APPLICATION OF SPATIAL-TEMPORAL CLUSTERING TO FACILITATE ENERGY SYSTEM MODELLING

Julien F. Marquant\textsuperscript{1,2}, Akomeno O. Omu\textsuperscript{1,2}, Kristina Orehoung\textsuperscript{1,2}, Ralph Evins\textsuperscript{1,2}, Jan Carmeliet\textsuperscript{1,3}

\textsuperscript{1}Chair of Building Physics, Swiss Federal Institute of Technology, ETH Zürich, Switzerland
\textsuperscript{2}Laboratory for Urban Energy Systems, Swiss Federal Laboratories for Materials Science and Technology, EMPA, Dübendorf, Switzerland
\textsuperscript{3}Laboratory for Multiscale Studies in Building Physics, Swiss Federal Laboratories for Materials Science and Technology, EMPA, Dübendorf, Switzerland

ABSTRACT

Network-level energy optimisation approaches aim to quantify the benefits of centralised and decentralised energy systems through the optimal siting of generation technologies within a region and the design of optimal energy distribution networks to link them. However, computational limitations arise at larger spatial scales, and difficulties emerge in the quantification of the different urban agglomeration levels when attempting to model varying network behaviours at multiple spatial scales. Nevertheless, a tractable multi-scale optimisation of the design and operation strategy for a large urban area is possible using a clustered, aggregated energy hub representation. In this paper, one such clustering method is applied to a district-scale problem consisting of multiple energy hubs. A spatial-temporal radius-based clustering method incorporating treatment of outliers is explored and adapted to the energy hub formulation. The clustering method is adjusted to account for geo-dependent demand data as well as typical daily load profiles. In addition to reducing the computational burden, the main benefit of this energy hub formulation is that it results in a more appropriate sizing of energy generation units, minimising part load operation and thus improving overall system efficiency.

INTRODUCTION

Achieving the European Commission’s 30% energy reduction target by 2030 will be difficult through demand-side measures alone (Corsepius, European Council, 2014). Supply-side interventions, such as an increase in the utilisation of renewable energy resources or the adoption of high-efficiency co-generation systems, will likely be required. In order to effectively integrate large amounts of renewable and high-efficiency technologies these decentralised energy systems must be able to manage fluctuating and distributed power sources, store energy, and convert energy from one carrier to another. This is also a multi-scale problem as it is not clear whether decentralisation is optimal if applied at the building scale, district scale, city scale, or even regional scale.

In order to tackle this complex problem, (Geidl et al., 2007) developed the energy hub concept which facilitates the optimisation of the interactions between multiple energy systems and carriers, using linear equations to describe the relationship between input and output energy streams with a conversion matrix representing technology efficiencies. The optimisation problem is formulated as a Mixed Integer Linear Programming (MILP) problem. MILP formulations are reliable and efficient, but the solving time can increase exponentially with the number of integer variables (Domínguez-Muñoz et al., 2011). Thus, the multi-scale optimisation of the design and operation of such energy systems is not trivial, especially when numerous distributed generation resource combinations are taken into account, resulting in highly integrated electricity, gas, and district heating network optimisation problems that become intractable as the spatial and temporal scales are increased.

It is possible to reduce the computational burden of these larger scale energy hub formulations without sacrificing the quality of the results by developing new techniques that allow for the increase of the spatial or temporal scale without increasing the number of decision variables. Increasing temporal scales can be modelled without increasing the number of time variables by employing typical days to represent the full-length problem. Different clustering methods are currently employed to select those days. (Fazlollahi et al., 2014a) use a k-means clustering method to select typical periods, (Morvaj et al., 2014) represent each month by an average day and (Domínguez-Muñoz et al., 2011) use a k-medoids clustering algorithm to select the best representative day of each computed cluster.

