DELFT UNIVERSITY OF TECHNOLOGY
FACULTY OF ARCHITECTURE AND THE BUILT ENVIRONMENT

MASTER OF SCIENCE GEOMATICS THESIS

SOLAR ANALYSIS ON BUILDINGS OF FAVELAS IN SÃO PAULO TO ESTIMATE PV POTENTIAL

DENIS GIANNELLI
DELFT, NOVEMBER 2021

DELFT UNIVERSITY OF TECHNOLOGY

FACULTY OF ARCHITECTURE AND THE BUILT ENVIRONMENT

MASTER OF SCIENCE GEOMATICS THESIS

SOLAR ANALYSIS ON BUILDINGS OF FAVELAS IN SÃO PAULO TO ESTIMATE PV POTENTIAL

DENIS GIANNELLI

DELFT, THE NETHERLANDS

2021

DENIS GIANNELLI

SOLAR ANALYSIS ON BUILDINGS OF FAVELAS IN SÃO PAULO TO ESTIMATE PV POTENTIAL

MSC GEOMATICS THESIS PRESENTED TO THE MENTORS AND GRADUATION TEAM TO OBTAIN THE TITLE OF MASTER OF SCIENCE.

1st MENTOR: GIORGIO AGUGIARO
2nd MENTOR: CAMILO LEÓN-SANCHEZ
CO-READER: OLINDO ISABELLA
DELEGATE: PETER KOORSTRA

DELFT, THE NETHERLANDS
NOVEMBER 2021

“Fortunately, among all works of art of mankind, cities are the masterpieces. And, fortunately, these are open and unfinished masterpieces.” Architect and Urbanist Professor Alexandre Delijaicov – translated by the author.

ACKNOWLEDGEMENTS

To the University of São Paulo, my everlasting alma mater, for opening my eyes to the complexity and challenges of the Brazilian society. I will always be grateful.

To REVOLUSOLAR, a non-profit organisation in Rio de Janeiro, for their outstanding hands-on experience with solar energy in favelas, and for the fruitful debate over this topic.

To Delft University of Technology, which embraced me during these two and half years, for all the technical knowledge and for expanding my horizons to an international environment.

To Giorgio Agugiaro and Camilo León-Sanchez, my MSc thesis mentors, for accepting the mentorship, supporting me with the research and transmitting the necessary knowledge to complete this academic and professional step.

To Olindo Isabella and Peter Koorstra, members of the graduation team, for the feedback and discussion during the graduation assessment.

To the colleagues Jérôme Kämpf (Idiap), Eric Duminil (HFT Stuttgart), Maximilian Sindram and Claus Nagel (Virtual City Systems), for the theoretical and technical support during the research.

ABSTRACT (EN)

GIANNELLI, Denis. Solar Analysis on Buildings of Favelas in São Paulo to Estimate PV Potential. MSc Geomatics thesis – Faculty of Architecture and the Built Environment, Delft University of Technology. Delft, the Netherlands, 2021.

In the Global South, large urban spaces resulted in the duality between the so-called 'formal' and 'informal' cities. It is the case of São Paulo, a twenty-two-million people metropolis and a financial hub in Latin America. Albeit a vast literature addresses the social-spatial segregation emerging from this dual built environment, the scarcity of spatial datasets regarding informal settlements also enforces a geo-information segregation, resulting in a terra incognita. This is exemplar in favelas, defined as precarious, spontaneous, and unorganised land occupation built on third-party property, most of which lack cadastral data. Since favelas are often not mapped, assessing urban phenomena becomes a technical challenge for several application domains, e.g., the energy one. A recent public initiative in Brazil estimates solar irradiation and photovoltaic potential for buildings at city scale, but favelas are intentionally excluded from the resulting web-based solar maps. Technicians believe that the absence of a spatial pattern in favelas calls for investigation on how to refine a roof mapping methodology. The main research question becomes: "How far is it possible to perform solar analysis on buildings of favelas in São Paulo, with the goal of estimating PV Potential?". The research is structured into two topics: 1) Roof Mapping, which investigates the data pipeline that leads to a digital reconstruction of favelas; 2) Solar Irradiation, which investigates how existing solar irradiation modules – GRASS GIS, ArcGIS, CitySim, SimStadt, Ladybug and the one developed by Virtual City Systems – perform when assessing buildings of favelas. From a roof mapping perspective, the experiments reveal that the absence of cadastral datasets represents a complex technical challenge. Nevertheless, the reconstructed and post-processed building footprints cover the extension of the cadastral footprints that are available for the study area, with building shapes that are satisfactory as a first approximation. Regarding the solar irradiation perspective, qualitative and quantitative analyses are carried out to compare the results coming from the six solar modules. The qualitative analysis indicates that each solar module offers potentialities but also limitations. Therefore, a straightforward choice is not possible, since the optimal solution will be derived from a data-driven approach that considers, among other factors: the scale of the favela, its topographical characteristics, the presence/absence of urban features other than buildings (such as vegetation), a possible pre-selection of buildings of interest, etc.; The quantitative analysis reveals that ArcGIS outputs an annual summation of irradiation values that is the closest to the one offered by the meteorological station of São Paulo, adopted as ground truth. Nevertheless, from an accuracy perspective, CitySim outputs a daily curve that best corresponds to the ground truth one. In conclusion, based on the geometrical model and the weather dataset criteria, the author expresses his preference for a raster-based solar module – GRASS GIS or ArcGIS – if, on the one hand, the reconstructed building footprints result in an unrealistic or excessively complex vector-based model. On the other hand, if the resulting vector-based model is simple enough and representative of the built environment of the favela, the author suggests the adoption of CitySim.

KEYWORDS: São Paulo, Favela, Solar Energy, 3D Modelling, Building Reconstruction.

RESUMO (PT)

GIANNELLI, Denis. Análise Solar em Edifícios de Favelas em São Paulo para Estimar Potencial Fotovoltaico. Tese para o Mestrado em Geomática – Faculdade de Arquitetura e Ambiente Construído, Universidade Tecnológica de Delft. Delft, Países Baixos, 2021.

No Sul Global, grandes áreas urbanas resultaram na dualidade entre as chamadas cidade 'formal' e 'informal'. É o caso de São Paulo, uma metrópole de 22 milhões de habitantes e centro financeiro da América Latina. Embora uma vasta literatura trate da segregação socioespacial resultante deste ambiente duplamente construído, a escassez de dados espaciais em relação a assentamentos informais também reforça a segregação geo informacional, resultando em uma terra incógnita. Isto é exemplar em favelas, definidas como ocupações de terra precárias, espontâneas e desorganizadas em propriedade de terceiros, para a maioria das quais não há dados cadastrais. Como muitas favelas não estão mapeadas, avaliar fenômenos urbanos torna-se um desafio técnico para diversas aplicações setoriais, por exemplo energia. Uma recente iniciativa pública no Brasil estima irradiação solar e potencial fotovoltaico para edifícios na escala urbana, mas as favelas são intencionalmente excluídas do web-mapa solar resultante. Alguns técnicos acreditam que a ausência de um padrão espacial em favelas requeira uma investigação sobre como refinar uma metodologia para mapeamento de coberturas de edifícios. Assim, principal pergunta da pesquisa é: "O quão possível é realizar análise solar nos edifícios em favelas de São Paulo, com o objetivo de estimar seu potencial fotovoltaico?" A pesquisa é estruturada em dois tópicos: 1) Mapeamento de coberturas, no qual é investigado qual é o fluxo de trabalho que reconstrói as favelas digitalmente; 2) Módulos de irradiação solar, no qual é investigado como módulos de irradiação solares existentes – GRASS GIS, ArcGIS, CitySim, SimStadt, Ladybug e aquele desenvolvido pela Virtual City Systems – performam na avaliação de edifícios em favelas. Na perspectiva do mapeamento de coberturas, os experimentos revelam que a ausência de dados cadastrais representa um complexo desafio técnico. Contudo, considerando a área de estudo, os polígonos 2D dos edifícios reconstruídos e pós-processados cobrem a superfície ocupada pelos edifícios existentes no arquivo cadastral, e possuem também formatos satisfatórios para uma primeira aproximação. Em relação à irradiação solar, análises qualitativas e quantitativas comparam os resultados oriundos dos seis módulos solares. A análise qualitativa indica que cada módulo solar oferece potenciais, mas também limitações. Portanto, conclui-se que uma escolha direta não é possível, visto que a melhor solução será derivada de uma abordagem baseada nos próprios dados e que considere, entre outros fatores: a escala da favela, suas características topográficas, a presencia/ausência de feições urbanas além dos edifícios (como a vegetação), a possível pré-seleção de edifícios de interesse, etc.; A análise quantitativa, por sua vez, revela que o ArcGIS produz uma somatória anual de irradiação que é a mais próxima daquela oferecida pela estação meteorológica de São Paulo, adotada como referência. Contudo, na perspectiva de acurácia, o CitySim é o software que produz uma curva diária que melhor corresponde àquela adotada como referência. Em conclusão, baseado nos critérios de modelo geométrico e de dados meteorológicos, o autor expressa sua preferência por um modelo solar matricial – GRASS GIS ou ArcGIS – se, por um lado, os edifícios reconstruídos resultarem em um modelo vetorial complexo ou excessivamente irreal. Por outro lado, se o modelo vetorial resultante for suficientemente simples e representativo do ambiente construído da favela, o autor sugere a adoção do CitySim.

Palavras-Chave: São Paulo, Favela, Energia Solar, Modelagem 3D, Reconstrução de Edifícios.

