Review and critical analysis of early-design phase evaluation metrics for the solar potential of neighborhood designs

Emilie Nault, Giuseppe Peronato, Emmanuel Rey, Marilyne Andersen

School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Abstract

The energy performance of a building is strongly influenced by its level of solar exposure, in turn affected by the climate, built context, and building morphological characteristics. Since these are typically fixed at the early-design phase, performance assessment methods based on solar considerations at the urban scale are essential to support early decision-making. As the adaptation of the well-developed building performance simulation methods to the urban scale lead to complexity issues, it is of interest to verify whether simpler metrics can act as performance indicators, as is often done at the building level with quantities such as form factor.

This paper first presents a review of existing methods that aim at evaluating aspects such as passive heating and photovoltaic potential. We classify each evaluation metric based on its level of complexity, before choosing a representative subset to be further examined. The goal is to test whether simple metrics fulfill their role as performance indicators when applied to neighborhoods. To do so, the selected metrics, including geometrical parameters (e.g. compactness) and solar exposure levels (e.g. annual irradiation), are compared with simulation results (e.g. heating need), taken as reference values. This is done for two sets of designs: six distinct typologies and a large number of variants from a base case typography.

Results, which show various levels of correlation between simple metrics...
and the reference values, highlight the limitations of using any one metric and indicate a need to revise the definition and evaluation metrics of various performance criteria for neighborhood designs.

**Keywords:** neighborhood scale, early-design phase, performance metrics, passive/active solar potential, simulation, correlation

1. **Introduction**

The level of solar exposure of buildings influences their heating, cooling and artificial lighting needs. Key elements affecting solar exposure can be divided into two main categories: (i) site-specific parameters, such as climate and existing obstructions; and (ii) design-specific parameters, such as building height and orientation. While site-specific parameters are imposed by the project’s inherent characteristics, design-specific parameters are defined by the designer during the early design phase. At the building scale, these decisions can be made in a rather straightforward way, based on designers’ intuition, experience, and widely available tools and methods, ranging from simple early-design phase indicators to complex detailed-phase simulation software. However, at a larger scale such as a neighborhood, the interaction between buildings blurs the distinction between site- and design-specific parameters and increase the complexity of the decision-making process. Despite recent developments in solar performance assessment methods for design-support, few are targeting the early-design phase of neighborhood projects. Moreover, from what we found in the literature, there exist no clear and standard definition and evaluation of the solar potential of a neighborhood.

In this paper, we are specifically concerned with the early-design phase of new neighborhood projects, when many decisions regarding building massing and configuration are to be made, yet little or no information is known about function, materials, etc. This is typically the moment when practitioners deal with multiple scenarios (or design variants), each having distinct advantages and disadvantages regarding social, environmental, financial and other aspects. From this standpoint, we advocate the need for a simple and rapid assessment method of passive and active solar potential, which is comprehensive and reflective of potential energy savings and production. This method should allow the ranking of design alternatives for a certain performance criterion based on a quantifier (or metric), thus providing sound
energy- and comfort-based guidance.
Considering this context, this paper attempts to address the following issues:

• Identify the existing evaluation methods of the performance of neighborhood designs and categorize them based on their metric\(^1\) (performance quantifier)

• Assess a subset of these metrics by applying them to case studies and comparing the results between each method

• Assess the validity of the selected metrics for comparative assessment of the energy/solar performance of early-design phase neighborhood alternatives

2. State of the Art

In this section, we present an extensive literature review of studies evaluating urban-scale designs using specific evaluation metrics for each performance criteria they consider. A performance criterion relates to an aspect of building performance, while an evaluation metric is a quantitative measure expected to represent a certain performance criterion. For example, daylight potential as a performance criterion is measured by various metrics such as the daylight factor, preferable sky window or daylight autonomy.

In Table 1, we summarize the relevant information from a series of selected papers by categorizing the performance criteria (columns) as:

**Passive:** related to the passive performance of buildings, e.g. potential to benefit from passive solar heating

**Daylight:** representing the potential to benefit from natural daylight

**Active:** expressing the potential energy production by active solar systems (photovoltaic and solar thermal)

**Energy:** representing the needs for heating, cooling and lighting

\(^1\)The term ‘metric’ is used to refer to a quantity that is a potential performance indicator.
Performance criteria

<table>
<thead>
<tr>
<th>Evaluation metric</th>
<th>Passive</th>
<th>Daylight</th>
<th>Active</th>
<th>Energy</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Geometry-based</td>
<td>40, 28</td>
<td>35, 28</td>
<td>23</td>
<td>39</td>
<td>[19, 28, 23] (urban phenomena e.g. heat island); [40] (daylight in streets); [28] (land usage)</td>
</tr>
<tr>
<td>2. External solar- and geometry-based</td>
<td>1, 3, 33, 49, 12, 38, 6, 7, 28, 23</td>
<td>51, 5</td>
<td>22, 18</td>
<td>50, 33, 49, 40</td>
<td>(outdoor thermal comfort); [19] (solar access); [40] (daylight in streets)</td>
</tr>
<tr>
<td>3. Full climate- and geometry-based</td>
<td>3, 23</td>
<td>22, 16</td>
<td>12, 38, 13, 21</td>
<td>21, 13</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: State-of-the-art (non-exhaustive) in (solar) performance evaluation at the urban scale, classified according to the type of evaluation metric used.

