Finding a Fair Heat Settlement Model

Kristian Kiradjiev\(^1\), Piotr Morawiecki\(^2\) and Robert Pisarczyk\(^1\)

\(^1\) Mathematical Institute, University of Oxford, UK
\(^2\) Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland

(Communicated to MIIR on 19 October 2021)

Study Group: ECGI 144, 17–22 March, 2019, Polish Academy of Sciences, Warsaw, Poland

Communicated by: Kamil Kulesza

Industrial Partner: Fortum

Presenter: Józef Augustynów

Team Members: Kristian Kiradjiev, University of Oxford; Piotr Morawiecki, Polish Academy of Sciences; Robert Pisarczyk, University of Oxford

Industrial Sector: Energy/Utilities

Key Words: heat transfer, heat settlement schemes

MSC2020 Codes: 80

Summary
In this report, we investigate optimal strategies for fair heat settlement schemes in apartment buildings heated by externally powered radiators and pipes. We use mathematical modelling to predict the distribution of heat among apartments taking into account heat flow through the building walls and external conditions. We use both deterministic and statistical approaches.

\(\dagger\) Corresponding Author: Kristian. Kiradjiev@nottingham.ac.uk
## Contents

1. Introduction 4

2. Heat Settlement Model 5

3. Deterministic Mathematical Model 8

4. Spatially Varying Model 11

5. Statistical Approach 14

6. Possible Generalisations and Extensions to the Model 25

7. Discussion and Conclusions 25
1 Introduction

Fortum is a multinational energy company that, among other things, delivers heat to individual blocks of flats in cities. Such buildings are heated using heat substations that are connected to a district heating network. Fortum charges the whole building based on the total heat $Q_n$ delivered to the heat substation. The cost is then split between the apartments using an appropriate heat settlement models. Current models suffer from a variety of drawbacks. A commonly used model of charging apartments based on their size in $m^2$ clearly does not reflect real energy usage and also does not encourage owners to energy efficient behaviours. Another model uses heat meters but those are installed only in the newer apartments for which such meters were put at the time of construction of the whole building. In the older building the cost of installation of such meters is too high or might even be impossible. Therefore, some tenant associations use heat cost allocators which unfortunately can be easily corrupted and their results can be easily manipulated. Moreover, both heat cost allocators and heat meters do not take into account heat fluxes coming from neighbouring apartments. This sometimes results in extreme situations in which people surrounded by hot neighbouring flat reduce or switch off their radiators and benefit from heat coming from neighbouring apartments as well as high temperature of water flowing in the building pipes and installation and they do not pay anything for that. On the other hand, other tenants have to pay for their neighbours.

There are many other challenges in developing fair heat settlement model that reflects real heat usage and also encourages tenants to use energy efficiently:

- it should discourage behaviours such as opening windows for long times and having radiators heating the whole apartments at the same time,
- depending on the objectives of the company it might want to charge different tenants who live in the corners of the buildings and therefore lose energy quicker to the outside world,
- account for home appliances (oven etc) that also produce heat energy,
Finding a Fair Heat Settlement Model

- it should also include cost of heating common areas like staircases.

Therefore, Fortum has decided to install thermometers in each apartment and would like to build heat settlement model that uses this information and also improves upon the previously used ones.

Some modelling on heat transfer in buildings has been done in, for example, [Champneys et al., 2013], [Madsen and Holst, 1995], [Lagree, 2018] and [Sakakibara, 1991].

In Section 2, we present our heat settlement model. In Section 3 and 4, we present two deterministic mathematical models for heat distribution based on the heat equation, and in Section 5, we discuss a statistical approach to the problem. In Section 6, we suggest possible further work, and in Section 7, we draw conclusions.

2 Heat Settlement Model

2.1 Proposed Heat Settlement Model

Here we propose a heat settlement model that uses all heat energy used by individual flats coming from: radiator, heat installation and pipes that carry hot water throughout the building, heat fluxes between apartments and cost of heating common areas.

Let us start with the notation. We denote by \( q_i \) the heat generated by all the heat sources within the \( i \)-th apartment. Also let \( \Delta F_i \) be the net heat flux coming into the \( i \)-th apartment from the neighbouring apartments. We stress that this is the heat flux between the apartments only and does not include heat exchange with the outside world. Thus, taking into account the internal net flux avoids paying double for heat used, and also apartments do not have to pay for heating from the outside when, for example, it is a very hot day. Finally, let \( N \) be the number of flats in the building and \( Q_{com} \) heat used in heating common areas. Then our heat settlement model charges
Figure 1: Schematic picture of a block of flats with a heat exchange between apartments.

each apartment by the amount proportional to:

\[ Q_i = q_i + \Delta F_i + \frac{Q_{\text{com}}}{N}. \]

Therefore, if the whole bill to the building is \( K \) then each apartment pays \( \left( \frac{Q_i}{\sum_{i=1}^{N} Q_i} \right) K \).