As the spatial scale is increased from buildings, to districts, to cities, and even countries, the number of variables used to represent the increasing system size can also be a challenge, and a clustering method is again required. In previous works, the clustering is often manually pre-defined taking into account building archetypes or energy demands. (Omu et al., 2013a) cluster buildings as a function of their yearly heating demand in order to optimize distributed energy resources. (Weber and Shah, 2011) optimise the district energy system for an eco-town in the UK and cluster buildings into single-type nodes (e.g. residential buildings, offices, schools) where the demand profiles are aggregated. (Keirstead et al., 2014).
also aggregate demand profiles in zones based on building locations in order to evaluate biomass energy strategies for the same eco-town. (Fazlollahi et al., 2014b) represent an urban area as a set of energy strategies for the same eco-town. (Fazlollahi et al., 2014b) apply the k-means algorithm. For different building characteristics and annual energy demand, (Fonseca and Schlueter, 2015) apply the k-means algorithm available in ArcGIS. Based on building location combined with the Getis-ord spatial statistic, they are able to identify the quality of the cluster using the intensity of clustering and its probability of occurrence. In this approach, outliers are manually detected based on their high energy demands.

In this work, a new clustering method is developed in order to facilitate the simulation of large-scale models while maintaining a reliable representation of the energy network interactions at the urban scale. The clustering method is spatially based, however the temporal variability in energy demand is also taken into account. The goal of the clustering is to be able to define the most representative cluster in order to correctly aggregate an integrated area and its district-heating network. Furthermore, the method facilitates the analysis of the trade-off between centralised and decentralised energy systems, resulting in more accurate sizing of a centralised energy system for a district-heating network.

A Non-Naive-typical day method based on (Domínguez-Muñoz et al., 2011) is developed. This method is used to select the optimal number of typical days as a function of the Davies-Bouldin criterion, which uses the ratio between the within-cluster and the between-cluster distances, and the error in the load duration curve computed from the typical days output demand profile.

**MODEL DESCRIPTION**

**Overview**

The objective of this study is to determine the optimal design and operating strategy of a distributed energy system for a city in order to minimise the investment and operating costs. The environmental impact of the different solutions is also analysed by calculating the carbon emissions associated with the minimum cost solutions. Different cases are analysed from fully decentralised, in which all buildings are independent, to fully centralised, in which all buildings are connected by a district-heating network. The study investigates the trade-off between centralised and decentralised energy systems and the sizing of the centralised energy system. The building with the highest annual heat-demand is selected as the anchor building for the clustering method. A cluster is spatially built from this anchor by aggregating the surrounding buildings as a function of a radial distance r. A district heating network is designed in the clustered area assuming that all the buildings are connected to the network. The district heating network price is calculated considering a cost for the pipes $C_{\text{pipe}}$. The shape of the network is computed with a minimum spanning tree algorithm. In order to better understand the impact of the aggregation, the number of technologies that can be installed has been limited to gas driven combined heat and power (CHP) systems and gas boilers.

The MILP formulation model is based on the energy hub models developed by (Evins et al., 2014) and (Morvaj et al., 2014). The buildings are linked to the electricity grid and can always buy or sell electricity. A more detailed formulation is used to adapt the cost and efficiency of technologies depending on the size. As in (Mehleri et al., 2013), the distribution losses within the heating network are assumed to be negligible and are not considered because the distances between the different energy hubs are sufficiently short, i.e. tens of metres.

In order to consider stochasticity in the occupancy schedules, an artificial variability is introduced using the method applied in (Mavromatidis et al., 2015). This permutation method (referred as loads permutation in Figure 1), maintains the specific energy patterns that are representative of the building types (e.g. morning and evening peaks for residential buildings), while randomly permuting the loads of each building across 3 hour temporal blocks. This method maintains the highest peak and total demand per building while randomly mixing the peak hours of all the buildings so that the peaks are not artificially amplified during the aggregation process.

**Objective function**

The cost function that is minimised is the sum of operating cost and investment cost per technology installed (t) in each energy hub i:

$$\min \{ \sum_i \sum_t (I_t + OC) \}$$

(1)

The investment cost consists of the sum of the cost of each technology installed in every energy hub. We take into consideration economies of scale by creating size tiers for the price $(LC^{\text{bound}})$.  

$$I_t = \sum_{\text{bound}} (CY_t^{\text{bound}} \cdot LC^{\text{bound}})$$

(2)

Equation 2, where $CY_t^{\text{bound}}$ is a variable linearising the multiplication of a binary variable $\delta_t^{\text{bound}}$ (1 if the size of the installed technology is between the upper $(UP)$ and lower bound $(LW)$ associated to the variable, else 0) and $C_t^{\text{max}}$, capacity of the installed technology.