**Other**: falling outside the previous themes, e.g. outdoor thermal comfort

We also categorize the evaluation metrics (rows) as being:

1. **Geometry-based**: metrics computed from the morphology of the buildings, based uniquely on the 3D model. Examples include the surface-to-volume ratio and the plot ratio.

2. **External solar- and geometry-based**: metrics computed from the level of solar exposure of external surfaces expressed in terms of irradiation (kWh/m²) or illuminance (klux), thus taking into account the interaction of buildings and their geometry. Examples include the average facade irradiation over the heating season, and the percentage of surface receiving an irradiation or illuminance level above a certain threshold.

3. **Full climate- and geometry-based**: metrics which are obtained through a more complex and computationally expensive yet conventional simulation, accounting for the climate and geometry in more de-
tail. Examples include the spatial daylight autonomy and the energy need for space heating.

For all performance criteria (columns), the majority of the metrics fall within the second category (external solar- and geometry-based metrics). Most of these methods are based on irradiation or illuminance levels, computed over a specific surface and period of time, e.g. over winter and summer to assess passive performance and outdoor thermal comfort [33, 49]. Some methods include threshold values [16, 38]. One example of the latter is the method by Compagnon [6, 7], which uses irradiation-based and illuminance-based threshold values to assess passive thermal, daylight and active (photovoltaic and solar thermal) potential. In [7], the method, applied to the current urban form of an area of Fribourg (Switzerland) and four hypothetical distinct designs, showed the percentage of adequate surface for each performance criterion. This method was also adopted in other studies [5, 29] including Swiss [31] and European [6] projects. However, to the best of our knowledge, no validation of the method’s results has been conducted, e.g. against actual daylight and energy consumption/production values.

Within the geometry-based metrics (category 1), studies combine various metrics including the passive space [39], visible sky area, south and west facade percentage [28], plot ratio, aspect ratio [23] and the more common surface-to-volume ratio [10] and its variants (e.g. relative compactness) [28, 37, 23], representative of building compactness. Despite widespread usage of compactness metrics as indicators of energy performance, their validity has been put in question by the work of Depecker et al. [10]. They found a strong correlation between the building’s shape coefficient and its energy consumption for the climate of Paris, but a much weaker correlation for the slightly milder climate of Carpentras (in southern France).

More accurate are the third category of metrics, which employ full simulations to compute conventional values such as heating and cooling needs [13], which are effectively well-acknowledged performance indicators. However, running such simulations at the neighborhood scale is a computationally expensive process requiring information not available at an early design stage (e.g. materials, occupation). This requires many assumptions and simplifications. As Beckers and Rodriguez note, “most programs are oriented to analysis and not to design. Thus, they are mainly used at the final step of the project, when the principal ideas are already defined, and all their possibilities
are not apprehended” [4, p. 475].

Independently of the metric category, we observe a disparity in the metrics used for some performance criteria. For the active solar potential, for instance, we observe not only a variety of metrics, but also a divergence within a single metric’s application. When a threshold is used, its value differs between countries and sources, e.g. the acceptable ratio of available to maximum irradiation for an optimally inclined and oriented surface ranges from 55% to 80% [9]. Assessment methods also differ in the number of metrics, as in the case of solar energy potential assessments based on solar maps, which offer from one output (i.e. irradiation levels) to multiple outputs (irradiation levels, solar systems output, categorization of suitable area, etc.) [17].

In addition to looking at the metric(s) used to quantify each performance criterion, it is of interest to examine how many and which criteria are considered. Many publications deal with a unique and specific criterion, for instance related to the daylight performance [51, 35], or the photovoltaic (PV) and/or solar thermal (ST) potential [22, 18, 9, 16]. Other studies also evaluate a unique metric, but one which is meant to represent the performance in a broader sense: Otis [33] and van Esch et al. [49] measure winter and summer solar exposure in terms of radiation, with this measure being linked to the passive performance of buildings and the thermal comfort of outdoor spaces.

Our main conclusions regarding the overall literature review are that both the methods and outcomes of most papers remain specific to the cases they studied and are hardly generalizable. We also observe a lack of validation when simple metrics (falling within categories 1 and 2) are adopted, to verify as to whether or not their interpretation truly reflect an improved performance. Approaches which use full climate- and geometry-based metrics (category 3) do not have this issue as these metrics are intrinsically acknowledged as performance indicators.

Based on the classification in Table 1, we have selected metrics to span all main performance criteria (all columns, except ‘Other’) and metric types (rows). The selection, detailed in the next section, was based on how widespread, adequate, and simple to implement the metrics were for evaluating and comparing early-design phase neighborhood designs. To infer the validity of each metric as a ‘true’ performance indicator, we make the assumption that the full climate- and geometry-based metrics (category 3) are to be considered as the reference values. Thus, results obtained by applying metrics from the first
2 categories should reflect the results obtained from the 3rd category. While the latter metrics are not adequate for early-design phase assessments due to the factors introduced earlier (computational cost, required information not yet known), they are valuable for proving or disproving the validity of alternative metrics, which are simpler and easier to calculate. This approach is based on the thinking that an increase in the performance as quantified by any metric should go hand in hand with energy consumption reductions or energy production increases.