SUMMARY

LIST OF FIGURES

1. INTRODUCTION

1 INTRODUCTION

1.1 Context

In the so-called Global South[1], the urbanization process throughout the past two centuries has resulted in the development of large-scale cities across the world, many of which have become regional economical hubs[2]. With a population of over 12 million inhabitants (2020)[3] – and 21 million across its Metropolitan Area (2018)[4] –, the city of São Paulo constitutes the financial core of Brazil and Latin-America.

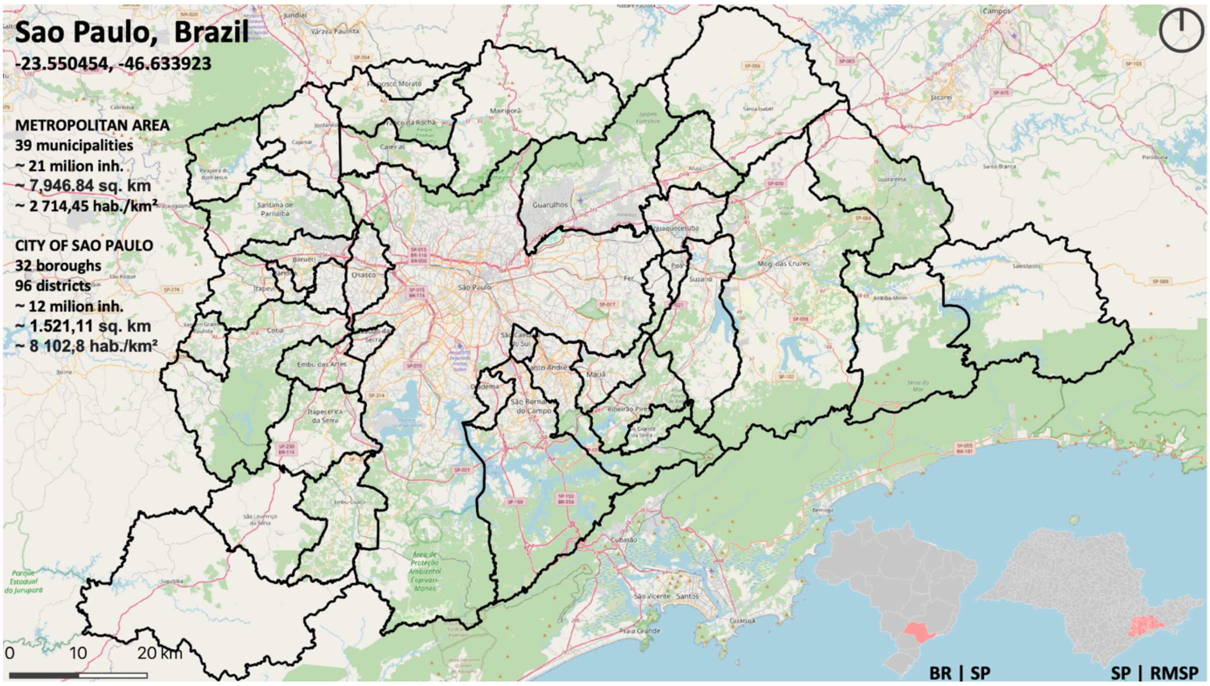


Figure 1-1 Municipalities in the metropolitan area of São Paulo (RMSP).

The City of São Paulo lies in the centre of the urban continuum.

Data source: Mapa Digital da Cidade de São Paulo, Open Street Map.

Despite its wealth, this Brazilian metropolis still faces challenges when it comes to the duality between the so-called *Formal City* and the *Informal City*[5]. A vast literature[6] discusses the

---

[1] ERIKSSEN (2015)

[2] DIRLIK (2015)

[3] IBGE Cidades – São Paulo

[4] EMPLASA

[5] FERREIRA et MOREIRA (2000)

[6] BONDUKI (2011), MARQUES (2014), MEYER et al. (2014), SOMEK (2000), VILLAÇA (2011), among others.

socio-spatial segregation effects emerging from the way that São Paulo and other Brazilian cities have been planned, economically produced, and designed in this "dual" built environment.

> "(…) unplanned settlements that have sprung up on the outskirts of Brazilian cities since the 19th century. More than 5% of the country's population lives in communities like these (…)" (DUARTE et al., 2021, p.1)

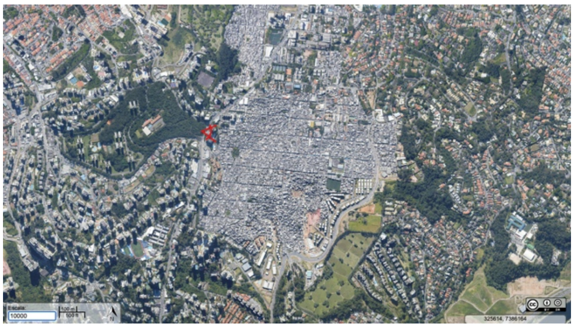


Figure 1-2 A formal neighbourhood (*Morumbi*) vs. an informal settlement (*Paraisópolis*) in São Paulo.

The red spot corresponds to an approximate position where the photograph on the right was taken

Data source: Mapa Digital da Cidade de São Paulo.

Figure 1-3 The social-spatial segregation between the *formal city* and the *informal city*.

The dwellings on the left side of the wall belong to *Paraisópolis*, whereas the condominium on its right side belongs to the luxurious *Morumbi* neighbourhood. Source: Tuca Vieira.

Nevertheless, it remains important to address, from a geo-information perspective, the impact that the scarcity of spatial datasets has over the understanding of the urban processes that take place in these informal settlements, which are, for many application domains, a '*terra incognita*'[7].

[7] The term 'Terra Incognita', derived from Latin 'Unknown Land' refers to areas that have not been mapped. According to Oxford Reference, 'Terra Incognita' is a "Latin phrase meaning 'unknown land' used by Renaissance cartographers to indicate uncharted territory", which "is now used metaphorically to describe unfamiliar or unexplored terrain of any kind."

Figure 1-4 Map of *Parque Brasilândia*, São Paulo.

Data source: Open Street Map.

Figure 1-5 Orthophoto of *Parque Brasilândia*, São Paulo.

Data source: Mapa Digital da Cidade de São Paulo.

The two images correspond to the same informal settlement in north-western São Paulo. While the mapping tool (on the left) represent the area as vegetation – a greenish colour –, the orthophoto (on the right) reveals that the area is occupied by informal dwellings.

According to the City Hall of São Paulo[8], there are four main types of informal settlements in the territory of this city.

A. ***Cortiços***, the Portuguese word for tenements, are collective and precarious housing blocks that are irregularly rented for a high amount of money per square meter. They often dispose of shared sanitary facilities among multiple rooms in a high-density environment, with precarious front door, common areas and facilities. All tenements are in the city centre.
B. **Irregular land parcelling** are informal settlements developed and/or commercialized by private entrepreneurs without previous approval by public authorities or, if such approval had been granted / was to be granted, the as-built situation is not in compliance with the specific legislation on the matter, or it does not correspond to the approved project. Low-income population and poor urban design implementation are typical of these sites.
C. ***Favelas*** are precarious settlements resulting from spontaneous and unorganized land occupation, without pre-defined land plots or road designing. They are built on both third-party private or public property, do not rely on sufficient infrastructure and the dwellings are in their majority poorly self-constructed by low-income vulnerable families.

[8] habitaSampa. Accessed on 21 September 2021.

D. ***Núcleos***, here understood as urbanized nuclei, are also favelas but fully equipped with water, sewerage, lighting, drainage, and garbage collection systems, which were sponsored either by public authorities or by private investors. Nevertheless, they are still not officially regulated in terms of land ownership.

The maps below indicate the spatial distribution of these four types of informal settlements across the territory of the city of São Paulo. By contrast, it is possible to notice a ring delimited between the city centre – the inner boundary – and the outskirts of the so-called expanded city centre – the outer boundary. Only a few informal settlements exist in this ring.

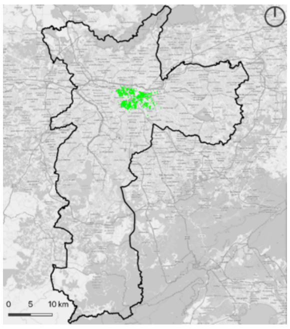


Figure 1-6 Location of *cortiços* (tenements) in São Paulo.

1.478 tenements, located exclusively in city centre.

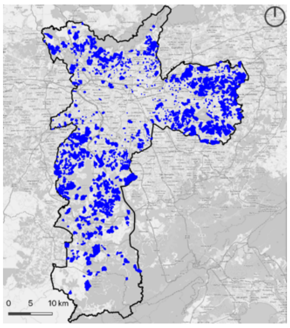


Figure 1-7 Location of irregular land parcelling in São Paulo.

1.999 irregular land parcelling, 388.459 irregular land plots

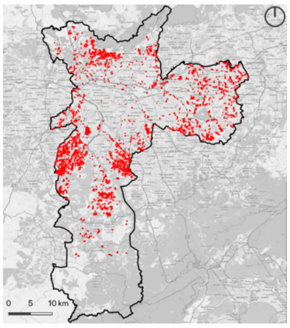


Figure 1-8 Location of favelas (slums) in São Paulo.

1.733 favelas, 391.939 households.

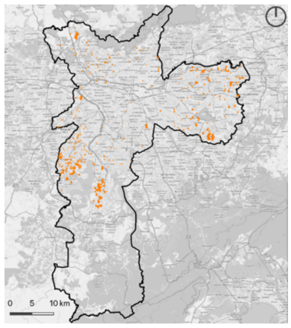


Figure 1-9 Location of *núcleos* (urbanized nuclei) in São Paulo.

435 nuclei, 60.855 families.

Data source: Mapa Digital da Cidade de São Paulo, Open Street Map.