$$CY_t^{\text{bound}} = \begin{cases} C_t^{\text{max}}, & \delta_t^{\text{bound}} \geq [UP \ LW] \\ 0, & \delta_t^{\text{bound}} \notin [UP \ LW] \quad (\delta_t^{\text{bound}} = 0) \end{cases}$$

(3)

Another binary variable $\delta_t$ assures that the term in Equation 2 is zero if the technology is not installed. The operational cost (Equation 4) consists of the cost of the electricity purchased from the grid $E_t^{\text{imp}}$, the
cost of the gas used in the natural gas boiler \( P_{\text{boiler}}/\eta \) and CHP \( P_{\text{CHP}}/\eta_{\text{CHP}} \) minus the income generated by selling electricity produced from the CHP system \( P_{self}^{CHP} \). In order to reduce the computational burden, a full year is represented by typical days extracted using a k-medoids algorithm (Domínguez-Muñoz et al., 2011). The term \( Day_h \) represents the number of days in the cluster \( h \). As in (Morvaj et al., 2014) and (Weber and Shah, 2011), the operating costs are calculated for a 20 years period assumed to be the average lifespan of equipment. The Net Present Value (NPV) is calculated in order to take into account the time value of money, and a discount rate of 3% is assumed.

\[
OC = \sum_t \sum_h \left( E_{grid}^{imp} \cdot C_{buy}^{grid} + C_{buy}^{gas} \right) \left( \frac{P_{\text{boiler}}}{\eta} + \frac{P_{\text{CHP}}}{\eta_{\text{CHP}}} \right) - \left( C_{\text{gas}}^{self} \cdot P_{self}^{CHP} \right) \cdot Day_h \cdot NPV
\]

The carbon emissions for the cost-minimised solution are calculated using Equation 5. Only emissions that are due to the operation of the energy systems during the 20 years horizon \( T \) are calculated, and embodied carbon is not taken into account. \( Carb_{elec}^{em} \) and \( Carb_{gas}^{em} \) are the carbon factors in kgCO2/kWh of importing electricity from the grid and using natural gas, respectively.

\[
Carb_{elec}^{em} = \sum_t \sum_h \left( E_{\text{imp}}^{grid} \cdot Carb_{elec}^{em} \right)
\]

\[
Carb_{gas}^{em} = \left( \frac{P_{\text{boiler}}}{\eta} + \frac{P_{\text{CHP}}}{\eta_{\text{CHP}}} \right) \cdot Day_h \cdot T
\]

Energy demand constraints

Additional constraints are required to ensure the supply of the buildings’ electricity and heating loads during every time period. Electricity is produced or supplied from the grid for each energy hub (Equation 6) where \( P_{self}^{CHP} \) is the electricity produced and directly used by an energy hub.

\[
L_{\text{elec}} = \sum_t \sum_h \left( E_{grid}^{imp} + P_{self}^{CHP} \right)
\]

Heat loads in each energy hub \( i \) are supplied from a CHP system and/or boiler:

\[
L_{\text{heat}} = \sum_t \sum_h P_{\text{CHP}} \cdot R_{\text{HP}} + P_{\text{boiler}} \cdot \eta
\]

The efficiency of the boiler \( \eta \) is assumed fixed at 0.8 for all sizes of boiler, just as the heat-to-power ratio \( R_{\text{HP}} \) for all sizes of CHP system is assumed to be 2. The model is heat-driven and electricity produced at a building can be used or sold to the grid with the following constraint:

\[
P_{\text{CHP}} = P_{self}^{CHP} + P_{self}^{CHP}
\]

As in (Morvaj et al., 2014), additional constraints (Equations 9 and 10) are added in order to represent the 50% minimum part-load of a CHP system, where \( M \) is an arbitrary large number, and \( \delta_{\text{CHP}}^{ON} \) is a binary variable which is 1 if the installed CHP is on and 0 if not. Finally, the CHP electrical efficiency \( \eta_{\text{CHP}}^{\text{bound}} \) is defined as a function of the size of the selected technologies. \( P_{\text{CHP}}^{\text{bound}} \) is a variable used to linearise the multiplication of the binary variable \( \delta_{\text{CHP}}^{\text{bound}} \) (same as in Equation 3) and \( P_{\text{CHP}}^{\text{gas}} \) represents the gas input to the CHP system.

\[
\eta_{\text{CHP}}^{\text{bound}} \cdot P_{\text{CHP}}^{\text{bound}} \leq M \cdot \delta_{\text{CHP}}^{ON} \tag{9}
\]

\[
0.5 \cdot C_{\text{CHP}}^{\text{max}} \leq \eta_{\text{CHP}}^{\text{bound}} \cdot P_{\text{CHP}}^{\text{bound}} + M \cdot \delta_{\text{CHP}}^{ON} \cdot (1 - \delta_{\text{CHP}}^{ON}) \tag{10}
\]

Constraints related to the sizing of the technologies are also added in order to account for lower and upper bound limitations.

