsulated Building

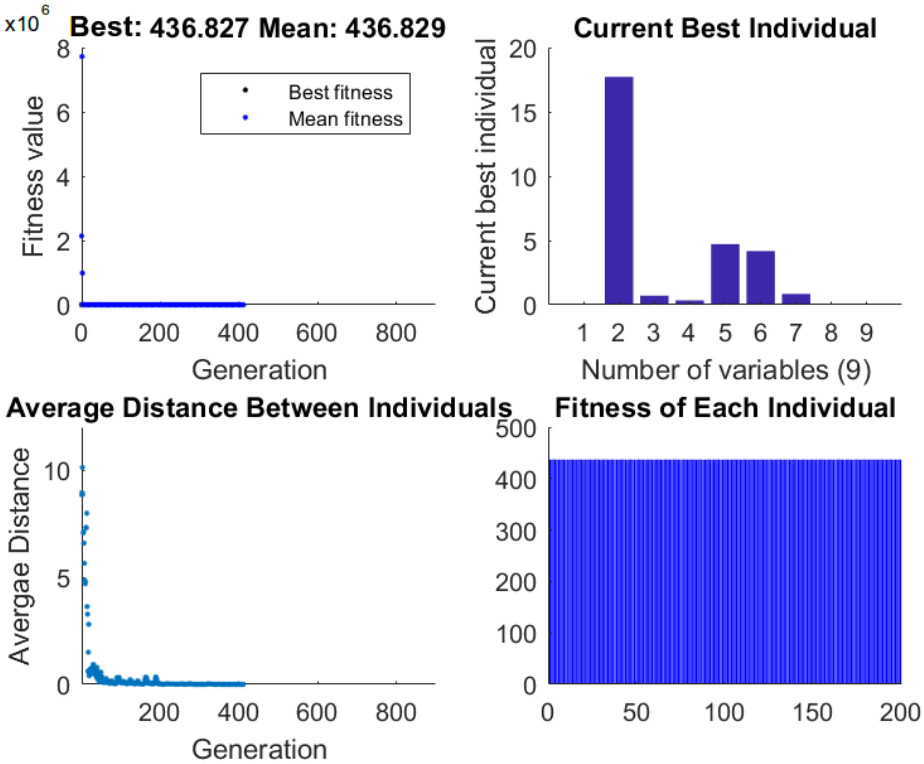


Figure 7.5: GA computed parameters for Single Pitched Roof House Without Infiltration - Heavy Uninsulated Three Node Network

GA parameter	Thermal Parameter	GA Value	Theoretical Value	Error
X(1)	$R_{ia,ei}$	0.099K/kW	0.32	69%
X(2)	C_{ei}	17.72 kWh/K	8.47	109%
X(3)	R_e	0.7097K/kW	0.83	14%
X(4)	$R_{ea,a}$	0.34K/kW	0.156	117%
X(5)	C_{ea}	4.72kWh/kW	6.08	22%
X(6)	$x_{sol,ea}$	4.17	–	
X(7)	$f_{sol,ei}$	0.84	0.99	15%
X(8)	$f_{h,ei}$	1E-4	0	
X(9)	$f_{int,ei}$	3.2E-5	0	

Table 7.7: Analysis of GA computed results for Heavy Uninsulated Model - Three Node Network

The converged fitness value for this network is much higher than global and two node network. Hence, three node network is not a better fit for the data when compared to the two node network. It can be seen in the error of the values predicted by GA. The high error in C_{ei} can be explained by the fact that theoretically the internal capacitance will include the capacitance of the second layer (concrete) because of thermal penetration depth. Penetration depth will play a higher role in heavy walls as they have heavy internal components which contribute to the heat store.

There is also errors in determining the convective resistances on both the surfaces.It can be due to error in calculation of external envelope surface temperature. The temperature for outer surface is calculated by area weighted mean of all surfaces including windows, doors and floor however, it is assumed that all windows are at the same temperature which could introduce slight errors.

The objective function for the three node network has nine parameters hence understanding the possible interplay whilst estimation could be difficult. However, the results could be explained similarly to the global and two node network. It is seen that C_{ea} is better estimated than C_{ei} because there is a presence of a free parameter ,$R_{ea,a}$ in the equation for node T_{ea}. In the node for T_{ei}, there is no free parameter as both $R_{ei,ia}$ and R_e are present in at least one more term.Though R_e is a parameter which needs to be more precise because of the possibly large bias it could produce leading $R_{ea,a}$ and C_{ei} to lower accuracy.

7.5.2 Light Built Best Practice Building

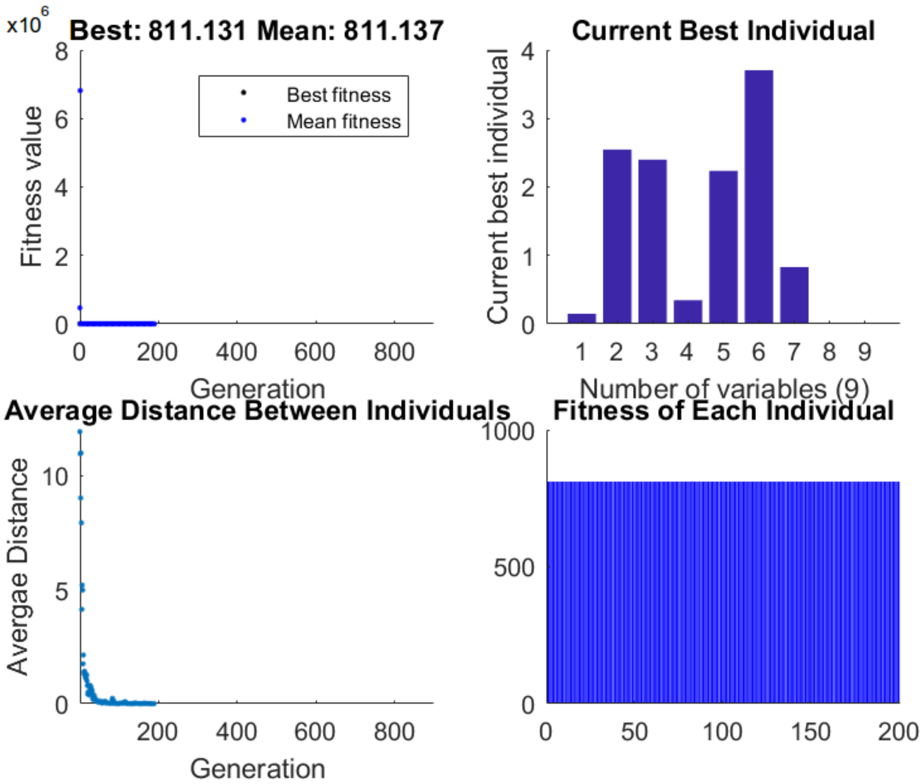


Figure 7.6: GA computed parameters for Single Pitched Roof House Without Infiltration - Light Weight Best Practice Insulation Three Node Network

GA parameter	Thermal Parameter	GA Value	Theoretical Value	Error
X(1)	$R_{ia,ei}$	0.146K/kW	0.32	50%
X(2)	C_{ei}	2.54kWh/K	3.83	33%
X(3)	R_e	2.39K/kW	1.49	60%
X(4)	$R_{ea,a}$	0.34K/kW	0.156	117%
X(5)	C_{ea}	2.23kWh/kW	1.956	14%
X(6)	$x_{sol,ea}$	3.7	–	
X(7)	$f_{sol,ei}$	0.82	0.99	17%
X(8)	$f_{h,ei}$	2.6E-5	0	
X(9)	$f_{int,ei}$	8.1E-5	0	

Table 7.8: Analysis of GA computed results for Light Best Practice Model-Three Node Network

The results for this case study show less error for C_{ei} as light built does not have a heat store beyond the first construction layer. Also, when the resistances are compared for the two cases, it is observed that GA could not find as good a fit for the data for light building as it did for heavy building. The lack in fitting is exhibited by higher fitness value. it is seen that there is a lower error in the values of convective resistances for the light house. Hence it demonstrates the resistance of envelope plays a bigger role in the building than the convective resistances as there is larger error present in R_e in light building. It is in line with the fact that without infiltration and ventilation, the convective resistances do not play a huge part in the inner dynamics.

The difference in results can also be explained by the fact there for an insulated light fabric ,the temperature differences between the surfaces and air are much lesser leading to lesser bias produced for the terms containing the particular temperature differences. Now that there is less dependence on those terms, the objective function is reduced to a global function look like. Similar to the reasoning of global network, there is more flexibility for R_e as compared to the capacitance in the equations. Amongst the convective resistances, $R_{ia,ei}$ requires to be more fixed as it has an independent term to optimise. The next step is to determine the change in performance of GA with change in building parameters and input data.

Chapter 8

GA Performance - Varying Parameters

An important question to answer before running the algorithm is: which data points will give the best result? Would it be only the day points when there is solar radiation? Would it be only when all the HVAC systems are all operating? Would it be when the house is occupied?Whats the least number of hours that can be provided such as to produce a good result?Will GA perform well under changing circumstances like inclusion of ventilation/infiltration.

8.1 Absence Of Solar Gains, Occupancy Gains and Heater Gains

As the heating and cooling schedule is combined with occupancy, HVAC Systems are switched off in moments of no occupancy.

Also in this case occupancy is programmed to be in the day time so for a subset of the occupancy hours, there is always a presence of solar radiation. The result presented in Figure 8.1 is for the times when none of the gains are present. The equation should be adjusted to exclude the heating demand, solar gains and the occupancy gains.

$$A = (-T_{ei,n+1} + T_{ei,n}) + \frac{(T_{ia} - T_{ei})}{X(1)X(2)} + \frac{(T_a - T_{ei})}{X(3)X(2)} + \frac{(T_{ei} - T_{ia})}{X(1)} \tag{8.1}$$

This study is particularly included to note whether the accuracy of results improve with the night data. At night there is no solar radiation, which would exclude one of the important factors in the equation and hence might lead to better results with lesser input. But the data set for nights will not be able to give the right results as the equation depends on the dynamic changes in node temperatures which will be wrong because of transitions between nights of each day. Also by excluding the day time data sets, we exclude the effect of heat capacitance in the building. The results added for this section are based on two node network equation as there was least error observed for that thermal network in the previous chapter.

Below are some results obtained after running the parameter estimation of one winter night.