Among these types of informal settlements, the scarcity of spatial datasets is typical of favelas[9], and most of these lack building cadastral data. When addressing favelas – but adopting the general term 'informal settlements' –, SALAZAR MIRANDA et al. (2021) provides detailed explanation about the urban morphology of favelas:

> "Informal settlements are aggregations of homes and businesses constructed by their residents, in an initially unplanned form of urbanization. Their complex morphology is the outcome of spontaneous and competitive building without official land tenure, which results in dense and multi-layered structures built up around labyrinthine street networks." (SALAZAR MIRANDA et al., 2021, p.2)

---

[9] SOUTHWICK (2016)

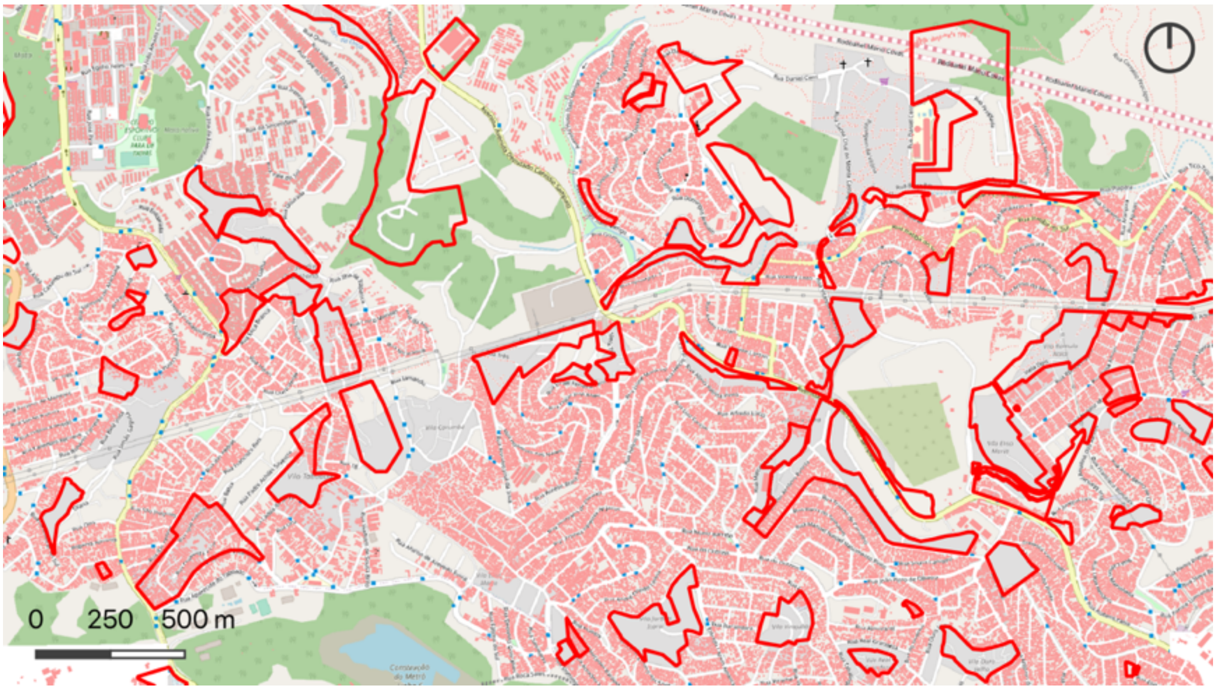


Figure 1-10 Official perimeters of some favelas (red edge) in the district of *Brasilândia*, north-western São Paulo.

Most of these areas lack official building footprints, such as the ones represented in light pink.

Data source: [Open Street Map](), [Mapa Digital da Cidade de São Paulo]().

1.2 Problem Statement

As favelas are often not mapped, assessing urban phenomena becomes a technical challenge for several application domains and, consequently, the provision of urban infrastructure is jeopardized for these areas.

> "For informal settlements in particular, measuring morphology is an entry point to broader inquiry about the tendencies of unfettered urban development and the challenges that attend it, including lack of accessibility and cadastral mapping, crowding, environmental health, and safety." (SALAZAR MIRANDA et al., 2021, p.2)

> "With little formal aid or administration and scant economic opportunities, favela residents have struggled to contend with unhealthy living conditions and frequent violence. A thick wall of social segregation means that resources from the city – including electricity and clean water – must take twisting, uncertain paths to make it inside." (DUARTE et al., 2021, p.1)

One of these applications domains is the energy one. A 2015-public initiative in Rio de Janeiro, another city in Brazil, estimates solar energy and photovoltaic potential for buildings at city scale, but favelas are intentionally excluded from the resulting web-based solar maps.

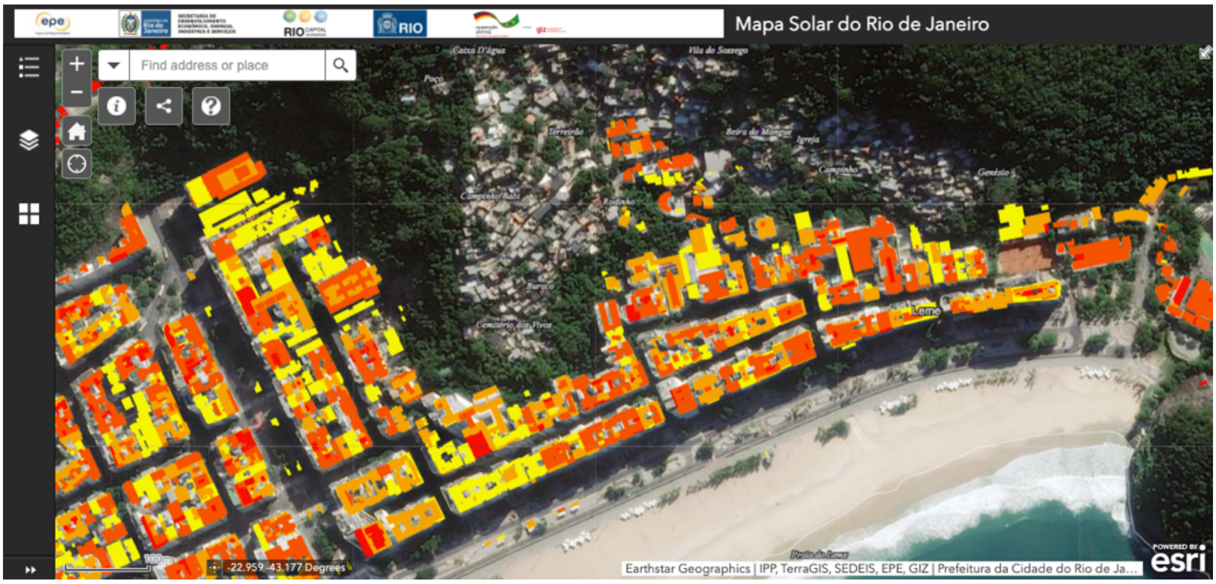


Figure 1-11 The formal neighbourhoods of *Leme* and *Copacabana*, in Rio de Janeiro, with their buildings modelled and classified by solar potential classes.

The favela *Morro da Babilônia* is excluded from such classified representation.

Data source: Mapa Solar do Rio de Janeiro.

FEITOSA et al. (2020) emphasise both political and technical reasons why favelas would not be present in the Web Solar Map of Rio de Janeiro. When strictly addressing the technical aspects, these authors claim that the absence of a spatial pattern in favelas would call for an investigation on how to refine a roof mapping methodology:

> “The lack of information from these regions on the Solar Map suggests, therefore, that GIZ-indicated refinement studies have not been carried out or that their products have not been utilized. It would bring greater methodological complexity to the construction of the Map to include such areas.” (FEITOSA et al., 2020, p. 36)

> “Including poor areas would require more complexity from the point of view of refining the roof mapping methodology, as already mentioned regarding the possibility of developing a more inclusive methodology.” (FEITOSA et al., 2020, p. 37)

Despite this 'greater methodological complexity' claim, the developers of the web-based Solar Map of Rio de Janeiro themselves admit the possibility of investigating a specific methodology for favelas.

> "And finally, there is unregulated residential occupation, which consists of slums. This type of occupation is not part of the roof mapping process in the present study. Its main characteristic is the absence of a spatial pattern, in addition to the impossibility of identifying streets, lanes or alleys, considering the proximity among the houses." (…)" (LANGE, 2015, p. 13 – translated by the author)

> "Subnormal areas[10] have characteristics that make it difficult or impossible to interpret or extract information through remote sensing techniques. This does not mean that it is not possible to use such techniques to investigate subnormal agglomerations. However, due to the limited time of this project, it was not possible to develop specific methods of remote sensing for the purposes of the study." (LANGE, 2015, p. 27 – translated by the author)

In parallel with this public initiative that maps solar irradiation and PV potential at urban scale, REVOLUSOLAR – a Rio de Janeiro-based non-profit organization – has been designing and implementing PV power plants in favela settlements.

At the present time, this organization is mainly focused on one community, *Morro da Babilônia e Chapéu Mangueira,* where three solar power plants were installed and a fourth one – in a cooperative model – is being implemented. In a meeting with technical directors, it was claimed that an expansion to other favelas in this city is foreseen by 2022, yet they do not have sufficient knowledge of these other sites in terms of solar irradiation and PV potential. This reinforces a demand for a specific roof mapping methodology for the purpose of solar analysis and PV potential, under the context of scarce cadastral data regarding favelas.

---

[10] Subnormal agglomerate is the official term adopted by the Brazilian Institute of Geography and Statistics (IBGE): "Subnormal agglomerate is a form of irregular occupation of land owned by others - public or private entities - for housing purposes in urban areas and, in general, characterized by an irregular urban pattern, lack of essential public services and location in areas with restricted occupation."

