Using a Large Floorplan Dataset to Study the Impact of Simulation Zoning Simplifications

Aurelien Bres¹
¹AIT Austrian Institute of Technology

Abstract
Aggregating distinct spaces into a single simulation zone is a typical way of holding down the complexity of building energy models. Such zoning simplifications may yield models that are easier to understand and parameterize and faster to simulate, but whose results differ from those of high-fidelity models with more simulation zones. This contribution proposes a systematic investigation of the corresponding differences based on a few hundred distinct building geometries translated from a large floorplan dataset.

Key Innovations
- Thermal simulations based on a large floorplan dataset
- Statistical investigation of the impact of simulation zoning on various output indicators

Practical Implications
The impact of zoning simplifications is shown to vary widely with the floorplan on the one hand, and on the considered output indicator on the other hand. The difference in annual energy demand with a one-zone-per-floor model in comparison to a one-zone-per-room does not exceed 5% in most floorplans with the selected simulation parameters. On the other hand, differences in maximum space temperatures frequently exceed 1 K.

Introduction
Zoning is a central concept in most building energy modelling tools. A simulation zone represents a spatial unit for which heat balance is calculated, based on the assumption of homogeneous air temperature. Many of the input parameters and output variables of simulation are defined at the zone level. The way in which spaces are grouped into zones, referred to as simulation zoning, is a key aspect of model complexity. In comparison to a detailed zoning with one zone per room, zoning schemes grouping rooms according to some properties may result in faster run times and a more understandable model, but at the price of some loss of accuracy in results. While this loss of accuracy has been quantified in individual cases, it can be expected to depend strongly on individual building characteristics and especially on their respective floorplans. This contribution aims at investigating the impact of simulation zoning on results in a more systematic way, making use of a large dataset of floorplans. Simulation models are created and run for 402 distinct floorplans and with different zoning schemes. Differences between aggregated results obtained with detailed one-zone-per-room zoning and with simplified zoning are calculated for all these models and statistically described.

Related work
Impact of simplifications
The impacts of zoning simplifications and other types of simplifications on the results of building performance simulation have been studied in various contributions.

Heo et al. (2016) studied the impact of sequential zoning simplifications on a building with two floors and 11 zones in the original model. In this investigation, modelling the entire house with a single zone lead to an underestimation of the annual heating demand by 24% in comparison to a detailed model with every room as a separate zone. Bres et al. (2017) looked at the impact of simulation zoning and HVAC zoning on 5 different floorplans. The results implied significant differences between floorplans and a significant impact of the way different space uses were modelled (homogeneous or heterogeneous internal loads). The bias in annual heating demand reached -30% in one case, while remaining within the ±10% range in most cases. Korolija and Zhang (2013) studied the impact of model simplifications on energy and comfort indicators under consideration of varying parameters and looking at five house designs. A mean absolute relative error of 10.6% was calculated between models with detailed versus simplified zoning, for 2000 simulations of each case. A recent review on thermal zoning for building HVAC design and energy simulation (Shin and Haberl, 2019) concluded that there was still a need to “validate methods of thermal zoning”. Most of the above-mentioned studies used a single building or a few (up to five) buildings to assess the impact of zoning simplifications. Beyond coarse simulation zoning, other simplifications are commonly made in building energy mod-
elling and have been investigated in the literature. For instance, surfaces are often simplified, especially for complex and curved geometries (Santos et al., 2019). Zoning simplifications may go along with simplifications in terms of internal gains and their scheduling (Heo et al., 2016). They may also eliminate or reduce the need to specify interior constructions and interzonal air exchange.

**Impact of layout**

Considering planning situations where a building’s interior structure is not defined yet, zoning can be seen as a planning decision. The impact of this decision on energy performance, as investigated in some of the studies reported in this paragraph, can be related to the impact of zoning simulation on simulation results. A recent survey of the effects of architectural space layout on energy performance (Du et al., 2020) reviewed ten articles and identified different mechanisms underlying these effects.

For instance, comparing energy demand predictions for three layouts for an office building in London, De Souza and Alsaadani (2012) found “significant variations in predicted energy demands”. Musau and Steemers (2008) compared five different office layouts using thermal and lighting simulation and also concluded on significant differences. These analyses in turn may show variable results depending on factors such as occupancy (Musau and Steemers, 2008), internal partitions (Bojić et al., 2007) and interzonal airflow (O’Brien et al., 2011).

As for the literature reported in the previous paragraph, most of the simulation studies addressing the impact of layout have used one or a few example buildings. An exception is the work of Dogan et al. (2015), where 25 real floor plans corresponding to distinct morphologies were used to compare results obtained with detailed architectural zoning and with the perimeter and core zoning paradigm. The results showed a wide spread in the distribution of percentage errors in energy use intensity.