We can now think about some particular scenarios to get the intuition behind this settlement model:

- Consider an apartment that is in the middle, has fewer exterior walls that completely switches off its radiators while its neighbours have them on. In the settlement model based on heat meters such apartment would not have to pay anything despite the fact that it might have warm apartment from its neighbours. In our model, it would still have to pay because the flux \( \Delta F_i \) would be positive. Moreover, our heating model that finds \( q_i \) includes all heat sources in the apartment i.e. installation, hot water in pipes etc. Also, such neighbouring apartments would not have to pay for their neighbour as for them the flux \( \Delta F_j \) will be negative and they will see the reduction in their cost.
• Consider an extravagant tenant that uses their heating and keeps their windows open with cold temperature outside. In order to sustain a comfortable indoor temperature they would have to run their radiator at a high setting, which would result in a large $q_i$ and in the end a high cost (as it should be). This is precisely the reason why we are not including heat flux with the outside word in $\Delta F_i$. Otherwise, this person would pay less despite wasting energy.

• Suppose it is a very hot day, and the temperature outside is higher than the desirable temperature inside. Then, if, for example, all apartments have the same temperature, which is lower than the outside one, and they have all switched off their radiators, according to our model, they don’t have to pay anything, since the sources and internal fluxes are zero. This is consistent with reality, as the apartments are using the solar power from outside.

2.2 Advantages and Disadvantages of the Heat Settlement Model

Our model improves upon the used settlement model for the following reasons:

• it includes real energy used by the apartment which is an improvement over a model based solely on the area of the apartment,

• it includes heat coming from installation and pipes, so that the flat that benefits from this heat which in the end comes from the heat substation still has to pay their share of the total cost of this heat.

• it accounts for the heat flux between the neighbouring flats

• it encourages good energy saving behaviours.

However, it still has some drawbacks.

• It neglects home appliances and other heat sources in the flat. A person that uses electric heater or has some other electric equipment on, which contributes to warming the flat, has to pay for this as if it was coming from a radiator.
• It does not discriminate between heat exiting the room because of bad insulation or wasteful practices like opening windows. Therefore a person who is in the corner of the building has to use more energy and therefore pay more. However, this is a matter to be resolved by the tenant association, whether such people should pay less.

Now we need to have a model that, given the temperatures and the geometry of the building, calculates \( q_i, \Delta F_i, Q_{\text{com}} \). In the following sections we present three such models. The first one is based on the heat equation and in addition to temperature variations across the building takes as input the geometry of the building, and physical properties of the walls (their width, heat conductivity and heat capacity). It assumes constant temperature within each apartment though. The second one assumes a steady state but includes spatial variance within apartments and where the radiators are located. The third model only uses temperature readings and uses linear regression to estimate the heat used.

## 3 Deterministic Mathematical Model

Below we present a simplified one-dimensional model for heat transfer in an apartment building that will help us estimate the various quantities introduced at the end of the previous section. Since this is a 1-D model, we can think of the building as a corridor with \( N \) apartments connected in series, or a vertical block with apartments stacked on top of each other.

### 3.1 Model Formulation

We denote by \( T(x, t) \) the temperature in the apartments and by \( q(x, t) \) the heat source (radiators, pipes) within each apartment. Assuming heat transfer due to conduction between the apartments, the temperature \( T \) satisfies the heat equation with a source term

\[
\rho c_p \frac{\partial T}{\partial t} = \nabla \cdot (k \nabla T) + q,
\]  

(1)
Finding a Fair Heat Settlement Model

where $\rho$, $c_p$ and $k$ are the effective density, specific heat capacity and thermal conductivity of an apartment. To simplify the situation, we assume that the temperature in each apartment is uniform across it, and the transfer happens through the walls. We also assume that the sensors are equally spaced throughout the whole building.