Figure 1-12 "*Babilônia* Rio Hostel" PV power plant (2016).

Source: Revolusolar

Figure 1-13 "*Estrelas da Babilônia*" PV power plant (2016).

Source: Revolusolar

Figure 1-14 "*Escolinha Tia Percília*" PV power plant (2019).

Source: Revolusolar

In the context of the city of São Paulo, however, neither literature nor hands-on experience was encountered when it comes to solar analysis and PV potential for favelas at an urban scale, even though favelas in São Paulo similarly lack urban infrastructure as the ones in Rio do.

This absence offers the possibility to expand the work that has been carried out in Rio de Janeiro to the city of São Paulo. In addition, the spatial data infrastructure of the City Hall of São Paulo allows a much broader investigation on the matter of solar analysis in favelas, considering the presence of useful data for the research.

1.3 Research Questions

Considering the context and problem statement, the research question that this MSc thesis addresses is:

"How far is it possible to perform solar analysis on buildings of favelas in São Paulo, with the goal of estimating PV Potential?"

This research question is then subdivided into two main investigation topics, which guides the workflow of the thesis and structures the present document:

Topic 1 – Roof Mapping

It is important to investigate, from a 3D modelling perspective, the data pipeline that leads to a digital reconstruction of favelas. Therefore, based on the existing spatial data infrastructure of São Paulo, this topic aims to respond:

- Considering the absence of cadastral data, what are the minimum geodata required to map buildings in a favela, in particular building roofs?
- What are the necessary algorithmic steps to achieve such goal?
- Considering the urban morphology of favelas, do existing algorithms satisfy the digital reconstruction task?
- If not, what are the important additional challenges to observe when setting up a specific methodology adapted to the environment of a favela?

Topic 2 – Solar irradiation modules

In parallel with the roof mapping, it is also important to understand, from a GIS perspective, how existing solar irradiation modules perform when assessing buildings of favelas. The selected software analysed in this MSc thesis, which contain solar irradiation modules, are:

- OSGeo **GRASS GIS**
- ESRI **ArcGIS**
- KAEMCO CitySim Pro / EPFL **CitySim** Solver
- HFT Stuttgart **SimStadt**
- Ladybug Tools **Ladybug**
- Virtual City Systems **3D Solar Potential Analysis**

Therefore, on a comparison basis, this topic aims to respond, for each solar irradiation module that is assessed in this thesis:

- What are the minimum data requirements to run the solar irradiation module?
- How automatized is the solar irradiation pipeline? From a user perspective, how complex is setting up the simulation?

- Is the solar irradiation module self-contained, or is it part of a greater package of solutions?
- What is the computational running time of the module if this metric is retrievable?
- What kind of urban features – among buildings, vegetation, and relief – does the module support, and how do these features participate in the simulation?
- What is the degree of flexibility offered by the solar module in terms of selecting which features of the city model will be assessed and which will take part only as shading elements?
- How is the output data delivered by the module in terms of time granularity and data model?
- With respect to the meteorological station of São Paulo, which offers ground truth data, what is the level of accuracy of the module in terms of solar irradiation?

It is important to emphasise that, for the present research, no data were collected on site by the author, and therefore all datasets come from secondary sources. Each dataset is subject to outliers, and therefore the presented results may be affected by error propagation.

It is also convenient to stress that this research does not intend to model PV panels for a specific favela, but rather investigate existing solar irradiation modules. In other words, there is no concrete project for a photovoltaic power plant.

Finally, it is opportune to remark that neither roof mapping nor solar irradiation algorithms were designed by the author. All software employed in this thesis were developed by third parties. Nevertheless, the author carried out programming assignments to provide each solar irradiation module with their specific required data.

1.4 Results Overview

From a roof mapping perspective, the conducted experiments reveal that the absence of a cadastral dataset represents a complex technical challenge. The 3D building reconstruction algorithm – adapted from the one that the 3D Geoinformation group implemented for the 3D Dutch cadastre – delivers a collection of 2D footprints that, in general, covers the extension of the cadastral footprints that are available for the study area. Nevertheless, these reconstructed footprints are very fragmented and complex with respect to their official counterparts. The post-processing steps can mitigate these undesired characteristics, but the shapes of the final building footprints are often much more complex than the official footprints. Still, the overall disposal of the reconstructed building roofs and their heights – which are close to the ground truth dataset – play a much more significant role than the precision of their edges.

Regarding the solar irradiation perspective, qualitative and quantitative analyses are carried out to compare the results coming from the six solar modules in analysis. The qualitative analysis indicates that morphological characteristics of favelas demand solutions that consider: the great number of building surfaces, the often-complex surrounding relief, and the shading effect that these multiple adjacent buildings produce into each other; The quantitative analysis, in turn, reveals that ArcGIS outputs an annual summation of irradiation values that is the closest to the one of ground truth values, the former only 0.89% greater than the latter. Nevertheless, from an accuracy perspective, CitySim is the software that outputs a curve that has the best correspondence to the one of ground truth.

1.5 Thesis Structure

The present MSc thesis is structured as follow:

- Chapter 1 situates the research in the urban context of São Paulo, Brazil, states the problems that are hereby addressed, presents the research questions and offers an overview of the results.
- Chapter 2 is composed of a literature panorama regarding the theoretical background and related work on the fields of Urbanism, Geomatics and Building Physics.
- Chapter 3 describes the method and technical procedures that are established to respond the main research question and all sub-research questions present in both topics.
- Chapter 4 offers detailed explanation regarding the data workflow that is carried out in each solar irradiation module. It also provides the reader with the algorithmic steps for mapping the building roofs in favelas.

- Chapter 5 exposes the results delivered by each solar irradiation module and compare them quantitatively and qualitatively. In addition, it reveals the reconstructed building footprints and their resulting CityGML model by means of visual interpretation and statistical analysis.
- Chapter 6 summarises the achieved results, responds the research question(s), suggests future related work, and presents a personal reflection of the author.

2. THEORETICAL BACKGROUND & RELATED WORK

2 THEORETICAL BACKGROUND & RELATED WORK

As the present thesis lies in a cross-domain field of research among Urbanism, Geomatics and Building Physics, groups of relevant literature are resourceful to sustain the topic and strength the theoretical background regarding the digital modelling of solar energy in favelas. For pragmatic reasons, the related work is hereby presented in groups, albeit these often intertwine among themselves. The scope of this chapter is not to present an exhaustive discussion of each domain, but rather to understand why 'embracing the informal' is fundamental and, moreover, to focus on meaningful concepts that are necessary to understand this MSc Thesis.

2.1 The *Global South*

The term *Global South* has been present in scientific literature for at least the last fifteen years. Its origin relates to the socio-economical transformations of the globalized world after the end of the Cold War and the establishment of the New International Order. The triad *first*, *second* and *third world countries*, or even the simpler dichotomy *developed* and *developing* countries no longer sustained the complex stakeholder positions that some countries play in the global economy, albeit their fragilities:

> "For what was Argentina? Or Turkey? (…) It makes little sense to speak of three worlds when there is only one game in town. Instead, during the last decade or so, scholars and enlightened commentators increasingly have begun to speak of the Global South and the Global North." (ERIKSEN, 2015, p.1)

At the present time, the term is world-wide spread, but consensus regarding its mean is still hard to be achieved. The same author claims that "any conceptual investigation of these classifications must inevitably lead to ambivalence. Global diversity is simply such that it cannot meaningfully be subsumed under a few, let alone two, concepts." (ERIKSEN, 2015, p.1).

Despite the plurality of semantics, it is a common understanding that the extents of the Global South are not limited to the Southern Hemisphere, as clear exceptions rise in both sides of the planet. Instead, a social-economic notion perceives:

> "While the countries of the Global North not only have stable states but also a strong public sector, the Global South is, to a far greater extent, subject to the forces of global neoliberalism, rather than enacting the very same forces. (…) This is why the prefix "Global" may be appropriate, as it signals the integration of the entire planet (well, nearly) into a single economic system (…) The Global South and the Global North represent an updated perspective on the post-1991 world, which distinguishes not between political systems or degrees of poverty, but between the victims and the benefactors of global capitalism." (ERIKSEN, 2015, pp.1-2).

This idea of a dialectical materialism relationship between North and South, in which one exists not despite, but rather because of the other, is key to comprehend not only global phenomena, but also local ones. This is especially true, at urban scale, in countries of Asia, Africa, and Latin America, in which wealth and privation of basic needs lie side-by-side.

> "Areas incorporated under the label Global South can also be found in the geographical North. Ethnic ghettos and barrios in US American cities are one example; the "Latinoization" of the US is another. And the gated communities of the cosmopolitan elite in Rio de Janeiro, Mexico City, or Santiago de Chile have more in common with their counterparts in Miami, L.A. or Chicago than with the surrounding barrios, marginales and favelas." (KALTMEIER, 2015, p. 1)

2.2 'Formal vs. Informal City'

In several countries across the world, at urban scale, the conflictive co-existence between rich and poor and, moreover, the social-spatial segregation between these counterparts has produced an abundant literature regarding what have been designated as the *formal city* and the *informal city*.

When addressing this contradiction in Brazil, FERREIRA et MOREIRA (2000) cite MARICATO (1996), who claims that the escalation of the duality between the formal city and the informal city has been a constant in the structuring of Brazilian cities. These authors characterize the former as the urban space which has always served the elite and their interests, whereas the latter is occupied by the working class in the context of the Fordism mode of production.