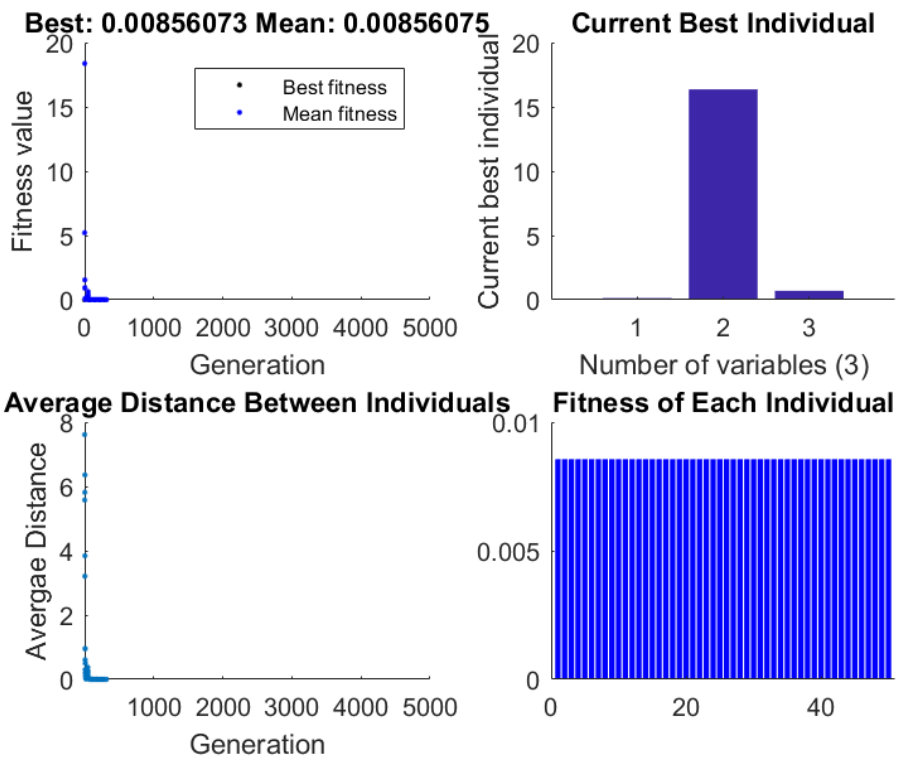


Figure 8.1: GA computed parameters for Single Pitched Roof House For One Night Data -Heavy Uninsulated Case Two Node Network

GA parameter	Thermal Parameter	GA Value	Theoretical Value	Error
X(1)	$R_{ia,ei}$	0.13K/kW	0.32	59%
X(2)	C_{eqv}	9.92kWh/K	10.82	8%
X(3)	$R_{ea,a}$	1.17K/kW	0.98	19%

Table 8.1: GA computed parameter values for For One Night Data -Heavy Uninsulated Case

As it can be seen, results are reasonably accurate and provide a good indication towards the results. Though because of limited input data, GA converged to plenty of possible results. The above result displayed is one of the closest results given by GA. In this case, only way to be certain of a result is to give strict bounds which can aid GA by limiting the search space. Nevertheless, GA will converge to more than one result which represents a huge shortcoming for this case. In this case, final bounds are given as [1,20,5].

This method does help to conclude that if data could be obtained when HVAC Systems are switched off and there is no occupancy, it decreases the process interdependencies and its easier to realise a good result.

8.2 Data-points

Light Best Practice Case Study is used for this purpose with the following equation,

$$A = (-T_{ei,n+1} + T_{ei,n}) + \frac{(T_{ia} - T_{ei})}{X(1)X(2)} + \frac{(T_a - T_{ei})}{X(3)X(2)} + \frac{(X(4)Q_{sol} + X(5)Q_h + X(6)Q_i)}{X(2)} + \frac{(T_{ei} - T_{ia})}{X(1)} + ((1 - X(4))Q_{sol} + (1 - X(5))Q_h + (1 - X(6))Q_i) \tag{8.2}$$

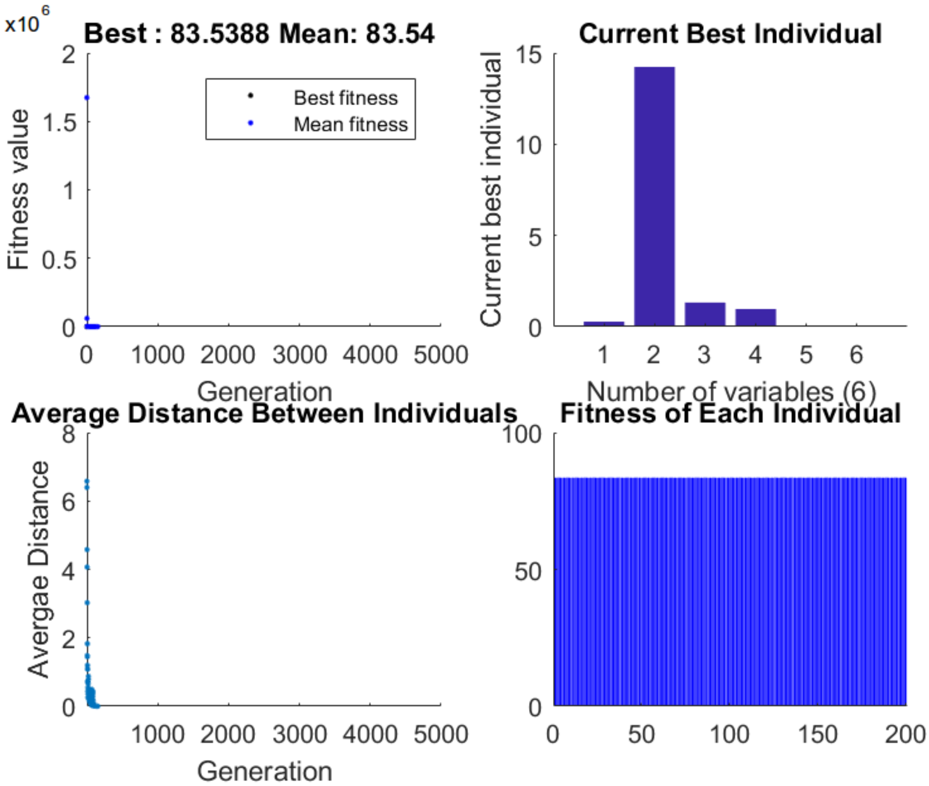


Figure 8.2: GA computed parameters for Single Pitched Roof House - Light Weight Best Practice Insulation Two Node Network Using 15 Days Data

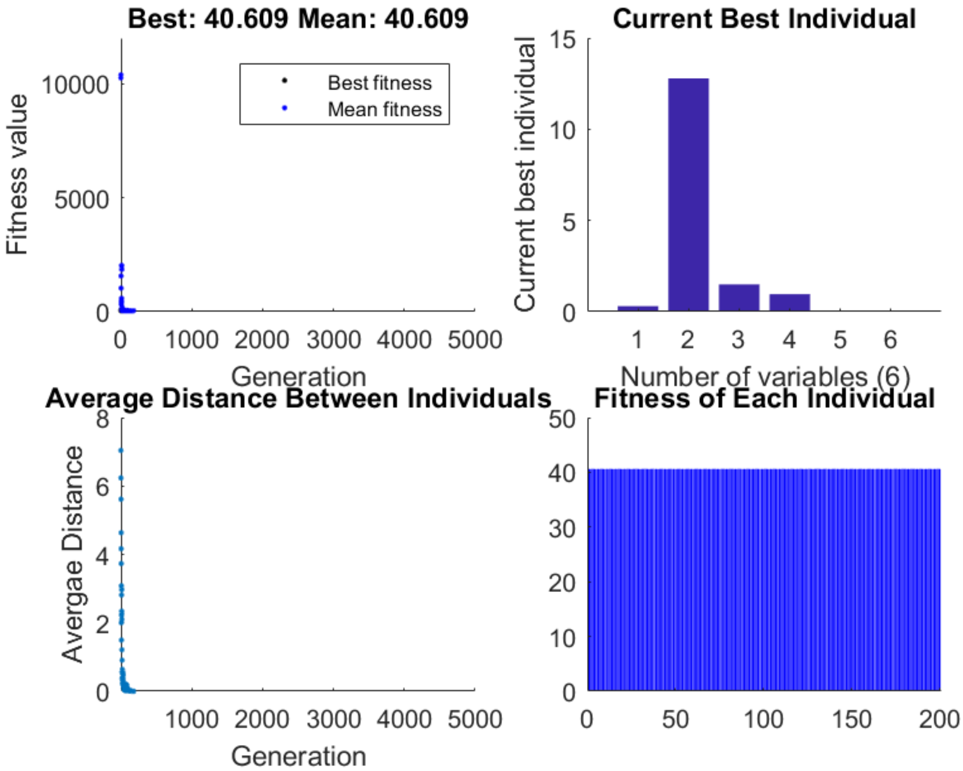


Figure 8.3: GA computed parameters for Single Pitched Roof House - Light Weight Best Practice Insulation Two Node Network Using 3 Months Data

It is to be noted that the cooling is switched off, hence during summers the equation would not hold true as the heating will not be zero but negative unlike the results of EnergyPlus. Hence, the study can maximum be applied to winter months data, the results for which are displayed below,

	$R_{ia,e}$	C_{eqv}	$R_{ea,a}$	f_{sol}	f_h	f_i	Overall Error
Expected	0.32	3.57	1.65	0.99	0	0	–
15 Days	0.279	14.23	1.32	0.96	4.8E-5	1.88E-4	110%
3 Months	0.302	12.78	1.49	0.95	5.3E-5	5.68E-6	90%
Winter	0.225	8.24	2.15	0.99	2.5E-7	4.6E-9	65%

Table 8.2: Variation of Results With Number of Data points

As it can be seen, with more data results are better especially when dealing with high number of unknowns. It would be ideal if winters months data from 1 November to 31st March is available for analysis.

Though, if only the error for resistances is taken into consideration , three months data produces only an error of 6% hence three months data could be enough to determine the resistances of the building.