We now discretise this heat equation in space and in time, treating each apartment as a grid point, for which we know the temperature value from the sensors’ readings. In particular, we have

$$\rho c_p \frac{T_{i}^{n+1} - T_{i}^{n}}{\Delta t} = k \frac{T_{i+1}^{n} - 2T_{i}^{n} + T_{i-1}^{n}}{\Delta x^2} + q_{i}^{n}, \quad (2)$$

where $T_{i}^{n}$ denotes the temperature in apartment $i$ and at time step $n$, $\Delta t$ is the time step (i.e., the time span between two consecutive readings of a sensor), and $\Delta x$ is the distance between two sensors in neighbouring apartments. Together with this equation, we have boundary conditions on the outside walls, which incorporate the effect of the outside weather conditions. For simplicity, we assume that the temperature on the outside walls is a given function of time, $T_{\text{out}}(t)$, which can, for example, reflect the diurnal oscillations in the temperature. In the case of the floor of the building, we can prescribe no-flux, treating the Earth as an insulator. We implement these conditions by placing ghost points (apartments) in our spatial discretisation at the boundary and using the same procedure. In particular, in the case of a multi-storey building, this means $T_0 = T_1$ (no flux at the bottom) and $T_{N+1} = T_{\text{out}}(t)$ (outside temperature at the top).

Rearranging, we can find the unknown heat source $q_{i}^{n}$. Thus, this turns into an inverse problem of finding the heat usage of each apartment given their temperature readings. We, thus, obtain

$$q_{i}^{n} = \rho c_p \frac{T_{i}^{n+1} - T_{i}^{n}}{\Delta t} - k \frac{T_{i+1}^{n} - 2T_{i}^{n} + T_{i-1}^{n}}{\Delta x^2}. \quad (3)$$

The flux term $\Delta F_{i}$ is easily obtained, using Fourier’s law

$$\Delta F_{i} = -k \frac{T_{i}^{n} - T_{i+1}^{n}}{\Delta x} - k \frac{T_{i}^{n} - T_{i-1}^{n}}{\Delta x}. \quad (4)$$

Finally, the heat from the common areas can be calculated by treating the common area as an apartment with its own temperature reading.
Figure 2: Temperature variations for 10 apartments and the outside temperature.

We note that this model can be straightforwardly generalised to 2-D and 3-D geometries, which are much more realistic.

3.2 Numerical Results

We simulate the behaviour of the above model on an artificial situation, where we have 10 apartments in series and a given dependence of the outside temperature on time for a single day. We use a set of manufactured temperature readings. We plot these in figure 2. We then use sample values for the material properties of concrete in internal and external walls. Using our model, we obtain the total heat used by each apartment for a day, assuming no common areas. The results are plotted in figure 3. As we see, apartment 5, which keeps the highest temperature, has to pay the most, but there are other apartments, which have to also pay a similar amount, even though they don’t keep such a high temperature, and this is because they are using the heat of their neighbours.
4 Spatially Varying Model

We now present a continuum model that takes into account spatial variations of temperature within each apartment due to the location of the radiators but is in steady state. For simplicity and ease of presentation, we consider a one-dimensional model, but it can be generalised to two- and three-dimensional geometries. We assume that there is one radiator in each apartment situated on the floor in the one-dimensional case, but again this can be generalised to include more radiators in more realistic geometries. In this model, we treat the radiators as point sources of a given strength, $Q_i$, where $i$ refers to the $i$-th apartment. We model each apartment in the one-dimensional case as an air gap of thickness $h$ and thermal conductivity $k$, situated between solid floor and ceiling of thickness $d$ and thermal conductivity $\tilde{k}$ (see figure 4). We denote the temperature in the gas phase by $T_i$, and in the solid phase by $\tilde{T}_i$. Then, these temperatures satisfy the following heat equations

$$k \frac{d^2 T_i}{dx^2} = Q_i \delta(x - ih), \quad \text{for} \quad ih < x < (i + 1)h - d \quad (5)$$

$$\frac{d^2 \tilde{T}_i}{dx^2} = 0, \quad \text{for} \quad (i + 1)h - d < x < (i + 1)h, \quad (6)$$

Figure 3: Total heat energy used by each apartment. Sample values: $\rho = 2400 kg/m^3$, $c_p = 880 J/kg K$, $k_i = 0.4 W/m K, k_e = 0.05 W/m K$. 
subject to

\[ T_i = \tilde{T}_i \quad \text{at } x = (i + 1)h - d, \]  
\[ k \frac{dT_i}{dx} = \tilde{k} \frac{dT_i}{dx} \quad \text{at } x = (i + 1)h - d, \]  
\[ T_i = \tilde{T}_{i-1} \quad \text{at } x = ih, \]  
\[ k \frac{dT_i}{dx} - \tilde{k} \frac{dT_{i-1}}{dx} = Q_i \quad \text{at } x = ih, \]

where the first and third conditions correspond to continuity of temperature across the solid and gas phases, the second condition represents continuity of heat flux across the ceiling, and the fourth condition describes the jump in the heat flux across the floor due to the radiator being there. The general solutions are

\[ T_i = a_i x + b_i, \]  
\[ \tilde{T}_i = \tilde{a}_i x + \tilde{b}_i. \]