> "The State has been historically absent and negligent to any obligation of ordering or regulating the occupation of urban space in the peripheries, as they in practice did not serve the interests of the capital. Therefore, there hasn't been an urbanization process in Brazil that would effectively and fairly distribute urban real estate and urban infrastructure to the society as a whole.
>
> (…)
>
> As a result of this process, there is the ascension in the dichotomy between the formal city and the informal city. A modernization paradigm overlaps an existing urban fabric that, paradoxically, still materializes the long-term social inequality and the social-spatial segregation resulting from the dependent industrialization." (FERREIRA et MOREIRA, 2000, pp.1-2, translated by the author)

In the contemporary debate about informal settlements in Brazilian cities, much attention has been paid to the different approaches that could reduce social vulnerability of inhabitants of the *informal city*. One approach is granting access to the formal real estate market by means of credit. In this sense, the financial market and the construction sector offer alternative social housing solutions for the inhabitants of poor areas of the city. This model has been supported by successive governments, with *Minha Casa Minha Vida* – Portuguese for 'my house, my life' – being the dominant housing policy, albeit the heterogenous landscapes and urban quality resulting from such projects.

Figure 2-1 Social housing in São Luís, MA, Brazil.

Source: Archdaily Brasil.

Figure 2-2 Social housing in São Paulo, SP, Brazil.

Source: Archdaily Brasil.

Despite the long-term effort on housing policies, the typical urban fabric of informal settlements has been present in Brazilian cities in high proportions. Constantly updated statistics coming from the Secretary of Housing of the City Hall of São Paulo reveal the figures for these phenomena. As of September 2021, the territory of São Paulo hosted 1.478 tenements, located exclusively in city centre; 1.999 irregular land parcelling with 388.459 irregular land plots; 1.733 favelas with 391.939 households; and 435 urbanized nuclei with 60.855 families.

> "In the 20th century, the Brazilian government attempted to eradicate favelas and replace them with more formal public housing, but the bulldozers could not keep up with the massive urban migration that made these settlements swell." (DUARTE et al., 2021, p.2)

Therefore, a second approach consists in the urbanization of informal settlements, more specifically of favelas, in which a lack of urban infrastructure is predominant.

2.3 Favelas

The effort of providing favelas with urban infrastructures of all sorts – such as official lighting, clean water, and sewage systems, but also green areas, health care facilities, educational institutions, etc. – has been a major policy not only in São Paulo, but also in other informal urban areas across Brazil and Latin America.

> "Efforts to eradicate these communities gave way to incorporation; the government choose them as the sites for new libraries and public parks. The Medellín model, despite some shortcomings, has since become the gold standard in Latin America and around the world." (DUARTE et al., 2021, p.2)

An illustrative example of such sort of urban project is the urbanization of *Cantinho do Céu*, reported in the website *Archdaily Brasil*. In this project, a whole community that lies in the skirts of São Paulo's water reservoirs was redesigned an equipped with the necessary infrastructure.

Figure 2-3 Urbanization of Cantinho do Céu, São Paulo, SP, Brazil.

Source: Archdaily Brasil.

Figure 2-4 Urbanization of Cantinho do Céu, São Paulo, SP, Brazil.

Source: Archdaily Brasil.

DUARTE et al. (2021) interprets the urbanization of favelas under a historical point of view, claiming that several renowned urban areas around the world had their origins as informal settlements. Following their rhetoric, the approach of providing favelas with urban infrastructure is not innovative per se, but rather a historical trend in city planning.

> “How we choose to respond to favelas, shantytowns, and refugee camps over the next few years will define the political and cultural attitudes that determine their long-term future. It’s worth recalling, then, that most cities were born from informality. Many parts of Paris were this way until the interventions of Baron Haussmann, whose 19th-century redesign was made possible with the force of a military strong-man. Fifty years ago, Singapore was still a city of shantytowns. New York once had more illegal settlements than anywhere else in the US, but waves of gentrification have allowed us to forget that messy history.” (DUARTE et al., 2021, p.2)

Albeit the positive efforts that have been made on the field of urbanization of favelas, the spontaneous characteristic of these sites, driven by an auto-construction designing process, makes them unofficial for the local authorities in terms of land ownership and building cadastre. Therefore, except for local inhabitants or social assistants, who act directly on the field, little information about these sites is known by city planning stakeholders. The scarcity of datasets regarding favelas, or even the limited methods of retrieving such urban data automatically, pose a critical problem in the work pipeline of the local technicians.

> “Having better knowledge of the favela’s physical layout could also improve living conditions. Urban designers could use this data to decide where to install stairs, or what structures to remove to allow in more air, sun or light.” (DUARTE et al., 2021, p.2)

Complementary, in an article on the effects of the 2016 Olympic Games in Rio de Janeiro, SOUTHWICK (2016) addresses the importance and challenges of putting favelas on the map. One of the reasons is also a sense of belonginess:

> “For me, the fact of not being on the map, creates a sense of exclusion. That we are not part of the city, that we are not part of the traditional script,” said Paulinho Otaviano, a resident and local guide in Santa Marta, in an interview in *Todo Mapa Tem Um Discurso* (Every Map Has a Discourse), a documentary that explores questions of mapping, geography, identity and representation in Rio’s favelas.” (SOUTHWICK, 2016)

With respect to the provision of energy infrastructure, FEITOSA et al. (2020) indicates what is known as energy poverty in the favelas of Rio de Janeiro: “These population groups also suffer from what economists are calling “energy poverty”, in which marginalized communities regularly lack legal access to energy utilities due to either poor service provision or generalized economic exclusion.”

Under this context, REVOLUSOLAR, a Rio-de-Janeiro based social organisation, was created with the task of spreading solar energy installations in low-income communities, therefore aiming to reduce the energy bills of households in favelas. In one of their reports, the phenomenon of energy poverty is described for *Babilônia*, one of Rio’s favelas that lies so close by the well-known Copacabana beach. The report demystifies widely spread common senses, such as the one that all households of favelas have illegal connections to the electrical system – popularly known as *gatos* (Portuguese for cats*)* –, or that their inhabitants simply do not for pay their energy bills. On the contrary, the research concludes that, due to energy losses in the network and high energy fees, the inhabitants of favela pay a high amount of money for what they consume, in comparison with their counterpart citizens in the formal neighbourhoods.

This contradiction justifies the hands-on experience of REVOLUSOLAR in the field of photovoltaic energy production. According to FEITOSA et al. (2020):

> “Pol Dhuyvetter, a Belgian activist that has lived in *Babilônia* and worked on the installation of the first photovoltaic systems in that favela, stated that it

> was possible to provide electricity to all households from photovoltaic systems on their roofs. Together with other community leaders, such as the electrician Adalberto Almeida, they founded RevoluSolar, an association to promote energy production in the favela from alternative sources, especially solar.”

At the present time, this organization is mainly focused on one community, *Morro da Babilônia e Chapéu Mangueira,* where three power plants were installed and a fourth one – in a cooperative model – is being implemented.

While this pioneer experience is extremely positive for the local community that benefits from new solar power plants, it is also true that assessing solar irradiation in a favela remains a complex technical challenge for the professionals that must decide where to allocate the PV panels for their optimal performance.

In a meeting with the technical staff of Revolusolar, the author questioned which criteria the engineers adopted for choosing the site of these four existing power plants. Their response was guided by several factors, such as vertical (slope) and horizontal (azimuth) angles of the building roofs, shading from the surrounding buildings, vegetation and relief, and space availability. Nevertheless, also according to the staff, this selection hadn’t been made by a computational approach, but rather by local experience from people in the field, and aerial photo interpretation.

The organisation, however, claimed that it is their intention to expand their activities to other favelas in the city of Rio de Janeiro by 2022. They claim, therefore, that in parallel with other fundamental criteria – such as the building structure and the presence/absence or organised militia in the area –, the solar irradiation of building roofs also needs to be assessed to optimise this research for new power plant sites.

This technical challenge justifies both topics of the present MSc thesis: how to map building roofs in favelas and how to assess solar irradiation in these building roofs.

2.4 Geomatics + 3D Modelling

A 2015 public initiative by the city hall of Rio de Janeiro estimates solar energy and photovoltaic potential for buildings at city scale, but favelas are intentionally excluded from the resulting web-based solar maps. The technical arguments provided by the developers rely on the fact that, in the absence of cadastral data of favelas, there would not be a viable solution that could extract the roof boundaries of buildings in favelas by means of remote sensing techniques, at least considering the limited timeframe of the project.

Despite this allegedly absence of a specific roof mapping methodology for buildings in favelas – in the context of the Rio's web-based solar map project –, the topic of measuring urban morphology of favelas has been present in more recent literature.

DUARTE et al. (2021) and SALAZAR MIRANDA et al. (2021) detail the recent work that has been developed by the Senseable City Lab of the Massachusetts Institute of Technology – MIT. In this project, *Favela da Rocinha*, in Rio de Janeiro, was surveyed by a group of experts with a terrestrial laser scanner.