8.2.1 Mechanical Ventilation and Infiltration

The EnergyPlus File belongs to the category of light construction and best practice insulation. There is inclusion of infiltration of 0.3 air changes per hour as in and an additional mechanical ventilation as per minimum requirements per person. The optimisation equation used for mechanical ventilation and external infiltration will stay the same as two node equation in the previous section except there will be an addition of infiltration resistance according to the equation given below.

$$A = (-T_{ei,n+1} + T_{ei,n}) + \frac{(T_{ia} - T_{ei})}{X(1)X(2)} + \frac{(T_a - T_{ei})}{X(3)X(2)} + \frac{(X(4)Q_{sol} + X(5)Q_h + X(6)Q_i)}{X(2)} + \frac{(T_{ei} - T_{ia})}{X(1)} + \frac{(T_a - T_{ia})}{X(7)} + ((1 - X(4))Q_{sol} + (1 - X(5))Q_h + (1 - X(6))Q_i) \tag{8.3}$$

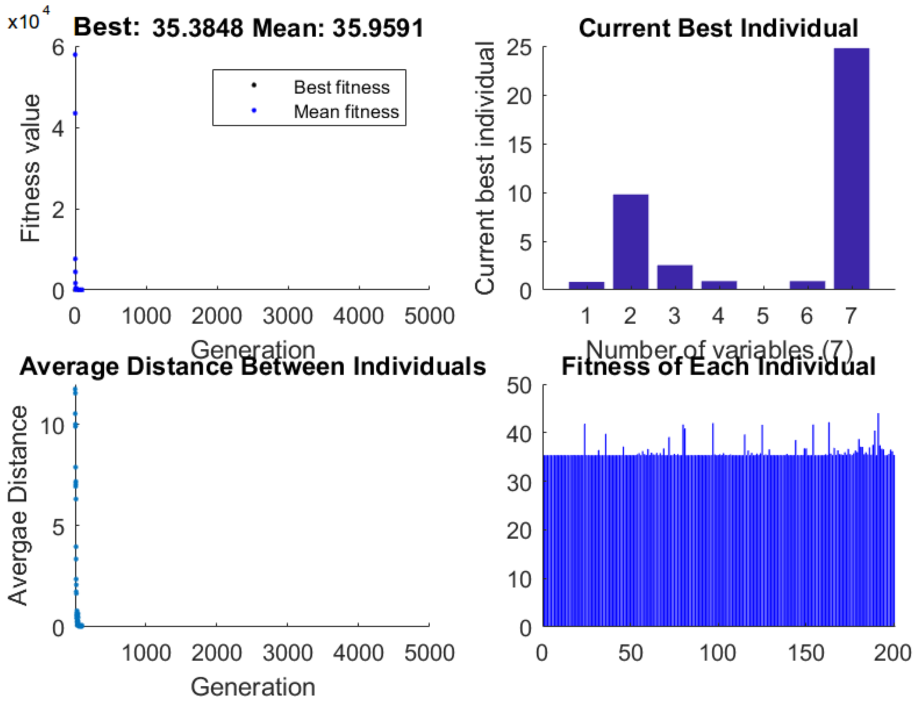


Figure 8.4: GA computed parameters for Single Pitched Roof House - Light Best Practice Two Node Network

$$\begin{aligned} R_{inf} &= \frac{1}{mC_p} \\ &= \frac{3600}{(0.35 * 1.225 * 1.002 * 321)} \\ &= 24.91K/kW \end{aligned} \tag{8.4}$$

GA parameter	Thermal Parameter	GA Value	Theoretical Value	Error
X(1)	$R_{ia,ei}$	0.82K/kW	0.32	156%
X(2)	C_{eqv}	9.7 kWh/K	3.57	220%
X(3)	$R_e + R_{ea,a}$	2.54K/kW	1.65	54%
X(4)	f_{sol}	0.91	0.99	8%
X(5)	f_h	1.6E-3	0	
X(6)	f_i	0.91	0	
X(7)	R_{inf}	24.78	24.91	0.5%

Table 8.3: Analysis of GA computed results for Light Best Practice Infiltration Ventilation Model - Two Node Network

The results for this case are observed to be nearly as accurate as the previous case study. The value of infiltration resistance is determined precisely to the expected value which is because of a large bias present in $(T_a - T_{ia})$ making it imperative for GA to accurately determine infiltration resistance.This can be very useful in application as infiltration is a difficult value to estimate for building engineers. This can provide much insight into the renovation needed for a building to make it energy efficient.

However, addition of infiltration term lead to higher errors in all the other parameters which is an indication that infiltration cannot be modelled simply as a term. There are more process interdependencies introduced because of infiltration which should be included in the thermal networks.

Chapter 9

Effect of Noise

Effect of Noise is studied for a twin heavy un-insulated case study to increase the scope of study to another architecture in Holland. The construction details of construction are same as the single heavy un-insulated house with an additional presence of internal partitions which add to the capacitance of the system.The buildings are modelled with addition of lighting and equipment internal loads. Infiltration of 1 ach and mechanical ventilation is also present for this case study.

9.1 Global Model

The results of GA for a noiseless input data is given in Figure 9.1. Results given by GA are R = 0.45, C = 20.4 as compared to the expected values of R = 0.816 and C = 22.04.

The displayed results are for annual data. They show very low convergence to results as the fitness values are very high.This can be explained by the fact that the equation is modelled for heating whereas for six months data cooling needs to be taken into account as well. Effect of noise is now studied with respect to this case. To analyse the effect of noise, the equation corresponding to a global R and global C values discussed in Section 4.0.1 is used. To assess the performance of GA to noise levels, random noise is added to the EnergyPlus data of Sensible Heat Gain, Indoor Temperature, Ambient Temperature and Solar Radiation. Noise is added to the values by using the random number generator function. The value generated by the function is added to each of the values after adjusting it for the necessary percentage of noise needed. Figures 9.2:9.5 depict the results of GA corresponding to a 5%, 11%, 30%, and 60% noise. Figure 9.6 shows the performance of GA with noise added to all input parameters.

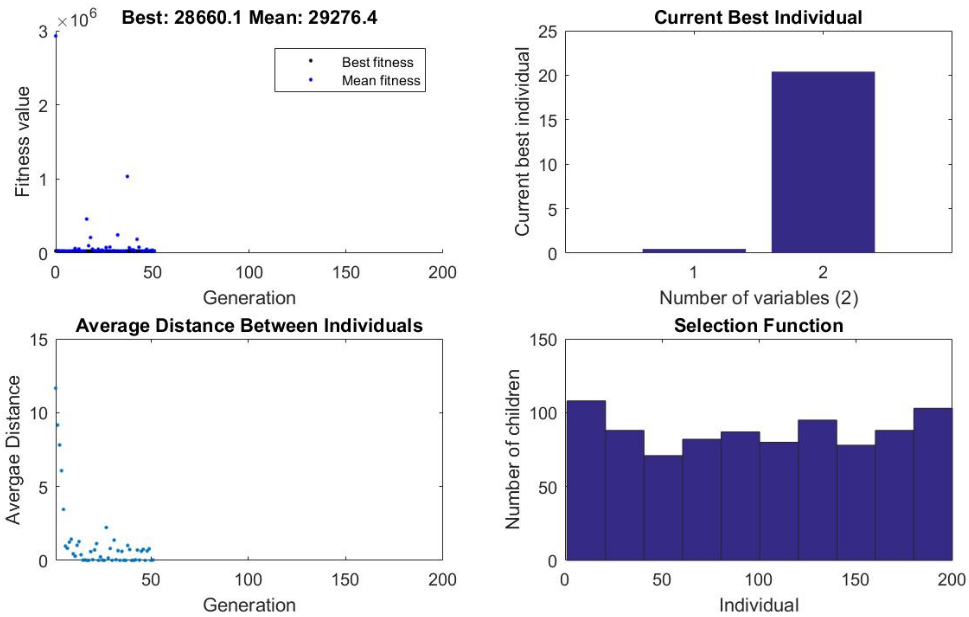


Figure 9.1: GA computed parameters for Twin Heavy Uninsulated Pitched Houses- Global Network