The boundary conditions give

\[ a_i((i + 1)h - d) + b_i = \tilde{a}_i((i + 1)h - d) + \tilde{b}_i, \]  
\[ ka_i = \tilde{k} \tilde{a}_i, \]  
\[ a_i(ih) + b_i = \tilde{a}_{i-1}(ih) + \tilde{b}_{i-1}, \]  
\[ ka_i - \tilde{k} \tilde{a}_{i-1} = Q_i. \]

Given \( Q_i \), this can be used to express \( a_i \) and \( b_i \) in terms of \( \tilde{a}_i \) and \( \tilde{b}_i \), and then get two recurrence relations for \( \tilde{a}_i \) and \( \tilde{b}_i \) in terms of \( \tilde{a}_{i-1} \) and \( \tilde{b}_{i-1} \). In practice, though, we do not know \( Q_i \) and wish to find it. Once we find it, then we can feed that in our heat
settlement model and calculate how much each apartment needs to be charged. What we do know are the temperature readings of sensors located in each apartment. Now, suppose the temperature sensors are located at the ceiling, so define

\[ T^*_i = \tilde{a}_i((i + 1)h - d) + \tilde{b}_i, \] (17)

which we, thus, assume is known. Then, we are able to obtain recurrence relations for \( a_i, b_i \) and \( Q_i \), for example. In particular, we have

\[ b_i = T^*_i - ((i + 1)h - d)a_i, \] (18)

\[ \tilde{a}_i = \frac{ka_i}{\bar{k}}, \] (19)

\[ \tilde{b}_i = T^*_i - \frac{k((i + 1)h - d)a_i}{\bar{k}}, \] (20)

\[ Q_i = k(a_i - a_{i-1}). \] (21)

Finally, we have a recurrence relation for \( a_i \)

\[ a_i = \lambda a_{i-1} + c_i, \] (22)

where \( \lambda = kd/\bar{k}(d - h) \), \( c_i = (T^*_i - T^*_{i-1})/(h - d) \).

Possible boundary conditions on an \( N \)-storey building are no-flux at the bottom, i.e., insulated at the bottom, and prescribed outside temperature, \( T_{out} \), at the top.

\[ \frac{d\tilde{T}_{-1}}{dx} = 0 \quad \text{at} \quad x = -d, \] (23)
\[ \tilde{T}_N = T_{out} \quad \text{at} \quad x = (N + 1)h. \] (24)

This gives

\[ a_{-1} = 0, \] (25)
\[ a_N = \frac{(T_{out} - T^*_N)\bar{k}}{dk}. \] (26)

Thus, \( a_n \) can be determined.

We note that the total heat supplied to the building is

\[ Q = \sum_{i=1}^{N} Q_i = \sum_{i=1}^{N} k(a_i - a_{i-1}) = k(a_N - a_0). \] (27)
We calculate \( a_i \) for the particular case when \( N = 1 \), i.e., we have a ground floor and a first floor. We have
\[
\begin{align*}
a_1 &= (T_{\text{out}} - T_{1}^{\ast})\frac{k}{d} \\
a_0 &= \frac{(a_1 - c_1)}{\lambda} = \frac{((T_{\text{out}} - T_{1}^{\ast})\frac{k}{d} - (T_{1}^{\ast} - T_{0}^{\ast})/(h - d))}{\lambda}.
\end{align*}
\]

We note that we can still incorporate time dependence in this model through the outside temperature \( T_{\text{out}} \) and then evolve our model quasi-statically, i.e., assuming equilibrium within each apartment at each time step for the changes in the outside temperature.

5 Statistical Approach

5.1 Linear Regression Model Assumptions

The approach presented in section 3 is based on a simple heat flow model based on the technical properties of the wall. However, the model is not possible to implement if the technical data is not available. Also, it may be inaccurate if there are any faults in the isolation or if other factors influencing heat conductivity are not included.