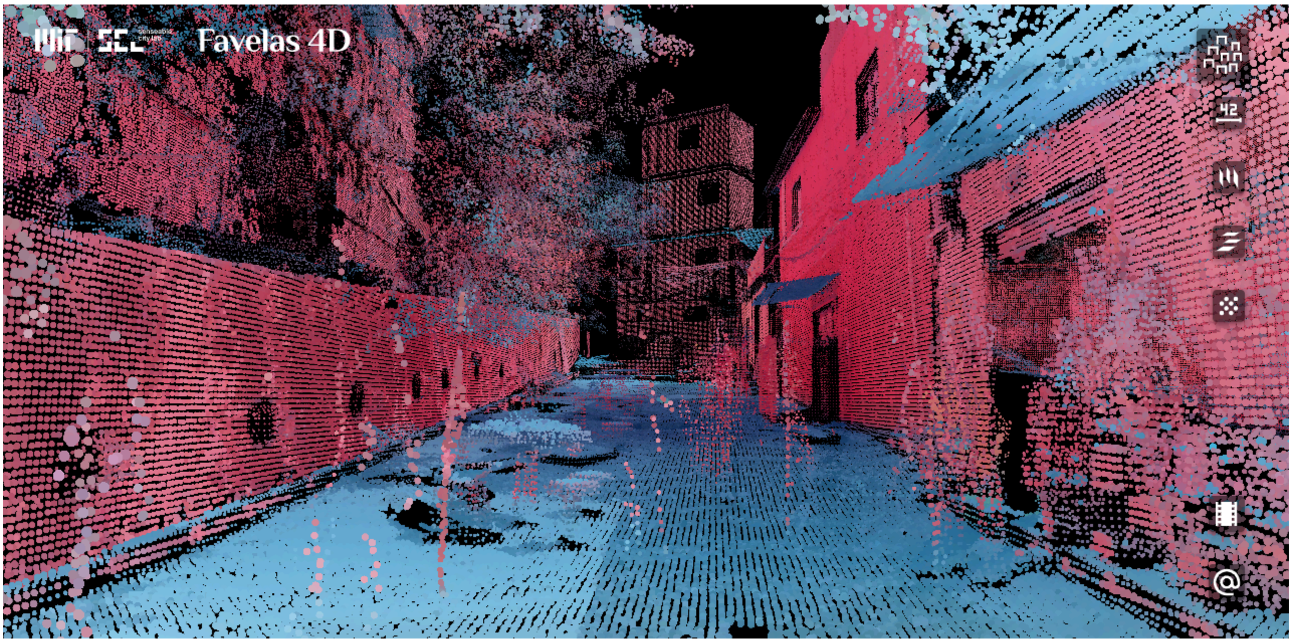


Figure 2-5 Web-viewer of the project Favelas 4D, designed by the Senseable lab of MIT.

The terrestrial laser scanner collected LiDar point cloud datasets, which made possible a further digital reconstruction of Favela da Rocinha, in Rio de Janeiro, Brazil.

Source: Favelas 4D.

Considering the present research, their work is useful for three purposes:

- Firstly, it demonstrates the necessity of creating digital models of informal settlements for several urban applications, i.e., “embracing the informal” (DUARTE et al.,2021)

- Secondly, it offers a recent panorama of related literature that deals with morphological studies regarding informal settlements from a spatial data perspective, which are in turn sub-divided into four categories:

    - “Remote sensing-based studies that have refined methods of identifying informal settlements and modelling their growth”; By remote sensing, the authors understand satellite imagery-based works, namely:
        - WURM et al. (2019), in which the boundaries of favelas are delineated based on their distinct urban fabric.
        - DUQUE, PATINO et BETANCOURT (2017), WURM et al. (2019) and STARK et al. (2020), in which machine learning is used to predict the presence and expansion of favelas via detection of typical building roofs.

    - A group that “focuses on the definition and classification of informal settlement topologies using street network maps”, including:
        - SOBREIRA (2007), in which the morphology of favelas is analysed in contrast to the formal city or as a globally consistent urban typology of its own.
        - LOUF et BARTHELEMY (2014), BOEING (2021), TAUBENBÖCK, KRAFF et WURM (2019), in which analytical methods produce wide-ranging taxonomies of urban plans.
        - ZAPULLA, SUAU et FIKFAK (2014), LOUREIRO, DE MEDEIROS et GUERREIRO (2017), with a focus on classifying the street networks of informal settlements.
        - LOUREIRO, DE MEDEIROS et GUERREIRO (2019), in which favelas are compared to the other topological conditions, such as those of the medieval Portuguese towns that could be considered their ancestors.

- A group that "focuses on studying informal construction at the building scale", including:
    - ALEXANDER et al. (1977), with a focus on pattern language.
    - KNIGHT et STINY (2015), which makes use of shape grammar.
    - KAMALIPOUR et DOVEY (2020), DOVEY et al. (2020), which analyse pattern of constructions in informal settlements.
    - DOVEY (2013), BARDHAN et al. (2018), DOVEY et al. (2010) with the purpose of growth modelling.
    - DOVEY (2013), BARDHAN et al. (2018), CHOKYU et DIAS (2018), VERNIZ et DUARTE (2020), focusing on parametric design.

- "A limited selection of LiDAR[11]-based analyses that have been conducted in favelas at small scales", at the expense of relying on manual data processing. This group includes:
    - TEMBA et al. (2015), in which roof boundaries are delineated to develop a digital surface model (DSM).
    - RIBEIRO et al. (2019), with a focus on counting the number of pavements in buildings of favelas for cadastral purposes.

- Thirdly, for their own development, the MIT researchers reported a methodology to extract urban scenes – i.e., to create a digital model of the built environment in *Rocinha* – in a street view perspective. According to the researchers, two main steps were followed.

    - Extracting planes from the point cloud, which consists of shape detection and plane regression.
        - In this step, a Random Sample Consensus (RANSAC) algorithm is adapted from SCHANBEL, WAHL et KLEIN (2007). According to the

[11] According to the definition in MIRANDA et al. (2021), LiDAR – Light Detection and Ranging – is "a 3D scanning technology that captures laser range measurements by measuring the level to which a laser pulse emitted from the tripod-mounted scanner reflects its beam. The resulting data are a set of individual distances stored as points with 3D coordinates registered spatially using the GPS position of the sensor, also known as a point cloud."

authors, RANSAC is suitable for filtering outliers in the point cloud, i.e., points that do not belong to the plane which is to be extracted.

- Shape detection: the RANSAC algorithm inputs the unorganised collection of points and five user parameters[12], and then clusters these candidates into sub-clouds that, once isolated, approximate to the shape of an existing surface in the built environment.
- Plane regression: for each clustered set of points, the RANSAC algorithm process a regression to obtain the optimal surface that fits these points, i.e., a structured 3D geometry represented in a vector model.

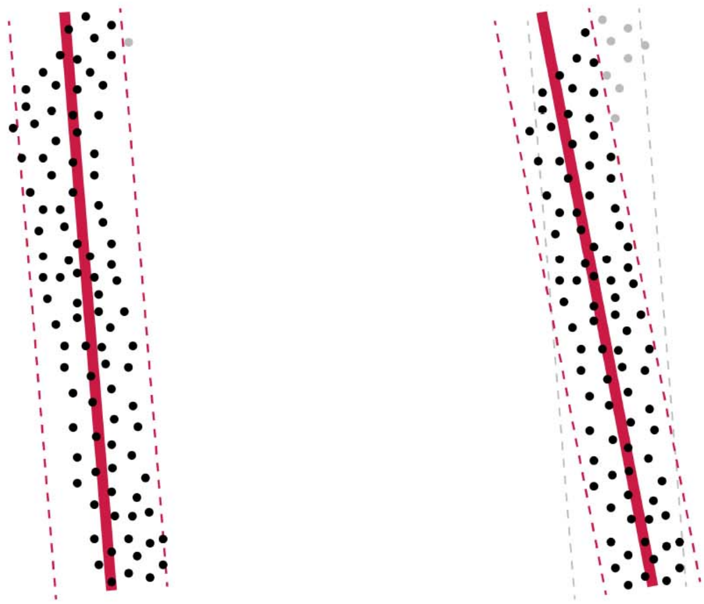


Figure 2-6 RANSAC algorithm, as described in SALAZAR MIRANDA et al. (2021)

Graphic explanation of the plane extraction from point cloud shown as a section view of vertical plane. Left panel shows the RANSAC shape detection step. The right panel shows the RANSAC plane regression step.

Source (image and caption): SALAZAR MIRANDA et al. (2021)

---

[12] According to MIRANDA et al. (2021): Minimum support points, i.e., "the minimum number of point required for a point cloud fragment to be considered a plane candidate"; Maximum distance to primitive, i.e., "the maximum distance of sample point to plane to be included in the corresponding set of points of a plane, given as a ratio of bounding box width"; Sampling resolution, i.e., the "distance between neighbour sampling points, equivalent to the point cloud resolution"; maximum normal deviation (alpha), i.e., "the maximum deviation of the normal of initial randomly selected points from the normal of the plane defined by them"; overlooking probability, i.e., "the probability that no other better primitive candidates are overlooked (…) during sampling".

- Constructing street scenes from planes, which makes use of a nearest-neighbours approach to connect horizonal (streets) and vertical (façades) surfaces into an aggregated street view.

Analogously to the work that has been conducted by the Senseable lab of MIT under the context of favelas in Rio de Janeiro, the 3D Geoinformation Group of the Faculty of Architecture and the Built Environment at Delft University of Technology created a new data pipeline in order to automatically reconstruct all buildings of the Netherlands in 3D.

Figure 2-7 The web-based viewer of the 3D building cadastre of the Netherlands.

Source: 3Dbag.nl

In OHORI, LEDOUX et PETERS (2021, 1), the authors present this 3D building reconstruction assignment with two important aspects to be observed beforehand, namely the model requirements and the reconstruction challenges:

According to these authors, the model requirements establish minimum criteria that the output model should attend to fulfil its purposes, and these may be summarised into geometrical complexity, accuracy with respect to the original point cloud, geometrical validity regarding pre-defined standards and level of detail (LoD), which in turn express the degree of generalisation in the roof structure of the building model, e.g., flat roofs in LoD1, multi-pitched roof shapes in LoD2, etc.