**Floorplan datasets**

In the present paper, a larger number of distinct floorplans is used to draw more generalizable conclusions. This is made possible by the availability of floorplan datasets. In recent years, floorplan datasets of increasing sizes have been used to train machine learning algorithms to extract information from floorplan images and have been made available. In 2015, Liu et al. (2015) used a dataset with 215 apartment floorplans for the reconstruction of 3D “virtual-tours” given two-dimensional floor plans and monocular images. In 2017, Liu et al. (2017) used a dataset containing 870 labeled floorplans to train an algorithm converting a rasterized floorplan image into a vector-graphics representation. In 2019, the CubiCasa5k dataset, a floorplan image dataset containing 5000 annotated samples, was released (Kalervo et al., 2019). In this last dataset, the annotations take the form of vector images in which a variety of objects and spaces are represented as polygons and tagged according to a given typology.

**Method**

**Objectives**

The present study aims at systematically investigating the impact of zoning simplifications on the results of building performance simulation, as measured by the differences in results between baseline models with one zone per room and models with coarser zoning. Using a few hundred - rather than a handful - distinct floorplans, a statistical estimation of this impact becomes possible. EnergyPlus (Crawley et al., 2001) version 9.3 is used as a simulation engine. This section describes the adopted methodology, which includes translating the two-dimensional vector images of the floorplan dataset into EnergyPlus input data files, performing systematic zoning simplifications according to various systematic zoning schemes and carrying out a number of simulation experiments.

**Translating annotated vector images into simulation models**

Annotated vector images from the CubiCasa5k dataset (Kalervo et al., 2019) are used to create a high number of simulation models with distinct realistic geometries. Although provided with rich annotations, these images need to undergo a number of transformations before they can be used for dynamic thermal simulation. The translation of annotated vector images into simulation models can be summarized with the following steps:

- Parsing of relevant polygonal elements: spaces, walls, windows and doors;
- Intersection of space boundary lines;
- Mapping of space boundary lines to exterior and interior walls;
- Mapping of windows and doors;
- Model checking;
- Extrusion in the third dimension;
- Enrichment with non-geometrical model content.

In the first step, the vector images in Scalable Vector Graphics (SVG) format are parsed with a generic Extensible Markup Language (XML) parser. The polygonal objects are filtered, leaving out objects not directly relevant for simulation such as furniture elements. External spaces, e.g. corresponding to balconies, which are also included in the original floorplans, are removed.

Segmentation of space boundary line segments is necessary in order to obtain pairs of interzone space boundaries as required for thermal simulation. The segmentation process is similar to existing methods for the transformation of first level space boundaries.
Figure 1: Example floorplan: space polygons. R: residential. U: unconditioned.

Figure 2: Example floorplan: clusters of space boundary lines (one color per cluster).

into second level space boundaries, as presented for instance in Rose and Bazjanac (2015) and Lilis et al. (2017)). However, the two-dimensional rather than three-dimensional setting makes it simpler. This segmentation process involves the following steps:

- Grouping of line segments into clusters of approximately collinear segments.
- For each line segment cluster, definition of an “average line” and projection of each line segment on this average line.
- Intersection of the projected segments. Let $a_1 < a_2 < \ldots < a_{n-1} < a_n$ be sorted one-dimensional coordinates of the projected points on the average line. An intersected segment is determined for each $1 \leq k \leq n - 1$. If the projected segment $[a_k, a_{k+1}]$ belongs to exactly one original segment, the intersected segment corresponds to an exposed (if adjacent to an exterior wall) or adiabatic space boundary. If it belongs to two original segments, it corresponds to a pair of interzone space boundaries. If it belongs to three or more original segments, the situation is considered to be pathological and the floorplan is rejected.

The model checking step involves verifying that all the previous steps could be carried out without errors and ensuring that the modeled floorplan satisfies some additional conditions that allow sensible comparisons to be made. Floorplans with atypical space types (e.g. sauna) are excluded, as are floorplans where the area of unidentified spaces exceeds 10%. Only floorplans with a floor area above a threshold of 50 m² and with at least 5 zones are selected.

In addition, fast visual checks are performed based on plots of surface types and space uses.

The extrusion in the third dimension is based on the assumption of standard heights for walls (3.0 m), windows (1.6 m) and doors (2.1 m). Also, floor and ceiling surfaces are modelled as adiabatic, following the assumption that each floorplan corresponds to an intermediate storey located neither on the ground floor nor on the last floor. The resulting space boundaries are illustrated in Figure 3.

Zoning schemes

For each floorplan, the following zoning schemes are considered:

- **room**: one zone per space in the original floorplans. This zoning scheme is the baseline against which differences are calculated. The results refer to the differences between indicators obtained with coarser zoning schemes and indicators obtained with this baseline.