The model proposed in this section allows to estimate the heat production and transfer when no technical data are available apart from the basic information about the geometry of the building (volume and location of the rooms). Also, it allows to estimate to some extent heat transfer through the outer wall (see figure 2).

It is based on analysing the time series representing the variation of temperature in each room in a certain period of time (e.g., day, week, month) and using it to estimate the effective heat conductivity for each wall and the average amount of heat produced inside the room. Former parameters can be used to calculate the current rate of heat transfer through each wall in the building.

As it is shown in section 5.2, the mathematical formulation of the model leads to a system of linear equations, which can be solved easily using simple numerical methods available in most computation environments (Python, R, MatLAB etc).

The model is based on the following assumptions:
1. The heat transfer through the wall is proportional to the temperature difference across it.

2. The temperature inside each room is uniform.

3. The heat produced in the room is uncorrelated with the temperature difference between the room and its neighbours.

Assumption 1, also called the Newton’s law of cooling, is very reasonable and is consistent with 1D heat conduction model of a wall with two sides kept at specified temperatures (see engineeringtoolbox.com)

Assumption 2 may be inaccurate, but without including the very details of the room geometry (for example locations of radiators, windows etc.) and complex heat simulations the variation of heat cannot be easily included. Also, in some extent we expect that in rooms of a standard size the temperature differences are insignificant.

Assumption 3 is the most problematic. The heat produced and temperature inside the room often varies periodically with time. For example residents will produce most heat in the evening, when the temperature outside is lower than average. From a purely statistical point of view such situation is indistinguishable from the situation of flat in which no additional heat is produced, but the walls are better isolated. This may cause the heat conductivity to be slightly underestimated. This inaccuracy can be limited by reinforcing the model with the technical data about the walls and heat provided to the building. Such approach is presented in section 5.3.

5.2 Linear Regression Model Formulation

Let’s focus on one particular room. The aim of the model is to predict the rate of heat transfer from and to other neighboring rooms and the rate of heat produced inside the considered room. The outside of the building is assumed to be a room with the known (outdoor) temperature.

In this section the following notation is used:
1. $\kappa$ is the amount of heat needed to rise the temperature in the room by $1\degree C$ (it can be estimated by $\kappa = V\rho c$, where $V$ is volume of the room, $\rho$ density of air and $c$ specific heat capacity of the air),

2. $T^n$ is temperature in the considered room in $n^{th}$ time step,

3. $T^n_i$ is temperature of its $i^{th}$ neighbouring room in $n^{th}$ time step,

4. $\Delta t$ is the time between consecutive temperature measurements,

5. $q^n$ is the rate of produced heat inside the considered room in $n^{th}$ time step

From energy conservation we can write that the rate of change of room’s internal energy is equal to the sum of rate of heat production within the room and rate of heat flow through the walls into the room (in case of heat flowing out of the room this term has negative sign).

Based on the 1st assumption the heat flow from $i^{th}$ neighbour to considered room can be expressed as $k_i(T^n_i - T^n)$ (see Figure 6). Therefore we get the equation (28) for

Figure 5: Example of thermogram. The statistical model includes impact of all factors on heat transfer like windows (building on right) and bad insulation (building in the back). [author: Passivhaus Institut]

Figure 6: Heat transfer in case of a room with three neighbours.
energy conservation within the room:

\[
\kappa \frac{T^{n+1} - T^n}{\Delta t} = q^n + \sum_{i=0}^{N} k_i (T^n_i - T^n) \tag{28}
\]

The key step in modelling process is to decide which factors should impact the rate of heat production \( q^n \). Based on these factors one can find the expected value for the rate of the heat production \( \hat{q}^n \) as a linear combination of those factors. The real heat produced in the room can be decomposed into its expected value \( \hat{q}^n \) based on the linear model and its error \( \epsilon \), i.e. \( q = \hat{q}^n + \epsilon \).

For choosing the linear model to describe the rate of heat production We suggest one of the following three approaches:

1. Constant heat production: \( q = k_0 + \epsilon \)
2. Heat linearly dependent on outside temperature: \( q = a + bT_{\text{out}} + \epsilon \)
3. Constant temperature of the radiator\(^1\): \( q = a + bT + \epsilon \)

Here \( k_0, a \) and \( b \) are constant parameters (most likely \( k_0, a > 0 \) and \( b < 0 \)), which can be found using linear regression method described later on. The choice of the model depends on the way that heat is regulated in the model and some more advanced model (dependent on more factors) may be proposed.