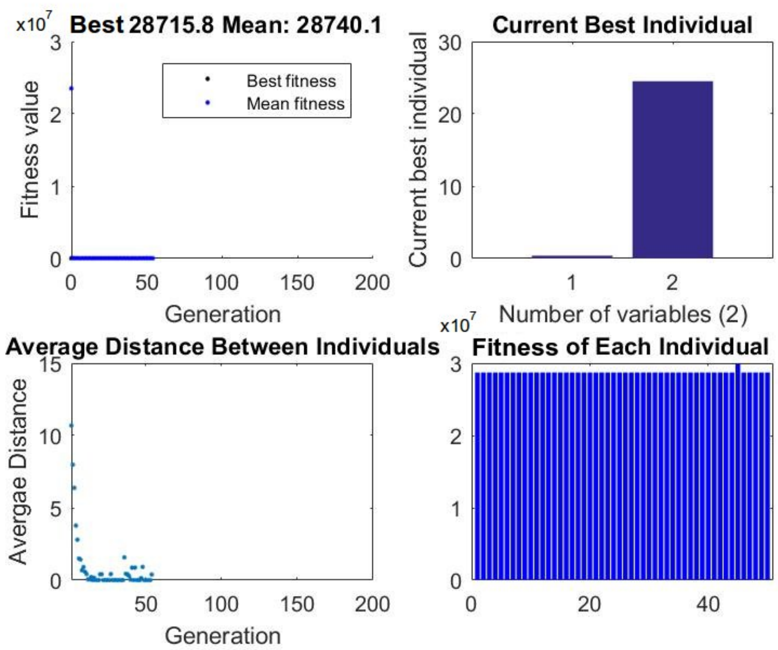


Figure 9.2: GA computed parameters for 5% Noisy Heat Demand

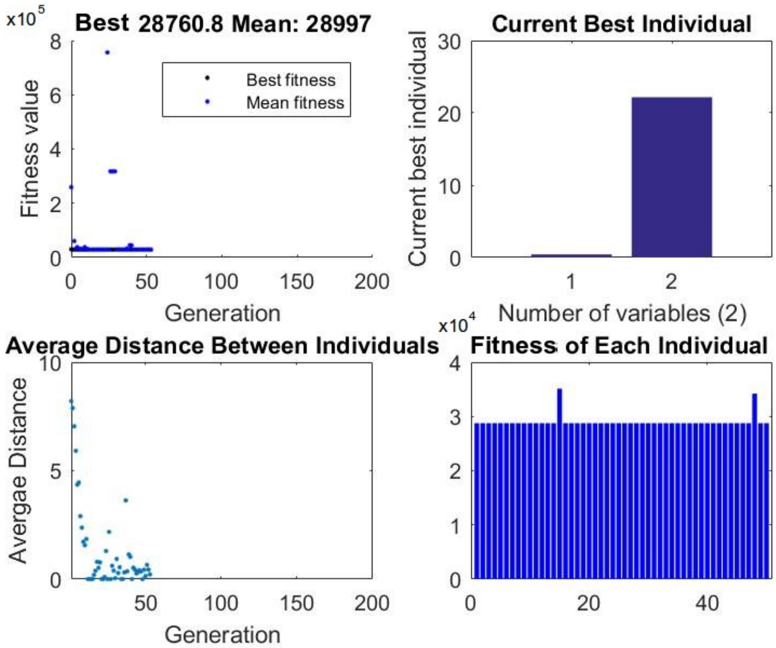


Figure 9.3: GA computed parameters for 11%Noisy Heat Demand

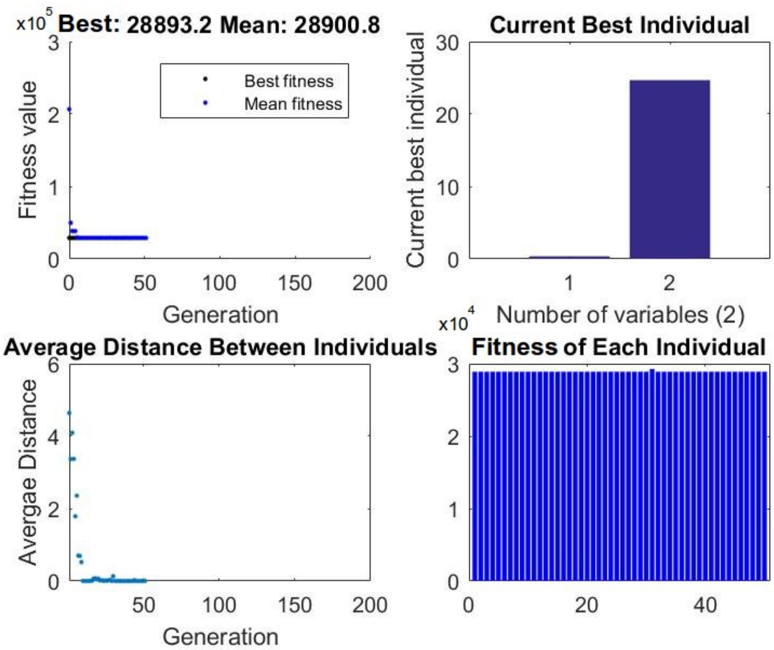


Figure 9.4: GA computed parameters for 30% Noisy Heat Demand

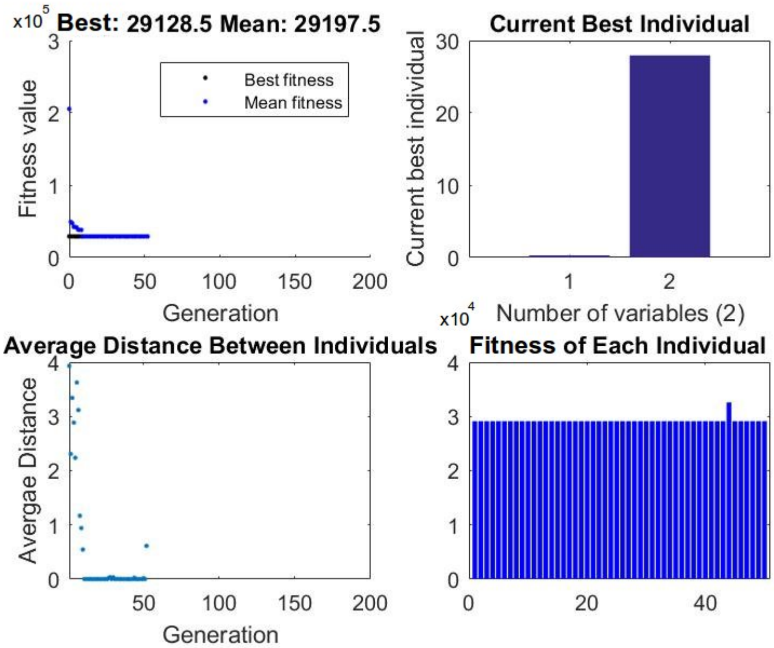


Figure 9.5: GA computed parameters for 60% Noisy Heat Demand

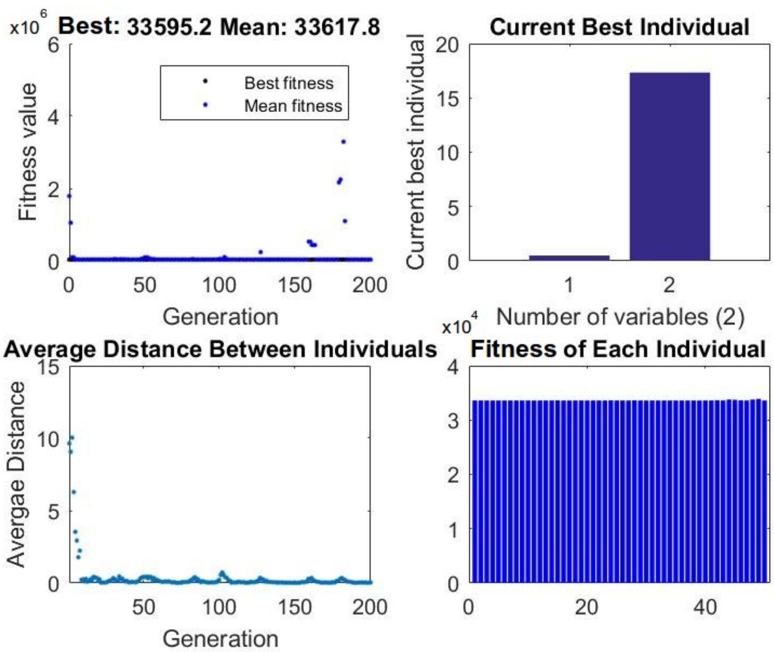


Figure 9.6: GA computed parameters for Noise in All Variables

	No Noise	5%	11%	30%	60%	All Vars 5%
R	0.45	0.391	0.4	0.362	0.307	0.483
C	20.4	24.51	22.12	24.682	27.948	18.09
Overall Error	26%	32%	26%	34%	45%	30%

Table 9.1: Comparison of Results With Variation In Noise

Results for noise when compared to the no noise results show that except for a very high noise level of 60%, GA gives comparable results. Also as this is the global network results, they are not expected to give accurate results hence the error might become less if the values for two node network are compared.

Chapter 10

Conclusion

Thermal networks have been successfully coupled with Genetic Algorithms for determination of Resistances and Capacitance of different types of buildings. Thermal networks of varied complexity were studied to identify an optimal network for modelling the thermal behaviour of the indoor space. The results seen in Chapter 7 suggest that the two node thermal network is optimal for determining the major resistances and equivalent capacitance of the building. Two node network has another advantage over the three node network as it requires less measurements for implementation.Global network was found to be oversimplified model for the building dynamics and the three node model was found to produce errors because of combination of nodes which should be modelled separately.

All three networks' results reported in the thesis have an inherent error which is possibly originating from the fact that for a case with no infiltration and ventilation, the capacity of indoor air can play a role in the dynamics of the house.

GA is found to be a promising technique for parameter estimation for thermal network equations with further analysis. GA's ability to estimate nine parameters for the three network model demonstrates its strength aptly. Although the performance of the three node network was slightly below par, it can be possibly improved by including additional nodes for floor and windows. In addition, GA estimated the given results with limited time and data. The results displayed in the thesis were obtained with a optimisation run time of less than five minutes. Also, as discussed in Section 8.2, considerably good results for resistance were seen with just three months of data which is another advantage of using GA. The ease of implementation is an added asset to GA as a simple addition of individual nodal thermal equations for multi-objective function provided good results.Also, it can be concluded from the discussion on GA results that GA does not produce accurate results in the presence of 'free' parameters so it is preferable that the objective function should have terms where the parameters to be estimated are present in more than one term and are present in conjunction with another unknown.

Different building architectures were also studied to ensure comprehensive application to varied constructions. It is concluded that GA provided better results for heavy built house because of lower differences in temperature of individual surfaces. Hence, it is necessary to input the temperatures of transparent facades separately to achieve good results. As GA is a blind process of finding an optimum within the provided parameter space guided only by the objective function, it is important that the objective function is formulated taking into account all the major possible energy flow-paths.The better results for heavy built are also possible because of higher heat instances in the heavy building providing GA more instances for data fitting. However, for better calibration of the results, thermal penetration depth should be included especially for heavy houses as the ability to store heat transcends the exposed layers.

Towards facilitating application to a real time scenario, where measurement of the internal temperature of surfaces and temperature of the indoor air are necessary which would include noisy biases, the effect of noise is also studied in this work. The results show that GA can manoeuvre the noisy input data and produce similar results as with a noise-free data-set. Also, in the data collected in real houses, presence of infiltration and ventilation cannot be negated hence, another case is studied in the thesis with inclusion of air flow paths. GA produced reasonable results for the convective and conductive resistances and exemplary results for the infiltration resistance which is a huge advantage of using GA. However, the work done in this thesis is based on input data generated by an emulation software and presents only a preliminary study towards the capabilities of GA in application to building parameter estimation. The results obtained for real time data might not be similarly accurate and can be established only upon application.

Chapter 11

Recommendations

The application of Genetic Algorithms for parameter estimation using thermal networks is largely unexplored. This thesis which delineates a preliminary application of GA towards building parameter estimation via thermal network modelling can form the basis for numerous promising avenues of future research. The work performed in this thesis has covered the basic concepts of GA and its working principles. The usage of GA coupled with building simulation data is thoroughly explored with the help of case studies.

Behaviour of every algorithm is expected to vary with different applications. Hence, the optimal formulation of a problem statement according to the limitations of GA is important. Further research into defining the limitations of GA with respect to building data can be useful for future implementation.

A self-coded GA can provide more insight and control on the results as compared to a black-box optimisation tools such as the GA provided by MATLAB. However, the optimisation toolbox of MATLAB gives a fair extent of flexibility in the process. It could be interesting to observe the performance of GA with dynamic mutation and crossover functions. One avenue would be to analyse the effect of varying mutation and crossover fractions as a function of change in fitness functions. This is interesting due to the fact that decreasing the mutation and increasing the crossover fraction with decreasing fitness value change would explore the narrow converged parameter space meticulously at the near end of the algorithm. Under heuristic crossover function, the next generation parameters will not deviate far from the parent parameters making it ideal for exploring around a specific global optimum area. But as mutation can lead the next generation pretty far from its area owing to random chromosome changes, it should be decreased as we approach closer to a minimal objective function value.