Reference model

The aim of the clustering method developed in this paper is to facilitate the optimisation of large-scale energy systems. To do so, the loads of the different buildings within a cluster are aggregated and the heating network lengths and costs for the cluster are calculated in a simplified manner using the minimum spanning tree algorithm. However, the error that arises from this aggregation must also be determined. Thus, a reference model, named the “DH Optimisation Network” case, based on the energy hub model developed by (Morvaj et al., 2014) is used for comparison. In this reference model, the district-heating network is no longer designed during the preoptimisation phase. Instead, each cluster is disaggregated and every building is able to install its individual energy system. Heat can be produced and transferred from one building to another through the district-heating network whose structure is designed during the simulation phase. Equation 6 is modified to Equation 11, and Equation 12 is added to account for the district-heating network. Furthermore, a term (\( C_{\text{pipes}} \cdot Dist_{ij} \)) which represents the district heating cost is added to Equation 2, where the price of the pipes \( C_{\text{pipes}} \) is 200 CHF/m, \( Dist_{ij} \) is the distance between building \( i \) and \( j \), and \( Q_{ij} \) is the heat exchange between building \( j \) and \( i \).

\[
L_{\text{heat}} = \sum_t \sum_h \left( P_{\text{CHP}} \cdot R_{\text{HP}} + P_{\text{boiler}} \cdot \eta \right) + \sum_j Q_{ij} - Q_{ij} \tag{11}
\]

Heat is exchanged only if there is a connection between the two buildings \( (\delta_{\text{pipe}}^{ij} = 1 \) if there is a pipe between building \( i \) and \( j \) ) and the connection mono-directional.

\[
Q_{ij} \leq \delta_{\text{pipe}}^{ij} \cdot M \text{ and } \delta_{\text{pipe}}^{ij} + \delta_{\text{pipe}}^{ji} \leq 1 \tag{12}
\]

Figure 1 illustrates the routine and the different comparisons in terms of assumptions made and software used. The “Aggregate cluster” routine consists of running the energy hub model with the district heating network optimised by the minimum spanning tree algorithm and assuming that there is a central location for the installation of the CHPs and...
boilers. The “Disaggregate” routine also employs the district-heating network that is determined by the minimum spanning tree algorithm, however the CHPs and boilers are installed at each building within a cluster. Finally, the “DH Network Optimisation” routine, which is the reference model, optimises the design of both the technology set and the district-heating network structure within the energy hub optimisation model.

The typical heating and electricity demand is extracted from these 10 buildings and a scaling factor is applied to compute a typical residential demand for other buildings in the neighbourhood as a function of their floor area.

The non-naive typical day method is employed to represent the full year demand profiles with the most representative days. The method is applied to select the typical days representing the aggregated cluster. The selection of the typical days of the outlier buildings is aligned with the cluster. The errors in load duration curve (ELDC) and Davies-Bouldin index are minimised. The best combinations give the optimal number of typical days that should be selected to represent the cluster. In this case the number of typical days selected is 12, 10 k-medoids typical days plus 2 peak days corresponding to the peak heating and peak electricity days. The corresponding days are selected for the other (non-clustered) buildings that are considered as outliers. Figure 2 shows the error in load duration curve for the heating demand for the anchor building. The red line is the load duration curve of the full heating demand and the blue line is the typical days load duration curve for the reconstituted year. The errors in reproducing the heating and electricity load duration curves from the typical days approach are 3.45% and 0.60%, respectively.

CASE STUDY

A fictitious case study has been built from the case study of (Orehounig et al., 2014) regarding the energy consumption of a neighbourhood. Electricity profiles were sourced from SIA 2024 (Merkblatt, 2006) and heating demand was simulated in (Orehounig et al., 2014) using the simulation software CitySim (Robinson et al., 2009). The demand profiles are simulated with an hourly resolution. The case study considers 103 buildings. One case examined a mixture of building types including supermarkets, office buildings, schools, and various types of residential buildings. Another case is built from 10 typical residential buildings.