From now on the linear regression model will be now presented for the 1st approach (assuming constant heat production), however applying any other of the suggested approaches is analogous. It is assumed that the heat produced in the room can be decomposed into its average value \( k_0 \) and time-varying part \( \epsilon \), i.e. \( q = k_0 + \epsilon \). Now the equation 28 can be expressed as:

\(^1\) After assuming that the radiated heat is proportional to the temperature difference between the temperature in the room \( T \) and the constant temperature of radiators (\( T_{\text{rad}} \)) one can easy see that it can linearly related to \( T \).
\[
\frac{T_i^{n+1} - T_i^n}{\Delta t} = k_0 + \epsilon^n + \sum_{i=0}^M k_i (T_i^n - T^n)
\]  
(29)

This form can be rewritten in a matrix form with each row corresponding to different time step \( n \):

\[
y = Ak + \epsilon
\]
(30)

where \( y = \begin{pmatrix} T^1 - T^0 \\ T^2 - T^1 \\ \vdots \\ T^N - T^{N-1} \end{pmatrix} \), \( k = \begin{pmatrix} k_0 \\ k_1 \\ \vdots \\ k_M \end{pmatrix} \) and \( A = \begin{pmatrix} 1 & T_1^1 - T_1^1 & \cdots & T_M^1 - T^1 \\ 1 & T_1^2 - T_2^2 & \cdots & T_M^2 - T^2 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & T_1^N - T_N^N & \cdots & T_M^N - T^N \end{pmatrix} \).

Vector of response variables \( y \) and matrix of explanatory variables \( A \) can be found based on available data. The parameter vector is the only unknown and can be estimated using the ordinary least squares method (minimising \( \epsilon^\top \epsilon \) term). The matrix aren’t large and can be easily computed using either direct (e.g. Cholesky decomposition) or iterative methods (e.g. Gauss-Seidel method). Direct methods finds the exact solution expressed as \( \hat{k} = (A^\top A)^{-1} A^\top y \). The iterative methods are iteratively approximating the exact solution and are much faster. A lot of information about regression methods and implementation can be found in the literature\[Björck, 1996, Watkins, 2002, Weisberg, 2005\].

After estimating the vector \( k \), the estimated rate of heat production in the room \( r_0 \) is already known. Now one can find the rate of change of heat transfer through the wall with neighbour \( i \) using equation (31).

\[
Q_{i}^{\text{now}} = k_i (T_{i}^{\text{now}} - T^{\text{now}})
\]
(31)

The positive value of heat transfer means that the heat is flowing into the room, and negative corresponds to flux out of the room. The same model can be applied for all other rooms to model the heat transfer in the entire building. Hopefully, the \( k \)
parameter between room A and B should not be dependent on which room is considered for the model. This fact can be used to estimate how precise the linear regression model is. If the precision is acceptable one can take the average of them as the best estimate of \( k \) parameter between these rooms.

### 5.3 Linear Regression Model with correction

One can try to improve the precision of the presented linear regression model by taking into account other estimations of either \( k_i \) parameters or heat production term \( q^N \). The former can be estimated based on technical data about the building’s structure. The latter can be estimated using heat measuring devices (which apart the limited precision do not cover some heat sources, like RTV/AGD). The aim of linear regression model with correction is to find corrections to these estimations to correctly describe the temperature variation in each room.

The idea is to decompose \( k_i = \overline{k}_i + k'_i \), where \( \overline{k}_i \) are independently estimated values and \( k'_i \) are (yet unknown) corrections to these estimations.

The equation (29) can be rewritten as:

\[
\kappa T_i^{n+1} - T_i^n + \Delta t - \kappa k_0 - \sum_{i=0}^{M} \overline{k}_i (T_i^m - T^m) = k'_0 + \epsilon^n + \sum_{i=0}^{M} k'_i (T_i^n - T^n) \quad (32)
\]

This form can also be rewritten in a matrix form (30). Here however vector of a response variable is defined as:

\[
y = \begin{pmatrix}
\kappa T_1^1 - T_0^n - \kappa \overline{k}_0 - \sum_{i=0}^{M} \overline{k}_i (T_0^i - T^0) \\
\kappa T_2^1 - T_1^n - \kappa \overline{k}_0 - \sum_{i=0}^{M} \overline{k}_i (T_1^i - T^1) \\
\vdots \\
\kappa T_N^N - T_{N-1}^N - \kappa \overline{k}_0 - \sum_{i=0}^{M} \overline{k}_i (T_{N-1}^i - T^{N-1})
\end{pmatrix} \quad (33)
\]