3. Methodology

3.1. Main steps

Our workflow is illustrated in Fig. 1. Considering the 3D model of a neighborhood design, we go through the steps from left to right; the first and simplest phase consists in extracting the geometry-based metrics. The second step is to run an irradiation simulation to obtain the second category of metrics. Finally, the reference metrics are obtained either through full simulations or further processing of results from the second phase.

Figure 1: Workflow adopted. Results obtained for the geometry-based and external solar- and geometry-based metrics (light gray boxes to the left) are compared with the full climate- and geometry-based metrics (dark gray box to the right) taken as reference values. The level of complexity, computational cost, and accuracy (accounting for some assumptions) increases from left to right. Best viewed in color.

This procedure was applied on various neighborhood designs (initial 3D
models in the workflow), introduced in the next section. The actual metrics for each category and performance criterion are presented in section 3.3.

3.2. Case studies

The series of neighborhood designs used as case studies for this paper are the outcomes of two projects: a bachelor architecture studio and summer workshop at the EPFL [43, 44], and a joint master’s thesis between EPFL and IUAV [36].

The first project is a comparative assessment of six urban visions based on the Waldstadt area in the city of Bern, Switzerland [45]. The schematic master plans of the six projects are shown in Fig. 2. The designs differ in typology, dominant building orientation, density and building height. Due to the time required for computer-simulations, only a representative sample of each design was simulated, taking into account the surrounding context (terrain and buildings’ reflectance/shading effect).

The second project is a parametric study that generated multiple design variants based on the master plan [34] for a neighborhood in the city of Yverdon-les-Bains (YLB) in Switzerland. Figure 3 shows a schematic shadow plan of the YLB base case [36]. Geometrical modifications were done by varying the height of buildings within the range of values prescribed by the...
master plan, and by setting their depth and setback to values expected to be representative of possible design choices. These variations were applied on each closed-block, giving 768 scenarios. To reduce simulation time and the extensive number of design variants, only buildings G, H and I were simulated, taking into account their surrounding context. The variations were applied equally on all three buildings, i.e. within each scenario all buildings have the same depth, setback, and height value, with the exception of additional stories located on the northern facade of buildings G and H. Despite these uniform parameter values, we obtain three distinct building sizes due to the initial outer perimeter being different for each block. This can be better understood by looking at Fig. 4, which illustrates the annual irradiation map of two example scenarios.

Figure 3: Schematic shadow plan of the Yverdon-les-Bains (YLB) base case. Due to practical considerations, only buildings G, H and I were simulated, taking into account their surroundings [36].

Most of the outputs associated with each phase of the workflow (Fig. 1) were obtained within these projects. We here make use of the data acquired through both projects and subsequently complemented by analyzing it, to fulfill the goal of this study. We will refer to the first and second project respectively as BE and YLB, which are acronyms for their location. While
the first project comprises six designs with different building typologies, the second set is composed of a much larger number of designs created from small variations applied to a single base case. Each set thus represents a distinct situation found in practice. For instance, the six urban visions could emanate from a design exploration phase or from an urban design competition, while the extensive series of variants for the YLB case study would come up at a subsequent phase, after fixing the building typology.

The 3D models of the designs were made in the Rhinoceros environment [25], in combination with the Rhino-integrated graphical algorithmic editor Grasshopper [26] for YLB. Simulations were conducted via various Rhino and Grasshopper plug-ins, as detailed in the next sections. The climate files used in the simulations were obtained from Meteonorm [27] for both cities (Bern and Yverdon-les-Bains).

3.3. Performance criteria and metrics

Following our goal of comparing the 2 first categories of (simpler) metrics to the reference metrics as mentioned earlier, we re-arranged the performance criteria into 2 groups: passive and active solar potential. Table 2 lists the selected evaluation metrics for each criterion.

3.3.1. Selected geometry-based metrics (cat. 1)

Geometrical or morphological parameters are often used in the literature to describe building shape and configuration. Many can also be computed for an ensemble of buildings. The geometry-based metrics examined in this study are:

• Density as expressed by (i) the plot ratio: the ratio between the total floor area and the land parcel, and by (ii) the site coverage: the ratio
Performance criteria

1. Geometry-based metrics
2. External solar- and geometry-based metrics
3. Full climate- and geometry-based metrics (Reference)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat gains</td>
<td>Envelope area with irrad. &gt; T_{heat gains} (%)</td>
<td>Annual irradiation per envelop or floor area (kWh/m²)</td>
<td>Annual heating need per floor area (kWh/m²)</td>
</tr>
<tr>
<td>Heat avoidance</td>
<td>Envelope area with irrad. &lt; T_{heat avoidance} (%)</td>
<td></td>
<td>Annual cooling need per floor area (kWh/m²)</td>
</tr>
<tr>
<td>Daylight</td>
<td>Facade area with illumin. &gt; T_{daylight} (%)</td>
<td></td>
<td>Spatial daylight autonomy (%)</td>
</tr>
<tr>
<td>PV-F</td>
<td>Facade area with irrad. &gt; T_{PV-F} (%)</td>
<td>Annual roof/façade irradiation per floor area (kWh/m²)</td>
<td>Annual estimated energy production on roof/façade per floor area (kWh/m²)</td>
</tr>
<tr>
<td>ST-F</td>
<td>Facade area with irrad. &gt; T_{ST-F} (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PV-R</td>
<td>Roof area with irrad. &gt; T_{PV-R} (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST-R</td>
<td>Roof area with irrad. &gt; T_{ST-R} (%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Metrics compared for each performance criterion. PV and ST: photovoltaic panels and solar thermal collectors installed on facades (-F) or roofs (-R).

between the total building footprint (ground floor) and the land parcel. These are common urban morphological parameters [23], linked to heat losses and solar and daylight availability [48].