Aggregation of states can form a very interesting study for thermal networks. This concept is explained in the Appendix Section A.1.

In case of unbound/loose bounds, GA is more likely to attain the global optima but many times it might converge at a local minima. To avoid this, one option would be to minimise the bounds according to the five fittest solutions attained after a specific number of runs of GA (say $\sim$20). This process could be repeated till GA converges to a certain solution. It could be useful to verify if such a procedure can lead to a good solution in the absence of strict limits.

The aim of this thesis was to assess the scope of implementation for GA for estimating parameters of existing buildings in Holland. The work needs to be followed by real life implementation of GA using sensor data for temperatures of facade, ambient, and thermostat temperatures along with heat demand from smart meters. For the implementation with real time data, GA would need to perform with a noisy data-set. More study into the performance of GA in a noisy environment is mandatory. In addition, the data preparation step would play an important role and could be explored.

One of the defining factors for the success of a novel technique is how it compares to the results of an established technique for the same application. Hence, comparison of the results given by GA with linear regression techniques would be an important step towards establishing the real scope of GA for building parameter estimation.

In a nutshell, the important areas of research to further the study include

- Modelling a ground exposed to the ground temperature and validate the results for more accurate results.
- Verify Aggregating of States.
- Check GA Performance with real time sensor data collected from buildings.
- Compare GA results with Multiple Linear Regression techniques.
- Include Direct Hot Water in Calculations as in smart meters, the heating value includes the heating needed for DHW.
- Check the improvement in GA with a dynamic mutation function.
- Check the improvement in results by employing a multi-objective function formulated by method of substitution.
- Extend the three node network to include independent nodes for windows, doors and floor.
- Evaluate the results for infiltration for real time scenario where it is affected by the wind velocity.

Bibliography

[1] "Key Figures on Europe", 2016 Edition, Eurostat yearbook, Eurostat european Commission, Eurostat Statistical Books.

[2] Olivia Guerra Santin, Laure Itard, Henk Visscher. "The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock." Energy and buildings 41.11 (2009): 1223-1232.

[3] Frits Meijer, Laure Itard, and Minna Sunikka-Blank. "Comparing European residential building stocks: performance, renovation and policy opportunities." Building Research & Information 37.5-6 (2009): 533-551.

[4] Vaidehi Parab. "Thermal Modelling of Existing Residential Buildings in North-Western Europe", Masters Thesis in TU Delft, (2016).

[5] Yuna Zhang, Zheng O'Neill, Bing Dong, and Godfried Augenbroe. "Comparisons of Inverse Modelling Approaches for Predicting Building Energy Performance", Building and Environment 86 (2015): 177-190.

[6] Xiang Li, Tingting Wang, Chakkrid Sattayatewa, and Moschandreas Demetrios. "Odor Emission Rate Estimation of Indoor Industrial Sources Using a Modified Inverse Modelling Method", Journal of the Air & Waste Management Association 61.8 (2011): 872-881.

[7] S. Thamilseran, and Jeff S. Haberl. "A Bin Method for Calculating Energy Conservation Retrofit Savings in Commercial Buildings", Proceedings of H&H - Symposium on Improving Building Systems in Hot and Humid Climates (1994). http://oaktrust.library. Tamu. Edu/handle/1969.1/6640.

[8] Shengwei Wang, and Xinhua Xu. "Simplified Building Model for Transient Thermal Performance Estimation Using GA-Based Parameter Identification", International Journal of Thermal Sciences 45.4 (2006): 419-432.

[9] Daniel Lawson, Glenn Marion. "An Introduction to Mathematical Modelling", Bioinformatics and statistics Scotland (2008): 3-13.

[10] K. Messa, and M. Lybanon. "Curve Fitting Using Genetic Algorithms", Report for the Naval Oceanographic and Atmospheric Research Laboratory, Stennis Space Center, (1991).

[11] Retrieved on 17th October,2017 from "http://nesa1.uni-siegen.de/wwwextern/idea/keytopic/3.htm".

[12] Retrieved on 20th October,2017 from "https://fet.uwe.ac.uk/conweb/hi4web/Principles/section1.htm"

[13] Goldberg. D.E ."Genetic Algorithms in Search, Optimisation, and Machine Learning". Chapter 1, 1st edition, Addison-Wesley Professional Publishers, (1989).

[14] Harb, Hassan, et al. "Development and validation of grey-box models for forecasting the thermal response of occupied buildings." Energy and Buildings 117 (2016): 199-207.

[15] Bacher, Peder, and Henrik Madsen. "Identifying suitable models for the heat dynamics of buildings." Energy and Buildings 43.7 (2011): 1511-1522.

[16] Retrieved on 8th October, 2017 from https://www.designbuilder.co.uk/

[17] Clarke. J.A ."Energy Simulation in Building Design".Chapter 1, 2nd edition, Butterworth-Heinemann Publishers. (2001).

[18] Laure Itard. "Heat transfer and room balances- Heating and Cooling Capacities/ energy demand", Lecture in Indoor Climate Fundamentals, Technical University Delft, Netherlands. 18 Sept. 2017.

[19] Kun Deng, Prabir Barooah, Prashant G. Mehta, and Sean P. Meyn. "Building Thermal Model Reduction via Aggregation of States", In American Control Conference (ACC), (2010): 5118-5123.

[20] Retrieved on 30th August, 2017 from *http://lpsa.swarthmore.Edu/Systems/Thermal/SysThermalModel.html*

[21] Jin Yang, Hugues Rivard, and Radu Zmeureanu. "Building Energy Prediction With Adaptive Artificial Neural Networks", Ninth International IBPSA Conference, Montral, Canada (2005): 15-18.

[22] Zhun Yu, Benjamin C.M. Fung, and Fariborz Haghighat. "Extracting Knowledge from Building-Related Data : A Data Mining Framework." Building Simulation, 6.2 (2013): 207-222.

[23] Imran Khan, Alfonso Capozzoli, Stefano Paolo Corgnati, and Tania Cerquitelli. "Fault Detection Analysis of Building Energy Consumption Using Data Mining Techniques", Energy Procedia 42 (2013): 557-566.

Appendices

Appendix A

A.1 Aggregation of States

The concept of aggregation of states is explored Deng *et al.*[19].It is a means for reducing complex networks consisting of various resistances, capacitance and voltages to simpler low order networks. Their work involves surjection based on Markov chains. Surjection is the mathematical term of mapping elements of a given set X onto another set, Y, such that every element in set Y is equal f(x) for at least one element in set X. They have derived a process for the reduced matrix elements relating them to the original complex network elements which are as follows, The principle could be helpful to study the optimal combination of states and further research could be conducted to explore the accuracy of the method. If found accurate, it could simplify the thermal network analysis. The drawback would be that the solutions would not be directly physically relevant.

step by step process for reduction of models is as follows,

- Construct a Transition Matrix of size n x n, where n is the number of nodes in the thermal network, A
- Calculate the Probability Matrix P, where P = I + AΔt, Δt=1(1 hour)
- Determine the π column matrix, π
- Convert the π column matrix to π Diagonal Matrix, π_d
- Construct a matrix Q with the relation $Q = \frac{1}{2}((\pi_d^{0.5})P(\pi_d^{-0.5}) + (\pi_d^{-0.5})P_t(\pi_d^{0.5})$,$P_t$ - Transpose Matrix of P,
- Compute the eigenvector and eigenvalues of matrix Q
- For the case of bi-partitioning the network, choose the second highest eigenvalue from the eigenvalue matrix

- Examine the corresponding eigenvector and the sign of the the elements in the vector determine the grouping of the network into 2 partitions.
- Compute the equivalent temperature, resistance and capacitance values

A.1.1 Creating Transition Matrix

Transmission matrix is a matrix of elements which form the coefficients of the corresponding nodes in the Thermal Network Equations. For example on comparison of equation determining T_{ia} and the thermal network,

$$C_{ia}dT_{ia} = \frac{1}{R_{ia,e}}(T_e - T_{ia})dt + \frac{1}{R_{ia,a}}(T_a - T_{ia})dt + \phi_{h,ia}dt + f_{conv}\phi_{sol}dt \quad (A.1)$$

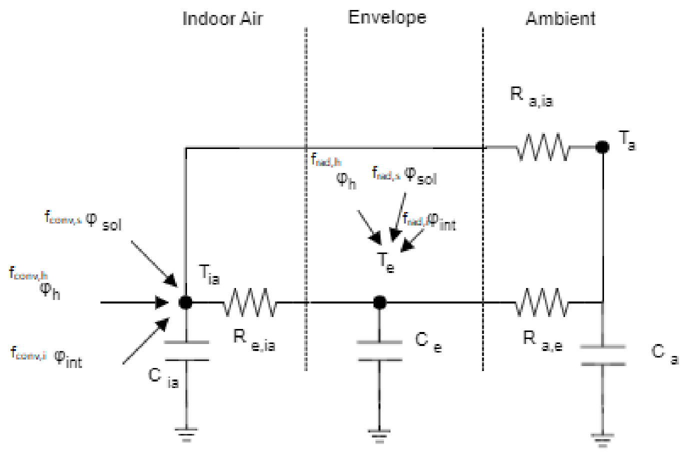


Figure A.1: 3R1C Model

where $node_1$ is Ambient, $node_2$ is Envelope and $node_3$ is Indoor Air.

a corresponding elements of transition matrix A,A_{33} would be the coefficient of T_{ia} in the T_{ia} equation. Hence, it would be $(\frac{1}{R_{ia,e}C_{ia}} + \frac{-1}{R_{ia,a}C_{ia}})$. Similarly

A_{31} would be the coefficient of T_a in T_{ia} equation which would be $(\frac{1}{R_{ia,a}C_{ia}})$ and A_{33}) would be $(\frac{1}{R_{ia,e}C_{ia}})$

A generic formula for the transition matrix elements would hence be,

$$A_{i,j} = \frac{1}{R_{i,j}Ci}, i \neq j \tag{A.2}$$

$$A_{i,j} = -\sum_{j=1}^{n} \frac{1}{R_{i,j}Ci}, \text{i= j, n is the number of edges for the corresponding node} \tag{A.3}$$