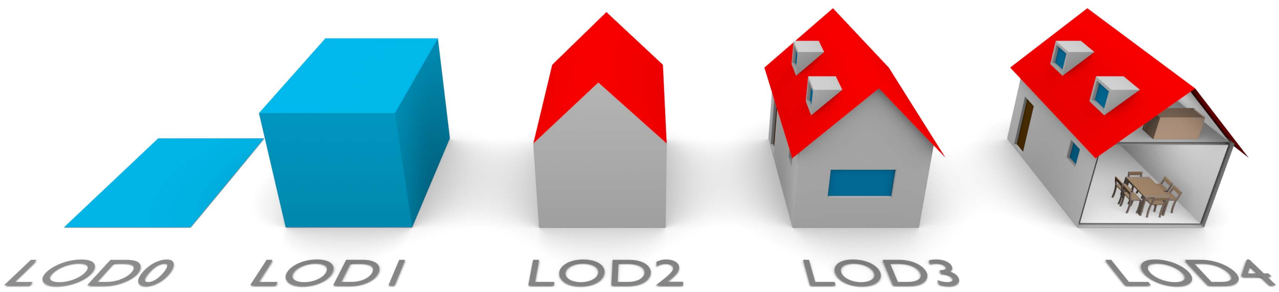


Figure 2-8 The five LODs of CityGML 2.0.

The geometric detail and the semantic complexity increase, ending with the LOD4 containing indoor features.”

Source (image and caption): BILJECKI et al. (2016).

In practical terms, for the case of solar analysis on an urban settlement, the reconstructed buildings need to be as little complex as possible, since the computing steps require heavy geometric processing, but accurate enough with respect to the edges and heights of their roofs, so that the shading effect is realistic. It also needs to be validated by the software that runs the solar module.

Moreover, for the specific context of favelas, modelling the buildings in ‘pseudo’ LoD2 is an acceptable assumption since most dwellings are designed with flat roofs. In this case, the geometries of buildings are created by extruding an LoD0 footprint, as it would be the case in a simple LoD1 model, but the resulting surfaces are semantically enriched such that they can be recognised as grounds, walls, and roofs, as it is the case of an LoD2 building.

Following the narrative of OHORI, LEDOUX et PETERS (2021, 1), the reconstruction challenges relate to the original input dataset from which the model will be derived and may be sub-divided into two main categories: firstly, the urban morphology of the built environment plays a significant role in how the algorithm outputs features, based on pre-defined criteria; secondly, the quality and completeness of the dataset also affects the performance of the algorithm.

Therefore, in the case of mapping building roofs of favelas, the urban morphology challenge relies precisely on correctly identifying and isolating buildings that are adjacent to each other, and whose (flat) roofs not always have a great height difference. Regarding the data quality, mapping roofs require a surveying technology that senses these roofs, which is the case of an aerial lidar dataset or an orthophoto, but not the case of a terrestrial laser scanner as the one used in the MIT's project in *Rocinha*.

Another important distinction mentioned in OHORI, LEDOUX et PETERS (2021, 1) is data-driven approach versus model-driven approach. The former relies on a strong data quality and completeness, and delivers features with a good data fit, but at the expanse of a great geometrical complexity. The latter, instead, relies on crucial geometrical assumptions to produce much simpler features – which are based on a pre-defined library of building shapes –, but with less dependency of the original input data.

For the automatic LoD2 reconstruction of the Netherlands, the 3D Geoinformation group combined both approaches and created a data pipeline that outputs buildings suitable for several environmental simulation applications, such as solar analysis.

The modelling assumptions are based on the characteristics of the Dutch datasets used by the researchers:

- A piecewise planar method approximates all surfaces of the buildings to planar faces, thus replacing non-planar shapes, such as a sphere, with a piecewise similar form.
- The roofs may have complex shapes but are always treated as 2.5D (one height value at every position in the XY plane), whereas the walls are always vertical.
- The input point cloud must be classified according to the semantics of the built environment, i.e., points representing building and grounds must be marked as such.
- The building footprint dataset must be present, so that points belonging to each building are clustered together beforehand.

When dealing specifically with buildings of favelas, the first two assumptions are satisfied by the morphological characteristics of favelas. The auto-construction designing process does not foresee great geometrical complexity, and buildings can be generalised as extruded blocks based on their footprint and height. The third assumption will depend on the spatial data infrastructure from which the dataset comes from: classified point clouds may be available for some cities, as it is the case of São Paulo, but not for others.

However, as exposed in the problem statement of this research, the reconstruction of buildings in favelas fails to attend the fourth assumption, since building footprints are often unavailable. Therefore, if this algorithm is to be implemented, it should be adapted so that the whole set of planes are extracted as if they would all belong to the same building. From the original algorithmic steps that are represented below, the present research focuses on extracting the boundary lines as in step 4, but without using the input footprints as in step 1.

In synthesis, the original algorithmic steps that are implemented in the 3D building reconstruction of the Netherlands are:

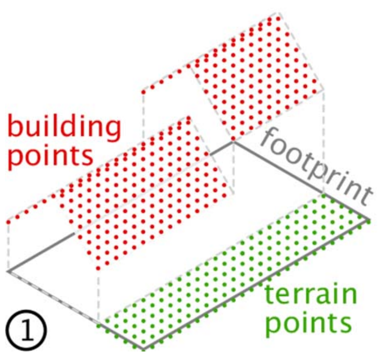


Figure 2-9 The classified (aerial) point cloud is cropped on the 2D footprint

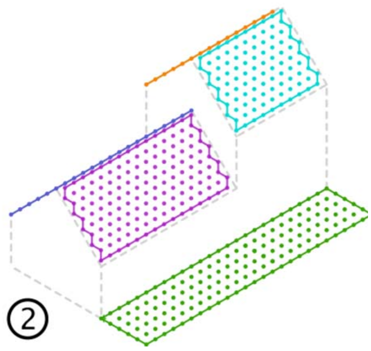


Figure 2-10 Planes and their boundaries are detected in the point cloud

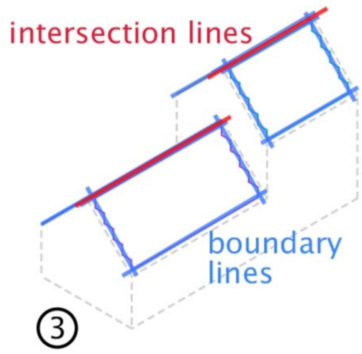


Figure 2-11 From the roof planes the intersection lines and boundary lines are extracted

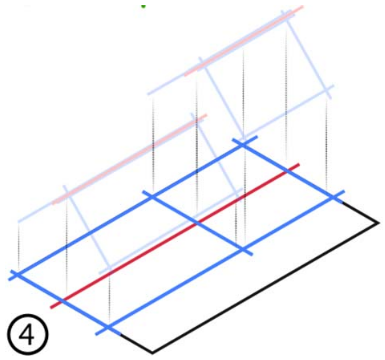


Figure 2-12 The lines are regularised and projected onto the 2D footprint,

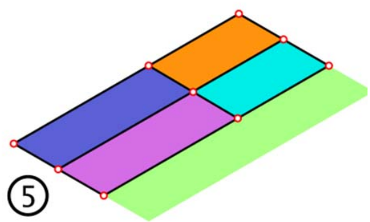


Figure 2-13 The roof-partitions is created. This is a DCEL where each face is labelled with the corresponding plane (from 2, compare colours).

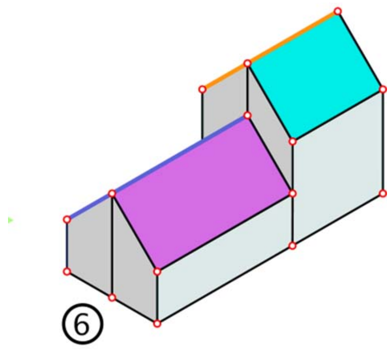


Figure 2-14 The roof-partition is extruded into a 3D mesh.

The six main steps of the algorithm for the automatic LoD2 reconstruction of the Netherlands, as described in OHORI, LEDOUX et PETERS (2021, 1).

2.5 Geomatics + Energy Simulations

The task of performing energy simulations with computational models requires a clear understanding of some physical concepts that are commonly found in the theoretical literature and software documentation.

Solar irradiance is a measurement of power (W) per unit area (m^2), i.e., W/m^2 in the international system of units. It expresses an average of electromagnetic energy (J) that comes from the sun in a certain time interval (s) per unit of area (m^2).

Solar irradiation is a measurement of energy per unit area, i.e., J/m^2 or Wh/m^2. It expresses the accumulated amount of electromagnetic energy (J) that comes from the sun per unit of area (m^2) within a certain time interval. Given the practicality of expressing the amount of solar irradiation in a period of one hour, the energy community often adopts the unit of measurement Wh/m^2, which is equivalent to 3600 J/m^2, i.e., the amount of energy accumulated in 3600 seconds (one hour) The literature often refers to solar irradiation as **insolation**. Adopting the unit of Wh/m^2 has a positive effect on computing processes, since 1 W/m^2 is $Wh/(m^2.h)$, i.e., the hourly solar irradiation over each surface, in Wh/m^2, without any conversion factor.

Solar radiation is a measurement of energy, i.e., J. or Wh. It expresses the amount of electromagnetic energy that comes from the sun. Moreover, solar radiation may be decomposed into: **direct (beam) radiation**, which is the short-wave radiation coming from the solar rays, or **diffuse radiation**, which is scattered by the elements of the built environment.

When reporting solar analysis on building surfaces, it is often common to express the results in solar irradiation terms, i.e., Wh/m^2 or kWh/m^2. This has to do with the fact that smaller or larger surfaces may express drastically different results in terms of solar radiation and, therefore, normalising these values per square meter is a good indicator of the phenomenon under analysis.

Several related work deal with automatic methods for modelling solar irradiation in digital buildings. The following ones were investigated thoroughly to understand their potentialities and limitations.