Using least squares method on this linear problem leads to the same solution as in the linear model without correction. However one can choose to use ridge regression which,
Finding a Fair Heat Settlement Model

unlike ordinary least squares, tries to minimise estimated \( k_i' \) parameters. It introduces an additional term to minimised function, which penalises high parameter values. The minimised function has the following form:

\[
\hat{k} = \arg \min_k \left[ (y - Ak)^\top (y - Ak) + \Gamma k \right].
\]  \hspace{1cm} (34)

The solution of the minimisation procedure is:

\[
\hat{k} = (A^\top A + \Gamma^\top \Gamma)^{-1} A^\top y. \hspace{1cm} (35)
\]

\( \Gamma \) is the \( p \times p \) matrix called Tikhonov matrix. It usually have a form of a diagonal matrix. The choice of diagonal terms is arbitrary. Generally higher values on the diagonal gives greater penalty for high parameter values. Additionally ridge regression avoids overfitting by assigning too high parameters. If reliable estimations are available the solution should be more accurate than in the model without the correction.

5.4 Numerical validation of the Linear Regression Model

5.4.1 Test 1. Linear Regression Model precision

To check the precision of the linear regression model a data from the real building should be performed and compared with the heat calculated using other models based on the physical modelling, such as the one described in section 3. However the real data were not available for the study group.

The only proof-of-concept that could be easily done was to check whether the rate of heat transfer and production could be reconstructed for a artificially generated data with known \( k \) and \( q \) parameters in the Equation 29. The data was generated for a 1000 array of rooms connected with a walls \( (k = 0.1) \) and outdoor \( (k = 0.05) \). Each was characterised by a different heating rate from 0 to 2 with an average of \( Q = 0.9 \). The outdoor temperature oscillated between 3\(^\circ\)C and 12\(^\circ\)C. The data were generated
for a period of one week with a time interval between the measurements of $\Delta t = 15$ minutes.

Three experiments were conducted. In each of those the data were modified with an $\epsilon$ randomly generated from a uniform distribution with standard deviation of 0.01, 0.1 and 1 respectively.

The results are presented on Figure 7. The three graphs on right hand side presents how is the estimated heating rate related to its real one for each of 1000 rooms. In all experiments high accuracy of the model was observed, with precision significantly decreasing in case of the highest variance of random term $\epsilon$.

The graphs in the centre and on the left presents what is the distribution of $k$ values for the external walls between the rooms and outside and interior walls between the rooms respectively. The values are distributed around the real value used during data generation (0.05 and 0.1 respectively).

Figure 7: Results of the tests of the statistical method on randomly generated data.
It’s worth noting however, that the estimations for the internal walls are much less precise than the one for external walls. This is caused by not significant temperature gradients between the rooms comparing to gradients between the rooms and outside, which are directly responsible for observed changes of temperature.

5.4.2 Test 2. Linear Regression Model accuracy

The results in the previous test are accurate (unbiased), which was expected as the data were also generated using linear model assumptions. If the assumptions are not met then the results may become biased. To observe that a second dataset was generated. This time the amount of heat produced in a flat was dependent on the temperature outside the building (the heating rate in each room decreases with the rise of temperature outside the building). The error term $\epsilon$ was set to be 1%.

Three different models were tested:

1. **Constant heat model** - the same as used in test 1

2. **Constant heat model with correction** - the expected value of heat conductivity and rate of heat production was assumed to be constant equal to its average value among all rooms and walls.

3. **Heat linearly related to the outside temperature** - based on the same assumptions as the algorithm used to generate the data set.

Each of these three models was used to find the heating rate for each room and effective heat conductivity for each wall (interior and exterior). The Figure 8 presents the comparison between the obtained estimations and the real values of these parameters.

The constant heat model lead to underestimated rate of heat production by approx. 10% and underestimation of exterior wall conductivity by approx. 5%. This observation can be easily explained. As the heat rate is increasing when the outdoor temperature falls down, the temperature is getting lower more slowly. As the model assumes that the rate of heating is constant the slow cooling of the air in the room is associated not with more intense heating, but with the better isolation of the exterior wall.
Figure 8: Results of three different models on randomly generated data.

As the model with correction was based on the average value the underestimated results only were corrected when the estimated value was lower than the real one. Otherwise, the results were slightly worse. Still such approach gives the better estimate of the real heating rate and heat transfer through the walls as the physical model based on the assumption that heating rate and heat conductivity is constant for all flats. The statistical model correctly shows how their real values vary from expected ones.