- **orientation**: adjacent spaces with similar orientation are aggregated into one zone. The orientation of each space is represented by the average of the normal vectors of outdoor-facing windows in the space, weighted by window area. Two spaces are considered to have similar orientation if the angle between the two corresponding orientation vectors is no more than 45 degrees. Spaces without windows are considered as a distinct group (core spaces).

- **space type**: adjacent spaces with the same space type are aggregated into one zone. Here, only two space types are considered: living spaces and other spaces with lower internal gains and heating setpoints.

- **one per floor**: all floorplan spaces are aggregated into one zone.
Experiments

Three consecutive experiments are carried out:

- A: In a first step, all spaces are assumed to have the same internal load density. This way, only the geometrical aspects are considered.

- B: In a second step, a distinction is made between living areas on the one hand (living rooms, bedrooms) and circulation and storage spaces on the other hand, in which internal gains and temperature requirements are assumed to be lower, as summarized in Table 1.

- C: In a third step, a local sensitivity analysis is carried out to evaluate the impact of some parameters, for a selected subset of the floorplans considered in the first two steps.

Annual simulations are carried out with the IWEC weather data for Brussels. Some simulation parameters common to all experiments are summarized in Table 2. The simulated buildings roughly correspond to new European building standards. Concrete and exterior insulation are used for the exterior walls, brick for the interior walls and double glazing for the windows, which are also equipped with external shades. No additional internal mass is modelled, and interior walls between merged spaced are not compensated with internal mass. Airflow between zones is neglected.

A subset of floorplans is selected based on analysis of the results of the previous step. Clustering with K-means on the 2D space of differences in heating demand and maximum temperature is used to select floorplans with different difference levels. Parameters considered for sensitivity analysis are summarized in Table 3, whereby the default values are used in experiments A and B.

The simulation results are aggregated to calculate the three following indicators:

- specific heating energy demand in kWh/(m²a)
- maximum heat load density in W/m²
- maximum temperature in occupied spaces in °C

Results

Statistical indices of key floorplan characteristics are summarized in Table 4.

The values of specific heating energy demand for the original and simplified zoning configurations are compared in Figure 5. Dots further away from the diagonal indicate discrepancies introduced by simplified zoning. As shown in Figure 5, in the case of uniform internal loads, specific heat demand as calculated with one-zone-per-room zoning and with simplified zoning configurations are generally very close.
Table 3: Sensitivity analysis parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Low</th>
<th>Default</th>
<th>High</th>
<th>Unit</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heating setpoint living spaces</td>
<td>20.0</td>
<td>21.0</td>
<td>22.0</td>
<td>°C</td>
<td>Tsp</td>
</tr>
<tr>
<td>U-value of interior walls</td>
<td>0.5</td>
<td>1.0</td>
<td>2.0</td>
<td>W/(m² K)</td>
<td>Uiw</td>
</tr>
<tr>
<td>Window g-value</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>g</td>
<td>g</td>
</tr>
<tr>
<td>U-value of interior doors</td>
<td>2.0</td>
<td>3.0</td>
<td>4.0</td>
<td>W/(m² K)</td>
<td>Ud</td>
</tr>
<tr>
<td>Max electric equipment gain</td>
<td>2.0</td>
<td>3.0</td>
<td>5.0</td>
<td>W/m²</td>
<td>Qeq</td>
</tr>
<tr>
<td>External shade transmittance</td>
<td>0.2</td>
<td>0.4</td>
<td>0.8</td>
<td>-</td>
<td>ST</td>
</tr>
</tbody>
</table>

Table 4: Floorplan statistics: number of zones ($n_z$), floor area ($A$) and window-to-wall ratio ($WWR$). 25%, 50% and 75% refer to percentile values.

<table>
<thead>
<tr>
<th>$n_z$</th>
<th>$A$ in m²</th>
<th>WWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>8.8</td>
<td>73.5</td>
</tr>
<tr>
<td>min</td>
<td>5.0</td>
<td>50.1</td>
</tr>
<tr>
<td>25%</td>
<td>7.0</td>
<td>58.3</td>
</tr>
<tr>
<td>50%</td>
<td>8.0</td>
<td>68.9</td>
</tr>
<tr>
<td>75%</td>
<td>11.0</td>
<td>80.6</td>
</tr>
<tr>
<td>max</td>
<td>22.0</td>
<td>167.5</td>
</tr>
</tbody>
</table>

With a similar scatter plot for maximum temperature, Figure 6 shows that the impact on maximum space temperature is more significant, and generally corresponds to an underestimation of maximum temperature with simpler zoning configurations.