Economic model

Table 1 shows the assumptions for the CHP and gas boiler $L_{\text{bound}}^\text{boiler}$ and $n_{\text{bound}}^\text{boiler}$ from (Beith, 2011), (Omu et al., 2013b) and (Pöyry, 2009).

<table>
<thead>
<tr>
<th>SIZE [KW]</th>
<th>EFF</th>
<th>ELEC CHP $\eta_{\text{elec}}^\text{bound}$</th>
<th>COST CHP [CHF/KW]</th>
<th>COST BOILER [CHF/KW]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 – 20</td>
<td>0.25</td>
<td>1128</td>
<td>211.5</td>
<td></td>
</tr>
<tr>
<td>21 – 50</td>
<td>0.27</td>
<td>775.5</td>
<td>176.25</td>
<td></td>
</tr>
<tr>
<td>51 – 180</td>
<td>0.3</td>
<td>564</td>
<td>131.13</td>
<td></td>
</tr>
<tr>
<td>181 – 350</td>
<td>0.3</td>
<td>564</td>
<td>111.39</td>
<td></td>
</tr>
<tr>
<td>351 – 500</td>
<td>0.3</td>
<td>564</td>
<td>91.65</td>
<td></td>
</tr>
<tr>
<td>501 – 5000</td>
<td>0.32</td>
<td>493.5</td>
<td>42.3</td>
<td></td>
</tr>
</tbody>
</table>
Clustering model

The radius-based clustering method is utilised for the different values listed in Figure 3. The different radius-based clusters (C₀-C₂₀) can be understood as an evolution between the fully decentralised neighbourhood with all buildings independent (C₀, all buildings are considered as independent energy hubs) to fully centralised neighbourhood (C₂₀, energy demand profiles of all 103 buildings are aggregated and considered as a single energy hub). This study is conducted for the two cases, mixed and residential, and the same anchor building is used in both cases. The aim of the analysis is to understand if the robustness of the clustering method is dependent on the types of buildings, and therefore diurnal demand patterns, that are being clustered. Thus highlighting the importance of the temporal demand aspect in the spatial-temporal clustering method.

RESULTS

In Figure 4, the economic results for the residential case are presented for the 20 clusters plus the fully decentralised case (C₀). The columns represent the investment costs for the technologies in blue (central plant) and red (decentralised plant) and the cost of the heating district network in green, while the purple line represents the operating costs.

Figure 4 Investment and operating costs for the residential case

The cost of the district-heating network is calculated assuming a network that connects all the buildings in the cluster. Figure 4 shows that the overall district technology investment cost decreases as more buildings are added to the cluster. Furthermore, the trade-off between the cost of the decentralised technologies installed at the outlier buildings and the centralised technologies installed in the cluster reflects the influence of economies of scale. However, as the centralised generation capacity increases (and thus technology investment costs decrease), the required district-heating network also increases significantly, as a longer network is required to connect more buildings. The analysis is carried out over a 20-year time horizon, therefore the operating costs dominate over the investment costs. As a consequence, the total costs of each cluster, calculated as the sum of the operating and investment costs, shows the same pattern as the operating costs. From cluster 13 (i.e. 91 clustered buildings) onwards, a centralised system in combination with a district-heating network starts to become advantageous when compared with the decentralised case (C₀). This is due to the fact that the number of aggregated profiles is high enough to smooth the aggregate loads, thus allowing the selection of a CHP unit that is sized in a manner that maximises CHP operation, and thus minimises the amount of electricity that must be bought from the grid. The increase in operating costs for a number of clusters is a result of the oversizing of the CHP systems due to the aggregation of a small number of buildings with higher demand peaks. Due to the minimum load constraints, the CHP system is unable to operate for a greater number of hours, thus increasing the amount of electricity bought from the grid.
Table 2 shows the details of the technologies installed in the optimal energy system for each cluster and the outliers. The first column also shows the percentage of the total cluster energy load that can potentially be met by the CHP, while the second column shows the length of the district-heating network that is required to connect all the buildings in the cluster. As also seen in Figure 4, as more buildings are added to the cluster, the length of the district-heating network increases, the capacity of the centralised CHP and boilers generally increases, and the capacity of CHP and boilers in the outliers decreases. Similar results are seen in the mixed case.