For the case of 3R1C model shown in Figure A.1 the transition matrix would be

$$A = \begin{bmatrix} \frac{-(R_{a,e}+R_{ia,a})}{(C_a R_{a,e} R_{ia,a})} & \frac{1}{(C_a R_{a,e})} & \frac{1}{(C_a R_{ia,a})} \\ \frac{1}{(C_e R_{a,e})} & -\frac{(R_{a,e}+R_{e,ia})}{(C_a R_{a,e} R_{e,ia})} & \frac{1}{(C_e R_{e,ia})} \\ \frac{1}{(C_{ia} R_{ia,a})} & \frac{1}{(C_{ia} R_{e,ia})} & -\frac{(R_{ia,a}+R_{e,ia})}{(C_{ia} R_{ia,a} R_{e,ia})} \end{bmatrix}$$

A.1.2 π Matrix

Π matrix is a column matrix of size n x 1 where n is the number of nodes and $\pi_i = \frac{C_i}{(\sum_{i=1}^{n} C_i)}$

$$\pi = \begin{bmatrix} C_1/(C_1+C_2+C_3) & C_2/(C_1+C_2+C_3) & C_3/(C_1+C_2+C_3) \end{bmatrix}$$

A.1.3 π Diagonal Matrix

Conversion of π column matrix to diagonal square matrix is done for ease of mathematical operations on the matrix.

$$\pi_D = \begin{bmatrix} C_1/(C_1+C_2+C_3) & 0 & 0 \\ 0 & C_2/(C_1+C_2+C_3) & 0 \\ 0 & 0 & C_3/(C_1+C_2+C_3) \end{bmatrix}$$

A.1.4 Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors are employed to represent a set of elements in a more concise and localised area. Eigenvectors retain the direction of the original subset and eigenvalues scale down the original values which can be later obtained via linear mapping of the reduced values. To explain, lets take an example of the case we have to employ them in. For the transmission matrix A, create a matrix P = I+At as mentioned above. Eigenvalues,λ would be the mathematical roots of the matrix. As A is a 3x3 matrix it is

derived from 3 equations and has 3 roots where the determinant is equated to zero to form the equations. a vector, v, is an eigenvector if
Ae = λe

A.1.5 Equivalent Reduced Matrix Elements

Once the groups are known, the equivalent circuit has to be constructed. The nodes have been grouped together hence the temperature represented also has changed and so has the resistive and capacitive values on the edges joining the nodes. The equivalent "super-nodes" and the grouped "super-resistor" and "super-capacitor" can be obtained from the original elements using the relations given as:

For each group,

$$C_{eqv} = \sum_{i}^{n} C_i, \text{n is the number of elements in the group} \tag{A.4}$$

$$R_{i,j,eqv} = \frac{1}{A_{i,j} C_i} \tag{A.5}$$

$$T_{i,eqv} = \sum \frac{C_i}{\sum C_i} T_i \tag{A.6}$$

$$A_{i,j,eqv} = \sum_{i}^{n} \sum_{j}^{m} \frac{\pi_i A_{i,j}}{\sum\limits_{i}^{n} \pi_i}, \text{n is the number of elements in the group} \tag{A.7}$$

To determine how to reduce the 3R1C model to global R and global C model, lets repeat the process while assuming the ratio of resistances and capacitance from the work of Vaidehi [4]as she also studied an old Dutch House.

$R_{a,e} = 2.734$
$R_{ia,a} = 7.676$
$R_{e,ia} = 0.246$
$C_a = 12.961$
$C_e = 987.98$
$C_{ia} = 12.961$

$$A = \begin{bmatrix} -2.76 & 0.0305 & 0.01 \\ 0.0003 & -2.76 & 0.0041 \\ 0.0109 & 0.3388 & -7.68 \end{bmatrix}$$

$$P = \begin{bmatrix} -1.76 & 0.0305 & 0.01 \\ 0.0003 & -1.76 & 0.0041 \\ 0.0109 & 0.3388 & -6.68 \end{bmatrix}$$

$$\pi = \begin{bmatrix} 0.0119 & 0.97 & 0.0119 \end{bmatrix}$$

$$Q = \begin{bmatrix} -1.76 & 0.0034 & 0.0109 \\ 0.0003 & -1.76 & 0.0373 \\ 0.0109 & 0.0373 & -6.68 \end{bmatrix}$$

The eigenvalue matrix,

$$b = \begin{bmatrix} -6.68 & 0 & 0 \\ 0 & -1.76 & 0 \\ 0 & 0 & -1.75 \end{bmatrix}$$

According to the theory the second highest eigenvalue is to be noted which in this case is -1.76 which is the second column. The second column of eigenvector matrix will determine the grouping.

The eigenvector matrix,

$$a = \begin{bmatrix} 0.002 & -0.72 & -0.69 \\ 0.007 & 0.69 & -0.72 \\ -0.99 & ,0.003 & -0.006 \end{bmatrix}$$

From the second column of eigenvector we can conclude that node 1 and node 2 should be grouped together as they are both positive signed and node 3 should be in another group.

Hence the equivalent circuit for global R and C would be,

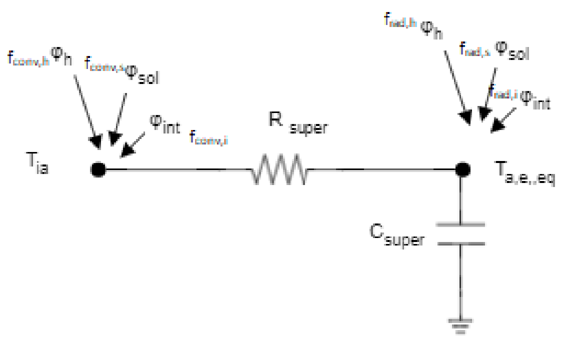


Figure A.2: Aggregated Model for 3R3C

$$\begin{aligned} C_{super} = f(C_a, C_e) &= C_a + C_e \\ &= 1000.94 \end{aligned} \tag{A.8}$$

$$A_{1,2,eqv} = \sum_{1}^{2}\sum_{3} \frac{\pi_i A_{i,j}}{\sum_{1}^{2}\pi_i} = \frac{\pi_1 A_{1,3}}{\pi_1+\pi_2} + \frac{\pi_2 A_{2,3}}{\pi_1+\pi_2} = 0.004 \tag{A.9}$$

$$R_{super} = f(R_a, R_e) = \frac{1}{A_{super}C_{super}} = 0.23 \tag{A.10}$$

$$T_{super} = \sum \frac{C_a}{C_{super}} T_a + \sum \frac{C_e}{C_{super}} T_e = 0.987T_e + 0.012T_a$$

$$A = T_{ia} + \frac{1}{R_g C_g}(T_{super} - T_{ia})dt + \frac{1}{C_g}\phi_h dt \quad + \frac{1}{C_g} Q_{sol} A_w dt + \frac{1}{C_g}(\phi_{int} + \phi_{mvent})dt - T_{ia,n+1} \tag{A.12}$$

A.2 Artificial Neural Networks

Artificial neural networks is a technique for finding the appropriate conduit for the parameters to combine to give the desired output. It attempts to mimic the brain process where information is sent through neurons in the form of a signal. When the process is applied to parameter estimation, the relation between the output and the parameters is given by tailing the strongest signal through the layers. In the first layer, parameters form the neurons and in the final layer, outputs are the neurons. The algorithm runs from the parameters combining in different relations through the intermediate layers consisting of the conditions guiding the decision to the final output. The solution is chosen based on the strength of the signal. Layout of Artificial Neural Network(ANN) is shown in Figure A.3 Yang *et al.* discuss [21]two ANN dynamic models, Accumulative models and Sliding Window approach. Former deals with modifying the large data set with few new entries and retraining the ANN while the latter use the technique of constant data range which is much smaller than the data set used in accumulative training and is also constantly updated with newer data. Yang found the sliding window approach to be overall better option as it holds up for both real and synthetic data. Though sliding window approach lacks the ability to predict the seasonal or annual variations as the data set used for training is small. Yang also applies Principle Component Analysis to improve the results by choosing only the parameters which are affecting the result more than 1%.

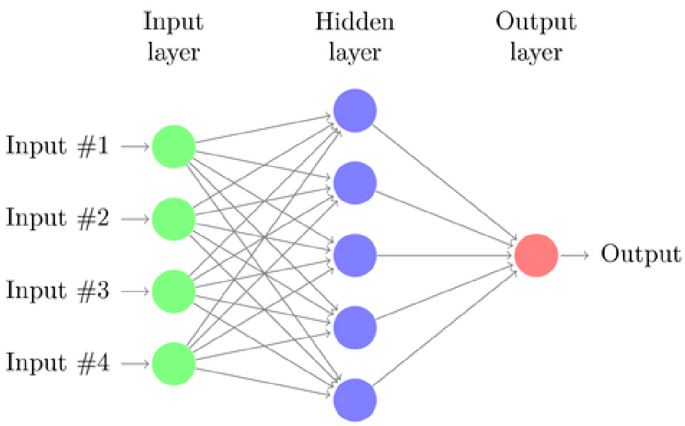


Figure A.3: Layout of a Neural Network

ANN process mostly can be tweaked by changing the number of neurons, number of layers, number of iterations and readjusting the weightage of the neurons. As the number of neurons increase, the complexity of the network increases thus giving good results for data which is difficult to relate with a simple network. The results of the fitting problem is measured in terms of Mean Square Error(MSE) and R(Regression Value). They both represent how close the target are to the outputs. The closer the R value is and the lower the MSE value is, the better is the result. The inputs given to the MATLAB code are normalised so as they are all represented between 0 and 1. The Neural NetworkToolbox is used to fit the data to the desired targets by giving maximum amount of inputs consisting of $Q_{sol}, T_a, T_{in}, R_{ia}, e, R_e, a, R_{in}, ia, C_e, C_{in}, R_{inf}$ to converge to an output of heater demand.