Next, differences between models with simplified and original (one-zone-per-room) zoning are calculated, as relative differences in % for energy demand and heat load, and as temperature differences in K for maximum temperature. Figures 7, 8 and 9 show the frequency distributions of the corresponding differences for the three output indicators, for the three zoning schemes and the two cases A and B.

As shown in Figure 7, the coarsest one-zone-per-floor zoning induces differences in heating energy demand between -9% and +9%. The distribution of difference values is narrower with homogeneous internal loads and setpoints (experiment A), whereas it is flatter and shows more outliers with differentiated internal loads (experiment B). In a vast majority cases, the deviation in annual energy remains between -5% and 5%, so that the loss of accuracy is likely to be acceptable in many applications.

The distributions of differences in maximum heat load (Figure 8 shows a similar picture. The difference
in maximum heat load seldom exceeds 4%. With homogeneous space parameters (experiment A), it is mostly negative, all the more so with the coarse one-zone-per-floor zoning. With differentiated space parameters, differences arise in both directions and reach higher absolute values.

As shown in Figure 9, the distributions of differences in maximum temperature have a broader shape. Underestimation is more frequent than overestimation, and it also reaches larger values. For the orientation-based zoning scheme, which seems to perform best, the difference remains between -2 °C and +1 °C. The results for the two other zoning schemes, with underestimation beyond -1 °C in about 25% of cases, do not seem reliable enough for most applications.

Sensitivity analysis results are shown in Figure 10 (maximum temperature) and Figure 11 (heating demand). The impact of various parameters depends on the output indicator: shade transmittance has the strongest impact on the difference in maximum temperature (Figure 10), with higher transmittance leading
to more spatial variations and larger deviations. In both cases, parameter variations have only a limited impact on the ranking of floorplans in terms of differences from the baseline. Internal gain density and heating temperature setpoint in living areas have the strongest impact on the difference in heating energy demand.

Discussion

By showing both variability and regularity with different floorplans, the results presented in the previous section justify the approach of investigating the impact of simplifications on a higher number of examples. Different levels of variability are shown depending on the output indicator and zoning scheme. Estimated probability distributions of differences in calculated indicators compared to a baseline fine-grained model are provided. Such distributions may be valuable for simulation practitioners. They could lead to caution in certain cases (e.g. when estimating maximum temperature, which does not appear to be a robust indicator with respect to zoning simplification) and confidence in others (e.g. when calculating heat loads).

Still, these distributions are empirical and cannot be assumed for simulation models deviating from the simple structure investigated here. Limitations of the present study are related to limitations of the used floorplan dataset on the one hand, and to assumptions underlying the simulation models on the other hand. Given the two-dimensional character of the underlying floorplan dataset, three-dimensional aspects of zoning as they may appear in buildings with varied floorplans have been ignored. Different results should be expected in the case of storeys for which the assumption of adiabatic floors and ceilings does not hold, and all the more with features introducing additional heterogeneity such as terraced roofs or skylights. Similarly, the consideration of material properties is limited, as the dataset did not include any information about them. However, the sensitivity analysis showed interrelations between some of these aspects, simulation zoning and simulation results.

Limitations related to the simulation models include simplified assumptions in terms of ventilation, lighting and shading. The assumption of known air change rates as a function of floor area arguably leads to more homogeneous results than natural ventilation as modelled for instance with airflow networks. Shading by neighbouring buildings or vegetation as well as shading from balconies may also lead to spatial differences not accounted for in the present study. Finally, the impact of daylighting has also been neglected. Neglecting such sources of spatial variability may result in some underestimation of the deviations introduced by zoning simplifications, which may be assessed in further investigations.

Finally, grouping rooms in more or less coarse simulation zones is only one of many simplifications which are done on a regular basis in dynamic thermal simulation. The present study may call for similar investigations into the effect of other simplifications, for instance in terms of building geometry and internal load schedules, which in turn can be assumed to interact with zoning and may require a more global sensitivity analysis.

Conclusion

This contribution has proposed a systematic investigation of the errors due to simplified as opposed to detailed simulation zoning. As opposed to previous investigations limited to one or a few building geometries, hundreds of distinct geometries from a large floorplan dataset are used. The impact of zoning simplifications is shown to vary widely with the floorplan on the one hand, and on the considered output indicator on the other hand. The difference in annual energy demand with a one-zone-per-floor model compared to a baseline one-zone-per-room model does not exceed 5% in most floorplans with the selected simulation parameters. On the other hand, deviations in maximum space temperatures frequently exceeding 1 K are obtained with the same parameters. A simple local sensitivity analysis carried out with selected floorplans gives first insights into the respective impacts of variations in geometry and simulation parameters. In future work, deeper analyses may be carried out with global sensitivity analysis methods.

References