- Compactness as expressed by the surface-to-volume ratio: the ratio between the envelope area and its enclosed volume (or floor area). The surface-to-volume ratio is indicative of heat losses [40]. The smaller the value, the more compact the building, and the lower the heat losses due to a smaller exposed surface area. In northern climates, this ratio should generally be minimized to avoid heat losses which dominate over heat gains in winter, and to limit heat gains (by having a smaller collecting surface) in summer [28].

- Passive area as expressed by the passive space ratio: the ratio between the passive floor area (surface within 6 m of a facade) and the total floor area. This value, adapted from [39], quantifies the potential of a space to use daylight, sunlight and natural ventilation. Ratti et al. [39] found it to be a better indicator than the surface-to-volume ratio for total energy consumption (heating, cooling, ventilation and lighting).

- Glazing proportion as expressed by the window-to-floor ratio: the ratio between the total window area and the total floor area. The window-to-floor ratio is used instead of the more common window-to-wall ratio
which was kept constant in the case studies. Glazing proportion characterizes heat gains and losses, natural ventilation, daylight, as well as other aspects not treated in this study such as glare [32].

These metrics are ordered as they would appear during the design process: density objectives are typically fixed at the beginning of the urban planning phase, while glazing proportion would come up in the building design stage.

All metrics were computed for each scenario (design variant). For example, for one specific scenario, we computed a unique compactness as an overall neighborhood value.

### 3.3.2. Selected external solar- and geometry-based metrics (cat. 2)

For this second category of metrics, we chose a threshold-based method based on [6, 7], in which a solar exposure threshold is associated to each performance criteria in the form of an illuminance level (for the daylighting criterion) or irradiation level (for the heat gains/avoidance and active solar potential criteria). The metric is the percentage of exposed surface area receiving an irradiation or illuminance level above (or below) the criteria-specific threshold. These thresholds are given in Table 3. The workflow for obtaining the solar exposure level and the surface percentage values are described in [30] and [36] for the BE and YLB case study respectively.

<table>
<thead>
<tr>
<th>Performance criteria</th>
<th>Threshold symbol</th>
<th>Threshold value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heat gains</td>
<td>$I_{\text{heat gains}}$</td>
<td>172.7 (140 for YLB) kWh/m$^2$ over heating period</td>
</tr>
<tr>
<td>Heat avoidance</td>
<td>$I_{\text{heat avoidance}}$</td>
<td>46.8 kWh/m$^2$ over non-heating period (upper limit)</td>
</tr>
<tr>
<td>Daylight</td>
<td>$I_{\text{daylight}}$</td>
<td>10 klx over working hours</td>
</tr>
<tr>
<td>PV-F</td>
<td>$I_{\text{PV-F}}$</td>
<td>800 kWh/m$^2$ over year</td>
</tr>
<tr>
<td>ST-F</td>
<td>$I_{\text{ST-F}}$</td>
<td>400 kWh/m$^2$ over year</td>
</tr>
<tr>
<td>PV-R</td>
<td>$I_{\text{PV-R}}$</td>
<td>1000 kWh/m$^2$ over year</td>
</tr>
<tr>
<td>ST-R</td>
<td>$I_{\text{ST-R}}$</td>
<td>600 kWh/m$^2$ over year</td>
</tr>
</tbody>
</table>

Table 3: Criterion-specific threshold value. PV and ST: photovoltaic panels and solar thermal collectors on facades (-F) or roofs (-R).

In addition to the threshold-based surface percentage values, the following metrics were computed: the annual irradiation received on facades, roofs,
and all exposed surfaces (facades and roofs combined) normalized against the envelope area (total exposed surface) and against the floor area.

3.3.3. Selected full climate- and geometry-based metrics (reference)

This third set combines conventional performance metrics taken here as a reference and obtained, for the heat gains/avoidance and daylight criteria, through more complex simulations requiring additional information about the designs. As for the previous metrics, the values are obtained per neighborhood. Due to the particularly high computational cost of the daylight simulation, only 32 variants of the YLB case study were simulated for this metric, representing the extreme values for every design parameter varied when generating the scenarios.

The heating need per floor area (kWh/m\(^2\)) was used as the indicator of passive heat gains potential. The annual heating need (kWh) for each building was first obtained for the six virtual projects of BE via UMI, a Rhino-based urban modeling interface [42], and for the YLB scenarios via Viper, a DIVA-for-Grasshopper component [47] also based on EnergyPlus [8]. These values were then summed and divided by the total floor area (all buildings, all levels). The UMI and Viper simulation settings are given in Table 4. More details regarding the use of Viper, which resulted in a more complex procedure than UMI, can be found in [36]. We refer to this metric simply as ‘heating need’, implying that it is the value normalized by the floor area (kWh/m\(^2\)).

The cooling need (per floor area), metric representing the passive heat avoidance, was obtained in the same way as the heating need.