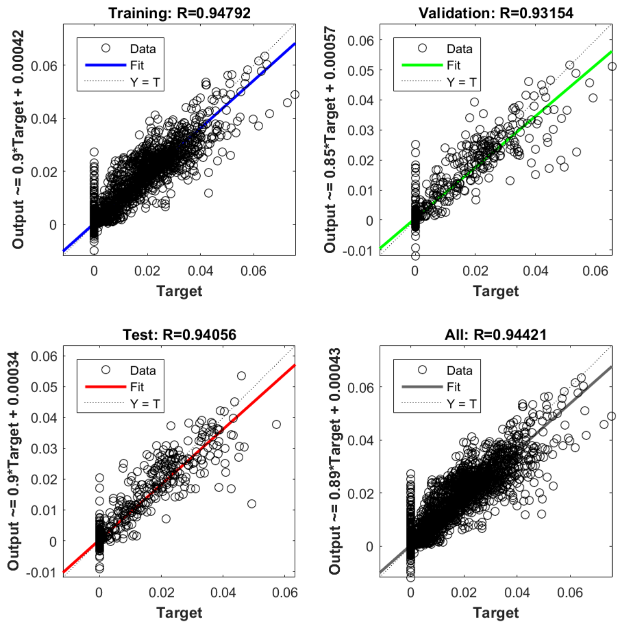


Figure A.4: Resulting Regression Plot of the Trained Artificial Neural Network

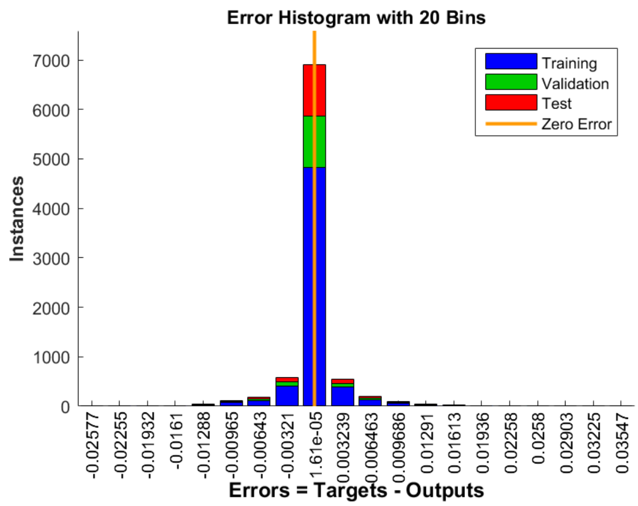


Figure A.5: Resulting Error Histogram of the Trained Artificial Neural Network

From the results shown in Figure A.5 and Figure A.4, it is certain that

ANN can predict the heater demand well given enough number for input parameters and data points. It can be seen that the maximum error on the values are 0.5% which is reasonably accurate. But the inverse process of determining the Rc value and capacitance of wall is not possible via ANN as for this case the outputs do not change with inputs and hence the algorithm has nothing to learn about how the output changes with input parameters. So there needs to be another approach to inverse modelling where outputs are constant.

A.3 GA Performance Results-Varying Parameters

In this section, the performance of GA for varying parameters is analysed.

A.3.1 Results for Line With a Constant

The results for GA estimation for a line with a constant is shown in table A.1

X	Y	Z	W	Objective Function Value
Bounds Adjusted to [0,0,0,0], [10,10,10,10]				
3.332	4.274	8.789	6.201	5.244E-06
3.753	4.257	7.982	7.867	2.4E-09
2.48	7.932	3.176	6.451	2.15E-05
3.398	7.332	2.54	9.523	1.44E-08
2.993	4.207	9.601	4.778	9.57E-06
Bounds Adjusted to [2,4,3,3], [4,6,8,8]				
3.357	5.042	7.202	7.071	2.60E-10
3.153	5.089	7.517	6.3	5.40E-09
3.412	5.641	5.895	7.887	1.44E-10
2.66	5.552	7.577	4.791	5.90E-10
Bounds Adjusted to [3.5,5,5,4], [4.5,6.5,8,8]				
3.089	5.39	7.041	6.547	4.30E-11
3.221	5.725	6.108	7.209	9.160E-11
3.367	5.165	6.937	7.231	7.6E-12
Bounds Adjusted to [3.5,5,5,4], [4.5,6.5,8,8]				
3.089	5.39	7.041	6.547	4.30E-11
3.221	5.725	6.108	7.209	9.160E-11
3.367	5.165	6.937	7.231	7.6E-12

Table A.1: Results for Linear Line with a Constant

A.3.2 Effect of Population on GA Performance

This section analyses the effect of population size on the performance of GA.

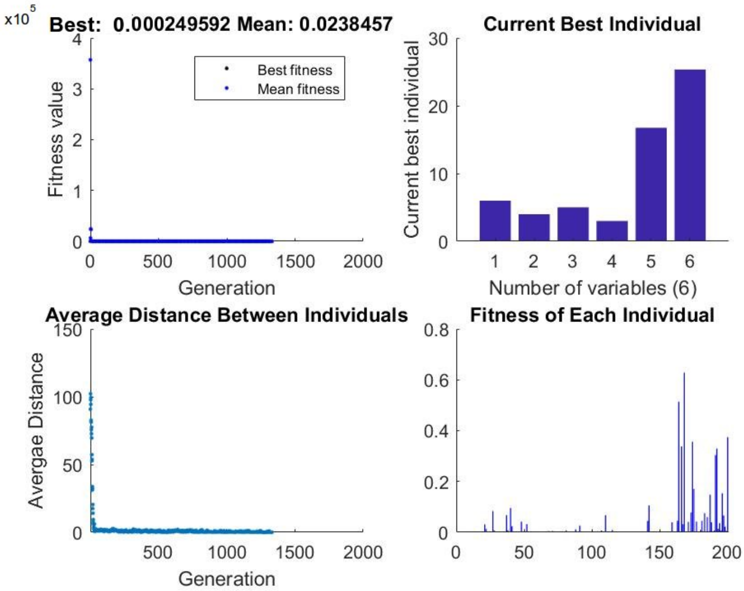


Figure A.6: GA computed parameters for Thermal Network Equation for Population Size of 200

A.3.3 Effect of 2500 Data Points on GA Performance

Results for 2500 Data points are shown in figure A.8

A.3.4 Pre-Converged Results for Thermal Network Equation with 6 Variables

Results for 6 parameter thermal network equation is shown in figure A.9.

Results estimated by GA (shown in the graph) can be seen to vary in accuracy.

u	v	w	x	y	z
6.043	3.955	5.118	3.014	49.577	26.711

Table A.2: Values of GA computed Premature parameters for Thermal Network Equation with 2000 Generations

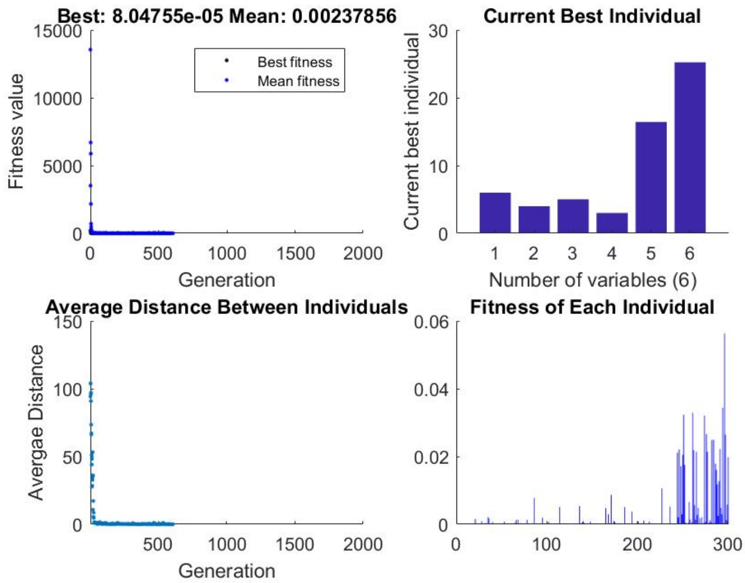


Figure A.7: GA computed parameters for Thermal Network Equation for Population Size of 300

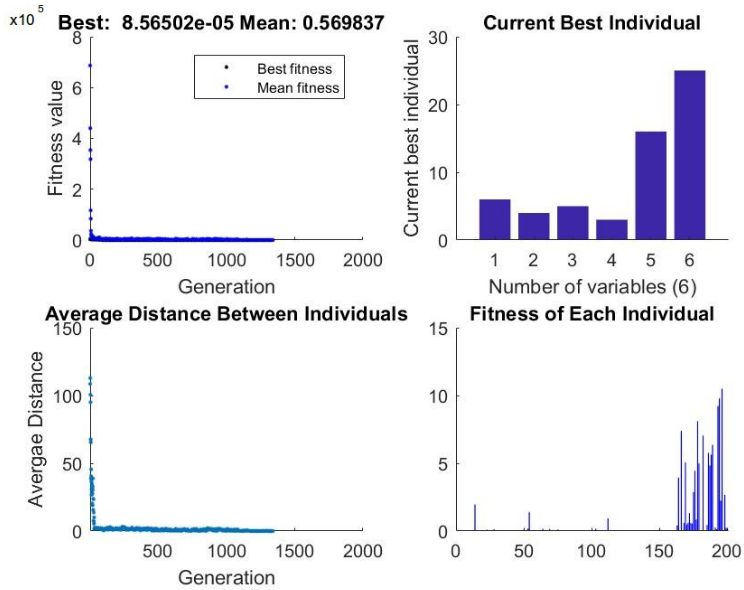


Figure A.8: GA computed parameters for Thermal Network Equation with 2500 Data Points

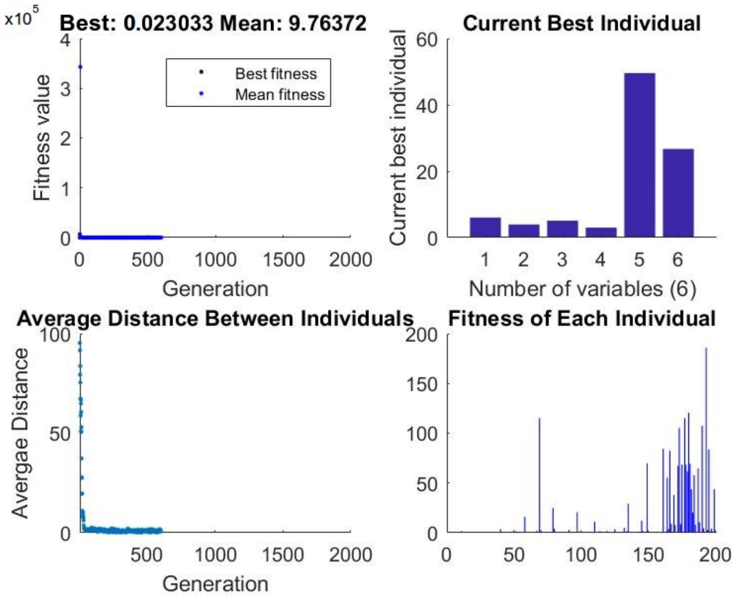


Figure A.9: GA computed parameters for Thermal Network Equation with 6 Unknowns

A.3.5 Case Study : Heavy Built Uninsulated Twin Houses

Figure A.10: Composition of External Walls

Pitched Roof - Uninsulated Lightweight is made of 3 layers, Clay Tile(Roofing) and Roofing Felt(Innermost)