The best results were obtained with the last model, which was not assuming that the average heating rate is uncorrelated with the outside temperature. Although the results are more accurate and unbiased the spread of the data has clearly increased, especially in case of exterior wall’s conductivity. This is caused by the fact that the outside temperature and temperature difference across external wall are highly correlated, which highly limits the precision of the linear regression model.
5.5 **Summary of the Linear Regression Model**

To summarize, although in the numerical tests the linear regression model turned out to work correctly for the generated data using the same assumptions it still need to be tested on real-world data to check whether this simplification is acceptable. As the second test demonstrated the results may turn out to be biased when the assumptions are not met. When applying any linear regression model one should compare the outcomes of different models based on real data for which the comparison between estimations and the real value is possible.

The **advantages** of the linear model presented in this section are:

- it does not require any technical data and advanced physical modelling,
- it includes all factors influencing the heat production and transfer (i.e. bad insulation and heat generated by RTV/AGD),
- the calculations based on linear regression are easy to implement and fast to perform,
- it can be used to estimate the heat losses through all types of walls including outer walls with windows, doors etc.
- in the modified version it can be used to correct the independent estimations of heat production/transfer (for example obtained using the model presented in section 3).

However the model also have few **drawbacks**. The most important are:

- the model becomes inaccurate if the heat production is highly correlated with the temperature difference between the rooms,
- imprecise and faulty data may highly influence the estimations.
6 Possible Generalisations and Extensions to the Model

Possible generalisation of the heat settlement model:

- include home appliances including the electric energy consumption in individual flats,
- use additional data such as humidity sensor that can for example give information about open windows etc.

Generalisation to the mathematical model of heat distribution:

- consider more general boundary conditions i.e. Robin-type boundary conditions with the outside world that might take into account convection due to wind, for example,
- use effective conductivity found by homogenisation of the heat equation for various geometries,
- consider convection and radiative transfer within apartments and more importantly through staircases and common areas,
- include solar radiation.

7 Discussion and Conclusions

We have proposed a heat settlement model based on the heat distribution in the building. The heat distribution has been modelled in two different ways: continuum modelling using the heat equation and statistical modelling based on linear regression. In the continuum modelling, we used a temporal and spatial discretisation of the heat equation corresponding to the time span and location of the temperature sensors. Then, we had an inverse problem concerning finding the heat sources in each apartment based on the temperature readings on the sensors. This model had a limitation
in that it assumed constant temperature within each apartment. Our other continuum model represented the actual radiators within each apartment as point sources, and we retained spatial variance of the temperature. However, this model assumed steady state, which is not too unreasonable given that we can still incorporate time dependence quasi-statically through a time-varying function for the outside temperature. Finally, we presented a statistical model, which aimed to estimate the heat usage in each apartment based on the temperature readings using linear regression. Our models show improvement upon previously used settlement models when looking at various scenarios of different human behaviour in terms of heat usage, in that they provide fairer charging strategies.

Acknowledgements

We are grateful to Prof. Hilary Ockendon and Prof. John Ockendon for numerous useful ideas and discussions, especially the part in Section 4. In addition, we would like to thank a number of contributors, who helped with a lot of helpful discussions: Piotr Bies, Filip Rupniewski, Jacek Szczytko.

References

[Björck, 1996] Björck, Å. (1996). *Numerical Methods for Least Squares Problems*, volume 135.

[Champneys et al., 2013] Champneys, A., Fresneda-Portillo, C., Hewitt, I., Liu, X., Hunt, M., Nagapetyan, T., Please, C., Shang, X., Tant, K., Witelski, T., Wood, D., and Zioło, P. (2013). Estimating the spread of fire in buildings. pages 1–27.

[Lagree, 2018] Lagree, P. Y. (2018). Homogenisation. *CNRS & SU*.

[Madsen and Holst, 1995] Madsen, H. and Holst, J. (1995). Estimation of continuous-time models for the heat dynamics of a building. *Energy and Buildings*, 22:67–79.
[Sakakibara, 1991] Sakakibara, Y. (1991). Numerical study of heat storage in a building. *Energy and Buildings*, 16:577–586.

[Watkins, 2002] Watkins, D. S. (2002). *Fundamentals of Matrix Computations // Iterative Methods for Linear Systems*, pages 521–601.

[Weisberg, 2005] Weisberg, S. (2005). *Applied Linear Regression // Simple Linear Regression*, pages 19–46. Wiley Series in Probability and Statistics.