For the daylight reference metric, we use the spatial daylight autonomy (sDA). The sDA represents the percentage of workplane-height space which receives at least 300 lux over 50% of the occupation hours over the year [15]. It was computed for the BE case via UD [11], a plug-in for Rhino which converts exterior radiation levels (obtained via Radiance/Daysim [20]) into interior illuminance distributions using an impulse response algorithm. For the YLB case, the sDA was computed via DIVA-for-Grasshopper which is directly based on Radiance/Daysim. Blinds are activated when direct sunlight exceeds 75 W/m\(^2\) and 1 klux for the BE and YLB case study respectively. Thus, while metrics in previous categories do not consider visual comfort, the sDA does take it into account through the integration of a dynamic blind system into the simulation.

For the active solar systems criterion, the metric builds upon the results
<table>
<thead>
<tr>
<th>Parameter</th>
<th>UMI settings</th>
<th>Viper settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window-to-wall ratio</td>
<td>0.5</td>
<td>~ 0.48</td>
</tr>
<tr>
<td>Building template</td>
<td>ASHRAE 90.1 Zone 5A</td>
<td>Custom</td>
</tr>
<tr>
<td>Building function</td>
<td>Office</td>
<td>Residential &amp; Office</td>
</tr>
<tr>
<td>Heating/cooling set point (setback)</td>
<td>Dual: 20(15)/26(30) °C</td>
<td>20/26 °C</td>
</tr>
<tr>
<td>Loads</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment</td>
<td>8 W/m²</td>
<td>6.59 W/m²</td>
</tr>
<tr>
<td>Lighting</td>
<td>11 W/m²</td>
<td>0 W/m²</td>
</tr>
<tr>
<td>Occupancy</td>
<td>0.045 people/m³</td>
<td>0.03 people/m³</td>
</tr>
<tr>
<td>Ventilation</td>
<td>0.0003 m³/sm²</td>
<td>0.0002 m³/sm²</td>
</tr>
<tr>
<td>Infiltration</td>
<td>0.5 ach</td>
<td>0.0002 m³/s</td>
</tr>
<tr>
<td>Material</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wall surfaces</td>
<td>Brick with insulation</td>
<td>Heavyweight concrete</td>
</tr>
<tr>
<td></td>
<td>(~ 18 cm thick)</td>
<td>with insulation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(~ 41 cm thick)</td>
</tr>
<tr>
<td></td>
<td>0.39 W/m²K</td>
<td>0.17 W/m²K</td>
</tr>
<tr>
<td>Glazed surfaces</td>
<td>Argon-filled double glazed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>U = 1 W/m²K</td>
<td></td>
</tr>
<tr>
<td></td>
<td>g = 58 %</td>
<td></td>
</tr>
<tr>
<td>Solar protections</td>
<td>Venetian blinds</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>(activated when I &gt; 75 W/m² - program default)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Settings for the heating energy usage simulations. The UMI [42] plug-in was used for BE, while Viper (DIVA-for-Grasshopper component [47]) was used for YLB.

from the threshold metrics (external solar- and geometry-based set) by making use of the surfaces, simulated as a series of evaluation points or nodes, adequate for each system, based on the same threshold value given in Table 3. As such, this metric does not require any additional simulation. However, we consider that it provides more trustworthy results, as the efficiency of each system is considered as well as the actual irradiation of each node which exceeds the relevant threshold. An estimate of the energy production by photovoltaic panels (PV) and solar thermal collectors (ST) on facades (-F) and roofs (-R) was obtained through the following simple calculation:
\[ E_{AS} = \sum_{n=1}^{N_{AS}} \frac{I(n) \times NA \times \eta_{AS}}{A_f} \text{[kWh/m}^2]\]  

(1)

Where

- \( AS \) subscript: active system in question (PV-F, PV-R, ST-F, ST-R)
- \( N_{AS} \): nodes for which the irradiation threshold is achieved for surface and system \( S_{AS} \)
- \( I(n) \): annual irradiation of node \( n \) (kWh)
- \( NA \): node area (m\(^2\))
- \( \eta_{AS} \): efficiency of system, 0.15 for PV, 0.70 for ST
- \( A_f \): total floor area (m\(^2\))

4. Results

We now present, for each performance criterion, the comparison of the geometry-based and external solar- and geometry-based metrics with the reference metrics, respectively in section 4.1 and 4.2. For conciseness we present only the main comparisons of interest.

4.1. Geometry-based against reference metrics

Geometry-based metrics are the simplest and fastest to compute, as no simulation is required and they can even be calculated by hand. However, are these metrics sufficient as urban performance indicators? Since they are independent of climate and thus of solar exposure levels, they are likely to show inaccuracies, at least in some cases.

4.1.1. Heat gains and heat avoidance

Figure 5 shows the heating need against each geometry-based metric for (a) BE and (b) YLB, along with the coefficient of determination (R\(^2\)) representing the goodness of fit for a linear regression model \[2\]. For the density metrics, only the plot ratio is presented, as the site coverage was found to show the same behavior.

For BE (graphs (a)), the strongest indicator of the heating need, i.e. with the R\(^2\) closest to 1, is the window-to-floor ratio, with higher needs corresponding to larger glazed areas due to more heat losses. Also powerful is the surface-to-volume ratio; as expected, the more compact variants (lower surface-to-volume ratios) have a lower heating need. Indeed, more compact
is the design, smaller is the exposed surface area and thus lower are the heat losses. The other geometry-based metrics are weaker indicators, at least when considered independently as done here.