Parameters	Specification
Occupation	0.01 people/m^2
Wall	Heavy Weight Uninsulated Wall [$Rc = 0.663m^2K/W$]
Roof	Light Weight Uninsulated Roof [$Rc = 0.341m^2K/W$]
Internal Walls	115mm Single Leaf Brick [$Rc = 0.510m^2K/W$]
Internal Floor	300mm Concrete Block [$Rc = 0.484m^2K/W$]
External Floor	Heavy Weight Uninsulated[$Rc = 0.506m^2K/W$]
Infiltration	0.3 ach(Infiltration Case)
Glazing Type	Double Glazing 3mm with 3mm spacing[$Rc = 0.316m^2K/W$]
Heating	Boiler with Radiator(Central Heating- Air)
Ventilation	Mechanical Ventilation with no Heat Recovery(Ventilation Case)
Schedule	8am to 6pm(Monday to Saturday)

Table A.3: Experimental Parameters of Heavy Uninsulated House

Figure A.11: Composition of Pitched Roof

Internal Partitions - 115mm single leaf brick(plastered both sides) is made of 3 layers, Gypsum Plastering(Outermost and Innermost)

Figure A.12: Composition of Internal Partitions

External Floor - Uninsulated Heavyweight is made of 3 layers, External Rendering(Outermost) and Timber Flooring(Innermost). The floor is built to be exposed to ambient conditions.

Figure A.13: Composition of External Floor

Internal Floor - 300mm concrete slab is made of a single layer of Cast Concrete.

A.4 Working of DesignBuilder

Model Options can be found in the Edit Menu of Design Builder. The "Model options" menu is the start to any simulation file. It defines the structure for modelling the working dynamics of the buildings. It has options to select Construction and Glazing Data to Pre-Design or General. In the former option, the user can choose the levels of insulation and thermal mass for the building while the latter loads a generic default template is loaded for the file. Followed by construction Gains Data can be chosen to be modelled

in Lumped, Early or Detailed models. In Lumped, all the internal gains are represented via one value representing the aggregated sums of occupancy, equipment, lighting gains etc. In Early or Detailed, the user can model the gains separately according to the specifics needed. Next tab is Timing which gives the user a choice to select their own schedules or a typical workday schedule allotted to all the modelling components. In the tab for HVAC, Simple HVAC can be chosen if the modelling of ideal load based system is adequate otherwise the user can opt for detailed. Natural Ventilation and Infiltration can be based on Scheduled or Calculations. For a simple implementation of fixing a given ventilation rate, scheduled is used. This rate changes according to the operation of the residence but is unaffected by pressure coefficients unlike the Calculations model. There are a few more parameters to be modelled like the simulation tab which can be used to specify the simulation period, control temperature, time steps per hour and Solar Distribution.
For our case, control temperature is set to Air Temperature to model the simulation closer to real life application. To determine the real life values of R and C, the indoor temperature data would be received via a thermostat which will record the temperature of the air surrounding it. Another option is to chose Operative Temperature which is the temperature felt by the occupant, it would lead to more accurate heating demand but would not be easy to determine in a house. Solar Distribution is set to Full Exterior which means that all the transmitted solar gains are assumed to fall on the floor.

Activity template has further tabs starting from 'All Gains" which is the window which can be used to specify the aggregated sum of internal gains from people, lighting and equipment, if the gain options which discussed before in Model Options is set to "Lumped".

If gain options is set to "Early" to "Detailed", the occupancy internal gains can be specified in the subsequent tab of "Occupancy". Schedules for occupancy can be set according to the schedules template.

The next tab "Other gains" programs the equipments and lighting gain. For the equipment the load and radiant fraction can be programmed. Radiant fraction is the fraction of the load expelled by radiation to the surrounding surfaces. Convective Fraction is $(1 - f_{rad})$ which is the heat gained by the indoor air via equipment. General Lighting can be switched on in the activity template but can only be programmed in the main screen.

Environmental Control is the last tab in the activity templates which can be used to setup the set point temperatures for HVAC Systems.

Construction template can be set if in the model options, the construction settings are set to pre defined. The construction template chosen for the simulation of heavy built uninsulated building is Uninsulated and Heavy Thermal Mass. The pre-settings in the template is as follows,

External - Wall -Uninsulated- Heavyweight which is made of 4 layers, Brickwork Outer(Outermost) to Gypsum Plastering(Innermost). The tab of "Surface Properties" can be used to fix the convective heat transfer coefficients which can override the surface convection algorithms selected before.The next tab "Calculated" is an important tab for our study as this window displays the calculated heat transfer coefficients used in energy simulations. The values displayed here are used only if SBEM convection algorithm is employed. For all other algorithms, the heat transfer is given separately for each surface by storing output of the surfaces which will be discussed in the outputs sections of EnergyPlus. If CIBSE algorithm is used the convective heat transfer coefficients displayed are used.

Internal Thermal Mass is the component which can be used to add extra internal thermal mass. Thermal mass can represent components like furniture and internal walls which do not take part in the simulation because of the setting for solar distribution in the software, which will be discussed in the "opening" sections of EnergyPlus.

Infiltration can be modelled by the "airtightness" tab . For the construction template being discussed, the cracks template preset is poor with 1ac/h infiltration rate.

Openings tab helps to set a glazing template. **Project Glazing Template has been chosen for the case studies in the thesis, where the glazing is specified with a 30% glazing ratio.Glazing Type of the external windows is specified to double glazing 3mm/6mm Air Cavity**

Lighting Here is where the lighting can be programmed after checking the lighting in the activity template. **template-No Lighting in included in the the work**

HVAC This is one of the main components as the systems results from this tab will provide us with a sensible heating demand which is the most important result for the energy simulation **template-Radiator Heating, Boiling Hot Water, Mech Vent Supply and Extraction with a central heating system. Heating is specified as a natural gas convective heating system type.Cooling energy analysis is not a part of this study. Schedule for HVAC designs was set to 8am to 6pm.**

A.4.0.1 Outputs And Export

Simulation Results include the following categories,

Site Data provides information about the surrounding environmental conditions for the model. The important data sets mentioned here are outdoor dry bulb temperature, Wind Speed, Direct Normal Radiation.

Comfort provides the user, information about the comfort of an occupant in the space which can be assessed by knowing the mean air temperature, the mean radiant temperature, the mean combined radiant and air temperature

Fabric and Ventilation mentions the heat gains through each component of the fabric which give the user detailed heat transfer occurring all over the building. Ventilation provides us with an overview of all exchanges of air between the indoor and outdoor. The notable parameters dealt with in this section are External infiltration-The heat gain from the air entering the zone from outside, Total Air Changes which include mechanical ventilation, natural ventilation and infiltration.

Internal Gains cover all different aspects of how heat could be gained in an indoor space. Hence, it includes lighting, occupancy sensible gains hence it also includes the radiant heat but not the latent, equipment, solar gains through the windows which is the transmitted solar radiation hence a convective fraction would be needed for node T_{ia}, zone sensible heating which includes the heat given by the entire HVAC System. It includes the heat gained through the mechanical ventilation. As this is also sensible heating, this value also includes the radiant part of heater.

A.4.1 Hessian MATRIX Approach

Hessian Matrix is a matrix consisting of double derivatives of the objective function,$H_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$. It is used to optimise a function f(x, y,..)by determining the role of all the critical points of the function. It does so by employing eigenpairs, a scalar value λ and a vector consisting of the same number of rows as the number of optimising variables. The property it is based on is that by multiplication of Hessian matrix by the eigenvector equates the multiplication of the eigenpair. This property makes the calculation of the matrix much easier. Optimisation of Hessian matrix is carried out by classifying the eigenvalues at the critical points of function f at all the point where $\frac{df}{dx_{i,j..}}$ is zero. If the matrix at a given point $(x_i, x_j, ..)$ is positive definite, the respective points form a local minima. If the matrix is negative definite, it is a local maxima.

A.5 Machine Learning

A.5.1 Data Mining

Data mining is the process by which the different patterns are inferred from a given dataset. These patterns can provide an inner glance into the energy consumption patterns of the building by deriving the occupancy changes, high equipment usage, and weekly activities like laundry. Data mining can be used to get information in the form of cluster patterns, outliers or anomaly detection and parameter inter-relations like pattern mining.

A.5.2 Pattern Recognition

Pattern recognition is forming patterns based on parameter inter-dependencies with the aim of learning pre defined rules such that if and when applied to a new set, inferences can be drawn. It can also be used to examine a data set such that the pattern formed can be used to remove the outliers. **Hence, it is very important to run pattern recognition before removing outliers because this would be the best way to find anomalies.**

There are two branches of pattern recognition, Supervised and Unsupervised Learning. Supervised learning involves learning a fitting function from an available data set and using the trends from the dataset to apply to the future data. This might be very helpful in case of faults and alarms or detecting unusual activity. Also this might be more useful when it comes to using the EnergyPlus data and using that for using it as a reference for the real time data involved. Unsupervised learning caters to the demographic where the data is fed to the algorithm with no reference but some conditional parameters. This strategy will be more suited to deduce the schedules of the occupants. One of the most used method in unsupervised learning is cluster analysis. The different cluster algorithms have been discussed by Zhun Yu[22],

- CART(Classification and Regression Tree)- This approach deals with splitting parent nodes into children nodes according to maximum homogeneity. Each parent node in the tree is the classification condition on the target variable which then splits into branches which signify the outcome of the classification condition. The branch further leads to child node which classifies the data on a subsequent sub-criterion.

- Cluster Analysis and K means Algorithm- This algorithm is sensitive to outliers detection(pre cleaning). It clusters data into spherical clusters on the base of distance. The points closest to each other are in one cluster, the rest in another. K means algorithm is a type of clustering where the algorithm divides a given n number of parameters into m clusters by placing each value into a cluster with the nearest mean value.
- DBSCAN-This is a non-spherical cluster forming approach. This approach is less sensitive to outliers and is the base of density clustering. It clusters according to the maximum number of points in a space and forms them into a cluster.

Imran *et al.*[23] compared the three approaches for cluster analysis for detecting faults and irregularities in lighting use and concluded that CART is the best of three.