Disaggregation
Figure 6 presents a map of the mixed case study area with the annual heating demand given for each building. The results of the three different optimisation processes (see Figure 1) are compared in the insets on the left of Figure 6 for C1.

Table 2 CHP and boiler capacities for centralised and decentralised plant for the residential case

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>AGGREGATE CLUSTER</th>
<th>OUTLIERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 Comparison between aggregated cluster and disaggregated case in time, total cost and capacity of the technology installed for the mixed case

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>C10</th>
<th>C4</th>
<th>C1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH Length [m]</td>
<td>2678</td>
<td>975</td>
<td>225</td>
</tr>
<tr>
<td>CHP [kW]</td>
<td>885</td>
<td>164</td>
<td>67</td>
</tr>
<tr>
<td>boiler [kW]</td>
<td>1228</td>
<td>501</td>
<td>124</td>
</tr>
<tr>
<td>Total cost cluster [kCHF]</td>
<td>14016</td>
<td>3657</td>
<td>1219</td>
</tr>
<tr>
<td>Computation Time [s]</td>
<td>1.55</td>
<td>1.41</td>
<td>1.23</td>
</tr>
<tr>
<td>Σ(i)CHP [kW]</td>
<td>1006</td>
<td>231</td>
<td>79</td>
</tr>
<tr>
<td>Σ(i)boiler [kW]</td>
<td>1035</td>
<td>269</td>
<td>95</td>
</tr>
<tr>
<td>Total cost cluster [kCHF]</td>
<td>15238</td>
<td>4085</td>
<td>1323</td>
</tr>
<tr>
<td>Computation Time [s]</td>
<td>50257 (14 h)</td>
<td>25846 (7 h)</td>
<td>49</td>
</tr>
</tbody>
</table>

While the Aggregated Cluster and Disaggregated routines use the same district heating network structure, the installation of generation technologies...
at each building results in a slightly higher overall installed capacity (but of smaller, most costly units), resulting in a higher technology investment cost. The total cost is likely higher for the Disaggregated routine because in the Aggregated Cluster routine district heating network the mono-directional constraint on the heating network is not considered. The Network Optimisation routine leads to similar investment costs as the Aggregated Cluster routine. This indicates that if generation units are to be installed at each building, it may be more accurate to either concurrently optimise the district-heating network thus allowing the inclusion of distribution losses in each section of the network, or iterate the minimum spanning tree network design process so that assumptions can be made about the locations of the generation units within the network. To further understand the impact of this minimum spanning tree network assumption, the Aggregated Cluster and Disaggregated routines were run in parallel for three of the clusters: C1, C4 and C10. Table 3 shows that in all three cases the Disaggregated routine selects an optimal energy system that is consistently around 10% more expensive than the one selected in the Aggregated Cluster routine, however, the computational time required increases drastically. In the C10 case, in which there are 65 buildings in the clusters, the Aggregate Cluster finds the solution in under 2 seconds, while the Disaggregated routine takes around 14 hours to solve.

CONCLUSION
This paper presents the first use of a spatial-temporal clustering method to facilitate energy system modelling at large scales. The results have shown that it is possible to model energy interactions at larger scales while maintaining computational tractability and without sacrificing the quality of the results, thus achieving time savings of several orders of magnitude. It is a first step toward a fully multi-scale approach energy hub modelling in which aggregations are performed at appropriate spatial and temporal levels in order to balance accuracy against run time.

In future work the method will be further developed to consider multiple cluster origins (i.e. multiple anchor loads), the location of the centralised energy system in the clustered area, and a density-based criterion instead of a solely radius-based criterion for the clustering method. A sensitivity analysis of the model to the economic assumptions will be performed and the energy hub model will consider the installation of more technologies such as long-term and short-term storage systems.

ACKNOWLEDGEMENT
This research has been financially supported by CTI within the SCCER FEEB&D (CTI.2014.0119).

REFERENCES


Meheri, E.D., Sarinveis, H., Markatos, N.C., Papageorgiou, L.G., 2013. Optimal design and operation of distributed energy systems:


Omu, A., Choudhary, R., Boies, A., 2013a. Distributed energy resource system optimisation using mixed integer linear programming. Energy Policy 61, 249–266. doi:10.1016/j.enpol.2013.05.009