![Graphs showing heating need against geometry-based metrics for BE and YLB cases.]

Figure 5: Heating need (reference metric) against geometry-based metrics for (a) the BE and (b) YLB case study.

Results obtained with the 768 scenarios of YLB (graphs (b)) also show a strong correlation of the heating need with the surface-to-volume and window-to-floor ratio, and a slightly weaker relationship with the plot and passive space ratio. The trend is particularly clear when looking at the datasets for 2-story and 3-story scenarios (dots vs crosses). The correlation observed across all four geometry-based metrics is caused by the dependency between these values that arose from the way the variants were parametrically generated. For example, when increasing the building depth parameter, the surface-to-volume and window-to-floor ratio followed, due to the parametric modeling set-up in Grasshopper. This can be proved by a Principal Component Analysis (PCA), which consists in projecting a set of data to provide a sequence of the best linear approximations to this data, thus lowering the dimension of the data space. Finding the best linear approximation amounts to maximizing the variance of the projected data [14, 46]. Results of our PCA analysis over the set of geometry-based metrics indicated that
88% of the variance is explained by the first principal component, which has coefficients showing that the surface-to-volume, passive space, and window-to-floor ratio are positively correlated to each other, as reflected by the same positive trend in the corresponding graphs. These 3 metrics are also inversely correlated to the plot ratio, as confirmed by the negative trend of the plot ratio graph. The consequence of this interdependency is that if any one of the geometry-based metrics correlates with the heating need (positively or negatively), as we would expect to see for the surface-to-volume ratio for instance, all other metrics will too.

The slope of each relationship indicates whether the geometry-based metric should be minimized or maximized to achieve a better performance. The positive slope for the passive space ratio is contradictory to the concept as used in [39], where non-passive zones consume more than passive ones. In our case, scenarios with values of 1 (100% of the space is passive) show the highest heating need. However, it must be noted that total energy consumption for heating, cooling, ventilation and lighting is used in [39].

When looking beyond the overall pattern of each graph, we observe clusters showing different or no trends, e.g. the group of red circled dots. Each cluster consists of scenarios with identical block depth (i.e. equivalent surface-to-volume and passive space ratio), but distinct setbacks, window-to-floor and plot ratio. The overall positive correlation with the window-to-floor ratio is not observed within the group of red points; the highest heating need is associated with a window-to-floor ratio of about 0.32, which is not the maximum value for this parameter. At constant surface-to-volume ratio, the influencing factor appears to be the higher heat gains provided by larger glazing area, as opposed to the general trend of the window-to-floor ratio graph. However, this parameter is not sufficient to clearly rank the design alternatives, as we still observe clusters of points with identical window-to-floor but different heating need values.

Results for the cooling need are not shown here. They were found, for BE, to present no trend except for the plot ratio ($R^2 = 0.58$) and, for YLB, show patterns very similar to the heating need. For the latter case, this means that both the heating and cooling needs can be minimized through the same geometrical features, e.g. a more compact design and a lower window-to-floor ratio.
4.1.2. Daylight

The spatial daylight autonomy reference metric showed trends which were very similar to the heating need for BE (graphs not shown here). $R^2$ values of 0.70 and 0.83 were obtained for the surface-to-volume and window-to-floor ratio respectively. As for YLB (Fig. 6), very little information can be drawn from the 32 data points due to the discontinuity in the data that comes from the scenarios’ characteristics (minimum and maximum parameter values). While results show what could be a pattern similar to the heating need, the actual indicative trend should be interpreted inversely: the sDA performance is to be maximized, not minimized like the heating need. This means that the passive space ratio now presents an overall behavior in accordance with its definition; buildings with only passive spaces (value = 1) present higher sDA values.

4.1.3. Active systems

No correlation was found between the production by active systems (reference metric) and the geometry-based metrics for BE, as well as for the roof-mounted systems for YLB (results not shown here). However, results for the facade-mounted systems were found to be similar to the heating need results (Fig. 5b) for both PV and ST. Thus, if comparing to the surface-to-volume ratio, as the exposed surface area expands for the same enclosed volume, more irradiation for energy production is collected per floor area. At constant surface-to-volume ratio, another metric should be identified to rank the design options.

4.1.4. Summary of findings

Not surprisingly, we have found that the main indicators of the heating and cooling needs are the surface-to-volume and window-to-floor ratio.
However, precautions should be taken when comparing scenarios: for very distinct designs, as in BE, a lower surface-to-volume and window-to-floor ratio is preferred; for very similar designs, as in YLB, a higher window-to-floor ratio within the lowest surface-to-volume ratio group is preferred. For the latter case study, selecting the lowest surface-to-volume ratio would in fact lead to a heating need 6.4% higher than the minimum value found within the variants.

For the spatial daylight autonomy, the passive space ratio shows the strongest correlation. This result is in line with the various rules-of-thumb defining the ratio between window-head-height and daylit zone depth. The general statement behind such rules is that “the depth of the daylit area usually lies between 1 and 2 times the size of the window-head-height” [41, p. 1017]. For energy production by facade-mounted systems, correlations are similar to the ones for the heating need for the YLB case study.

For all performance criteria, we would expect the observed correlations to be weaker in the case of dense neighborhood designs with high buildings or located in other climates. Such cases were not present in our dataset, but a side experiment was done to verify this hypothesis by varying the building height and spacing for project 2 of the BE case study. Results, presented in Fig. 7 for the climate of Bern and San Francisco, show a weaker correlation than what was previously observed between the heating need and the surface-to-volume and window-to-wall ratio.

Moreover, we must remember that for YLB, the correlations are affected by the inter-dependency of the 4 geometry-based metrics, as proven by the PCA. As such, it is important to note that it is practically impossible to detect causal relationships. For example, although we observe a strong correlation (high R² value) between energy production by PV-F and window-to-floor ratio, this relationship is clearly superficial and caused by a non-physically real phenomenon.

4.2. External solar- and geometry-based against reference metrics

We continue our analysis of early-design phase performance indicators by looking at the external solar- and geometry-based metrics, which are one complexity step higher than the geometry-based metrics, as they require irradiation simulations to be executed on a virtual 3D model of each scenario.
4.2.1. Heat gains and heat avoidance

We here present the results only for YLB, since for BE all results showed no correlation. Figure 8 shows the heating need against each metric in this category: the percentage of surface receiving an irradiation level above the heating threshold (left), and the annual irradiation per envelope area (center) and per floor area (right). In the first graph, the overall trend is opposite to what it ought to be considering that a higher surface percentage (x axis) should correspond to a better performance, but is instead associated here with a higher heating need (lower performance). However, the trend within scenarios of equal surface-to-volume ratio (clustered points) illustrates the ‘correct’ correlation.

The center graph also has a double behavior: a negative versus positive trend for the overall versus clustered data. More irradiation per envelope area leads to lower heating need as one would expect, except at constant surface-to-volume ratio, where lower exposure levels are preferred, likely due to the window-to-floor ratio as previously explained. The last graph shows more consistency, with the heating need increasing with the annual irradiation per floor area both within and across clusters. However, this trend is
Figure 8: Heating need (reference) against percentage of exposed surface exceeding the irradiation threshold (left), mean irradiation per envelope (center) and per floor area (right) for the YLB case study.

opposite what one would expect. This can be explained by the fact that more irradiation per floor area means more exposed surface area and less compact buildings: heat losses dominate over solar gains for this climate.

Each of the three metrics has a limitation; while the first (surface percentage) does not consider by how much the threshold is exceeded, the annual irradiation per envelope area does not take into account how much of this irradiation is transmitted to the interior, and the annual irradiation per floor area is overruled by the heat losses dictated by the surface-to-volume ratio, to which it is correlated.

Results for the cooling need were not obtained for the surface percentage metric, while they were found to be similar for the annual irradiation metrics.

4.2.2. Daylight

Figure 9 presents the spatial daylight autonomy (reference metric) against each of the three metrics: the percentage of surface receiving an illuminance level above the daylight threshold (left), and the annual facade irradiation per envelop area (center) and per floor area (right). The latter metric presents, for BE, the only correlation observed ($R^2 = 0.74$).
Figure 9: Spatial daylight autonomy (reference) against percentage of exposed surface exceeding the illuminance threshold (left), mean irradiation per envelope (center) and per floor area (right), for the (a) BE and (b) YLB case study.

4.2.3. Active systems

Figures 10 and 11 present, respectively for BE and YLB, the reference energy production metrics against the three external solar- and geometry-based metrics. For YLB, only the results for PV are presented, as very similar trends were obtained for ST.

For roof-mounted systems, the annual irradiation per floor area is the best indicator with $R^2$ values close to 1 for both systems and case studies. This conclusion holds for facade-mounted systems except in the case of PV-F for BE ($R^2 = 0.54$), where we observe higher correlations for the surface percentage ($R^2 = 0.66$) and annual irradiation per envelop area ($R^2 = 0.72$).
Figure 10: Energy production by active solar systems (reference) against percentage of exposed surface exceeding the corresponding irradiation threshold (left), mean irradiation per envelop (center) and per floor area (right), for photovoltaic (PV, top) and solar thermal (ST, bottom) systems installed on facades (-F) and roofs (-R). BE case study. Best viewed in color.

Figure 11: Energy production by PV (reference metrics) against percentage of exposed surface exceeding the corresponding irradiation threshold (left), mean irradiation per envelop (middle) and per floor area (right), for roof-mounted (-R, top) and facades-mounted (-F, bottom) systems. YLB case study.
4.2.4. Summary of findings

Within the 3 external solar- and geometry-based metrics compared in this section, the overall best indicator for all performance criteria appears to be the annual irradiation per floor area. However, it is important to note the direction of the trend; in the case of the heating need (right graph of Fig. 8), the irradiation metric should be minimized. This counter-intuitive result is caused by the dominating heat loss factor as previously explained. If that metric was erroneously used as an indicator to be maximized, i.e. if the scenario with the highest irradiation per floor area was selected as the ‘best’, the corresponding heating need would be as much as 65.7 % higher than the minimum value found within the variants.

For the other performance criteria, daylight and active systems, the irradiation metric should be maximized. This leads to conflicting goals which cannot be solved with simple calculations and require an integrated assessment approach, as discussed later.

5. Synthesis

In this study, we considered that an increase in the passive solar potential, as quantified by any metric such as the ones tested here, should go hand in hand with reductions of energy consumption. Similarly, an increase in the active solar potential should go together with higher production levels by PV and ST systems on roofs and facades. These correlations should be respected for both contexts of study: (i) when different typologies are compared, as for the six urban visions of the BE case study, and (ii) when comparing variants based on a fixed typology, as in the YLB case study.

Before highlighting the main outcome of this study, it is important to mention its limitations. The data acquired through simulation is subject to the intrinsic limitations of the software used, especially since some tools are still in development. However, this issue is mitigated by the comparative nature of the study, which makes relative values more important than absolute ones. The small number of data points for BE restricts the depth of the analysis that can be conducted on this dataset, and the conclusions that can be drawn from it. While the YLB case study offers a much larger dataset, it must be recalled that the variations between the scenarios are smaller. Results from the two datasets differ significantly due to distinct simulation settings, tools used, climate-files, and built context. For instance, the heating need is much larger for BE, as theses scenarios were modeled with less
insulation and with shading devices. This discrepancy in the results prevents us from merging both datasets, but also highlights the potential issues when comparing completely different designs. A robust performance assessment method is needed to address such issues.

We can merge all comparison results by visualizing, for each pair of metrics, the $R^2$ value averaged over both case studies. This is done in Fig. [12] for the (a) passive and (b) active performance criteria.

For the passive criteria, the surface-to-volume and window-to-floor ratio perform best as indicators of the heating need and spatial daylight autonomy. However, the relationship with each performance criterion must be well known. As was observed for the YLB dataset, this relationship was often two-fold; e.g. an overall positive trend with an inverse pattern within a cluster (points of equal surface-to-volume ratio). The irradiation per floor area metric showed to be complimentary to the geometry-based metrics, helping to further rank design variants within clusters. Thus, there are clear limitations associated to using only one metric. These limitations are further highlighted by the results of the test made with denser designs and higher buildings which showed much lower correlations between heating need and surface-to-volume and window-to-wall ratio, especially for a warmer climate.

No strong correlation is found for the heat avoidance criterion. For the active criteria, the strongest indicator is the annual irradiation per floor area.
The lowest correlations are found for the metrics involving thresholds (surface percentages) for both passive and active criteria.

It is also important to look beyond the $R^2$ value at the source of the data being compared and the shape of each graph. As was highlighted earlier through a PCA analysis, the interdependence between parameters can lead to ‘indirect’ correlations, which are hard to detect. Caution must also be used when drawing conclusions on the cause of the observed trend, as in the famous saying *correlation does not imply causation*.

It is difficult to establish a correlation that will hold for any dataset, e.g. based on a different climate or neighborhood typology. Due to this factor, and because of limitations associated with using any one metric, current work conducted by the authors aims at integrating both metrics (that were compared to the reference metrics) to capture the essential features of a neighborhood design, giving more accurate assessments of performance. This will be done through the following general steps:

- **Data acquisition**: enlarging the dataset of scenarios through parametric modeling of various base case designs
- **Mathematical model**: use a combination of metrics to train one or more mathematical model(s) able to predict the performance of a neighborhood for each performance criterion
- **Validation**: investigate the robustness and boundaries (e.g. climates, typologies) within which the mathematical model(s) remains valid

By merging various simple metrics together via machine learning techniques, we hope to overcome the limitations associated with using a unique metric, while making the assessment method more robust to various base case designs.

An additional challenge will be the combination of conflicting performance criteria, as highlighted previously. Weighting and multi-objective optimization techniques will be investigated to tackle this issue.

6. Conclusion

This paper aimed to determine whether simple neighborhood-scale and early-design phase metrics fulfill their role as performance indicators. A set of such metrics was selected for performance criteria classified as passive
(heat gains/avoidance, daylight) and active (photovoltaic, solar thermal) solar, and were compared to what we identified as reference metrics, i.e. values obtained through full simulations or further processing of the simple metrics results. Applied to various neighborhood designs, we examined the level of correlation between each pair of metrics compared, to see how reliable simple metrics could be in allowing a ranking of design alternatives in terms of each performance criterion.

Results show levels of correlation that are case-specific, with trends that are sometimes counter-intuitive, and relationships that can be fictitious. They highlight the risk, when using a unique metric, of incorrectly or incompletely ranking design alternatives. This risk is more or less significant depending on the context of the analysis, e.g. how do the design variants differ, and in which climate we are located. The assumed simplicity of using easily computable metrics as performance indicators can have significant drawbacks. Depending on the design alternatives to be compared (differing in their geometrical differences), using the surface-to-volume ratio or irradiation per envelope area as an optimization indicator can lead to heating need increases of 6% and 20% respectively, based on our YLB case study.

To overcome this risk, we emphasize the need to revise and refine the definition of early-design phase performance metrics for neighborhood designs. Current and upcoming work will attempt to address this gap by integrating multiple simple metrics into a metamodeling platform, with the aim of predicting performance in a more reliable way. This shall remove the need to run full simulations which are computationally expensive and not appropriate for the early-design phase.

7. Acknowledgement

This work was supported by the EPFL with additional support from the EuroTech Universities alliance (E. Nault). The authors would like to thank Jan Wienold and Parag Rastogi for their constructive feedback.

8. References


URL http://www-bcf.usc.edu/~rknowles/sol_env/sol_env.html