LANGE (2015) adopts ESRI ArcGIS's Area Solar Radiation module when designing the web-based solar map of Rio de Janeiro. In their work, the author uses an official three-meter resolution raster-based digital terrain module (DTM) coming from Rio's local authorities, as well as the vector-based building cadastre coming from the same source, which were first updated by the researcher.

To achieve their final go, the researcher:

- patched multiple DTM tiles into a 'continuous' dataset and increased the DTM's original spatial resolution to one meter. The author does not mention if the last procedure was either oversampling or interpolation.
- elevated the building footprints to the height of the roofs, based on the value of the DTM increased by the height attribute of these buildings,
- rasterised the buildings and incorporated these to the DTM, generating a digital surface model (DSM).

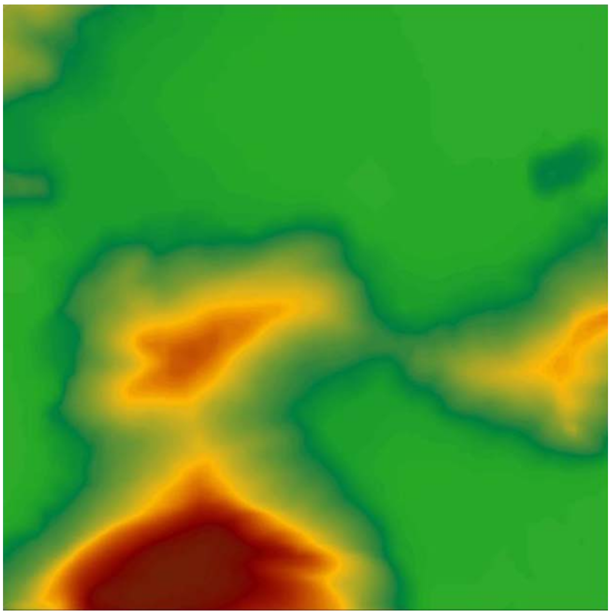


Figure 2-15 Sample of the input DTM of Rio de Janeiro.

Source: LANGE (2015), adapted.

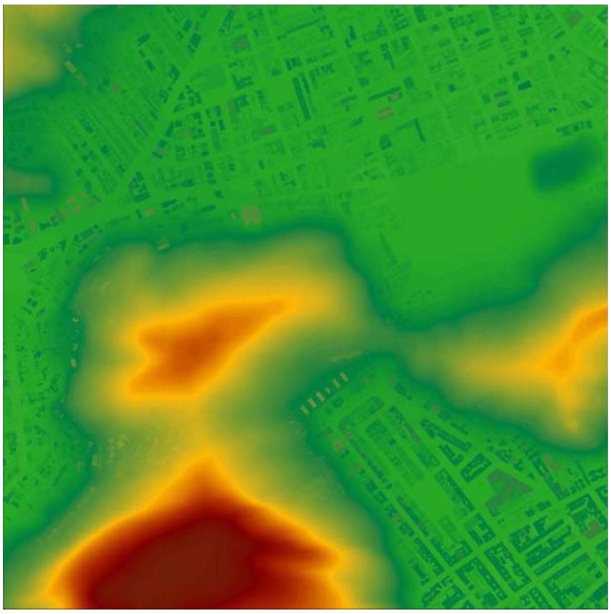


Figure 2-16 Sample of the output DSM of Rio de Janeiro.

Source: LANGE (2015), adapted.

The resulting DSM served as the input for the Area Solar Radiation model, which output global irradiation values at every pixel of the dataset. According to their report, the following parameters where adopted:

- Timeframe: daily analysis only for the summer solstice (December 22nd) and winter solstice (June 22nd)
- Diffuse irradiation rate, fixed at 0.4 according to the experiments conducted with the meteorological data captured at Santos Dumont Airport, in Rio.
- Solar irradiation transmittivity through the atmosphere, fixed at 0.5 according to the same experiments.

The author warns about how computationally expensive the module is when processing all input data. They opted, therefore, to fragment the original dataset into multiple tiles and process these individually, at the expense of ignoring possible shading effect from the surroundings. This poses a critical problem in the city of Rio de Janeiro, known for its dramatic hilly features in the landscape.

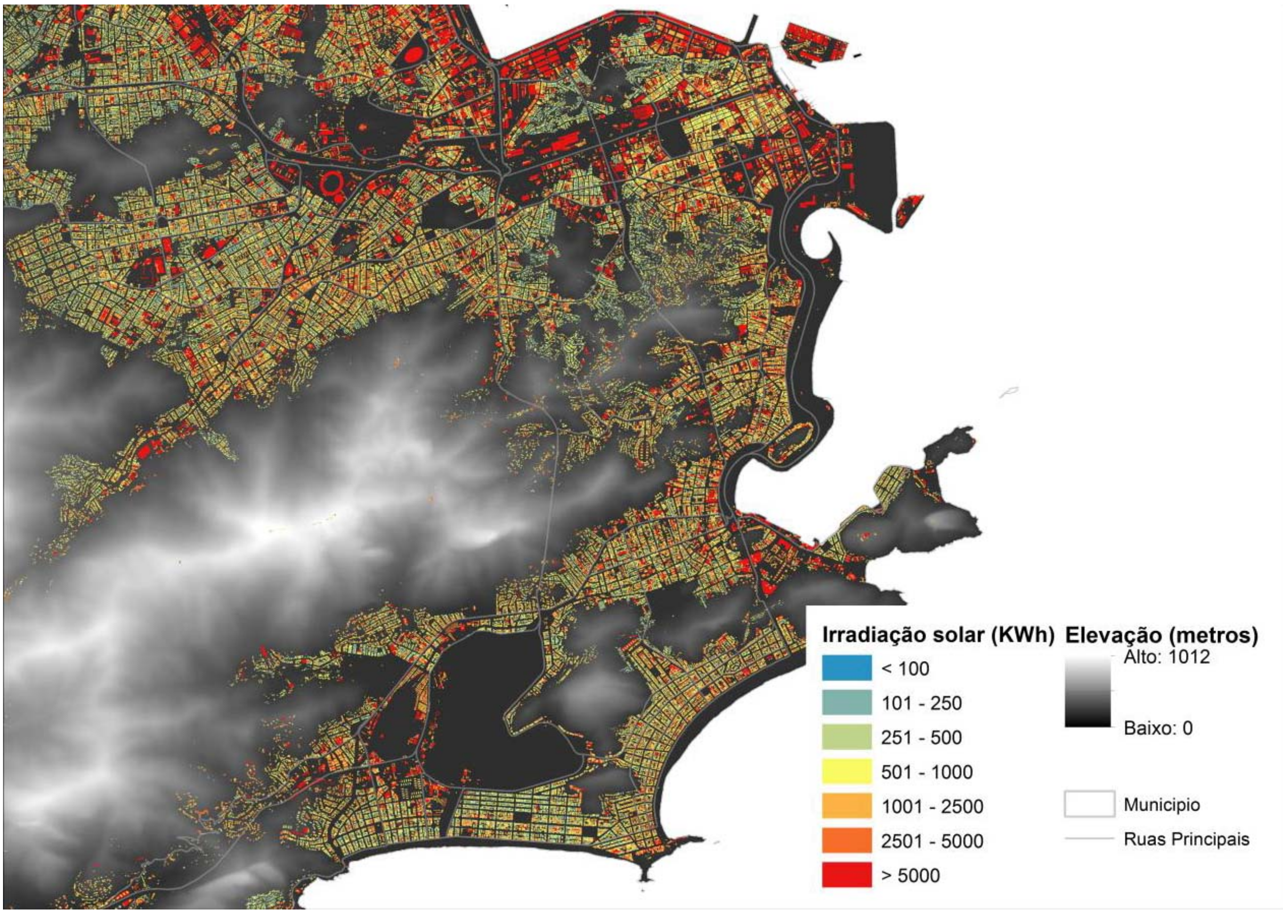


Figure 2-17 Output map of solar irradiation values computed per building roof for the 22nd of December.

The shading effect from the hilly landscape of Rio de Janeiro plays a significant role in the resulting irradiation values.

Source: LANGE (2015), adapted.

The shading effect caused by the relief in solar analysis is also investigated in AGUGIARO et al. (2012). The authors proposed a data pipeline to perform solar irradiation estimation at a city scale in an alpine environment, for which the surrounding mountains must be considered to achieve realistic results.

The authors combined a series of vector and raster-based datasets, among which a digital surface model (DSM), building cadastre datasets and meteorological data. These were pre-processed and integrated in a database environment to optimise the workflow and allow scalability for a large area of study.

Considering the necessity of a free-of-charge and open-source software that could handle an integration with the database, the authors adopted OsGeo GRASS GIS and its modules to process the data in multi steps. When addressing their methodology, the authors describe the following steps:

- Definition of the extents of the area to be analysed,
- query of the DSM tiles covering the area and a buffered zone,
- merge of these tiles into a unique dataset,
- computation of aspect (azimuth) and slope maps,
- computation of horizon maps, which inform the far field obstructions in the landscape,
- values of monthly air turbidity coefficients

These datasets allow the implementation of the r.sun module which, just like in the case of ArcGIS, also operates with a raster-based data model. According to the authors, “in total, 36 maps were obtained: for each month, direct, diffuse, and global irradiance maps were calculated, yielding the average daily irradiance value (in Wh/m^2) for each month. The yearly average values were also calculated.” It is worth mentioning that the simulated values are adjusted using measured data from a pyranometer.

The buildings are then classified according to the global solar irradiance values computed for their roofs, and the final material is presented in a web-based map, which allows the user to retrieve data regarding specific buildings.

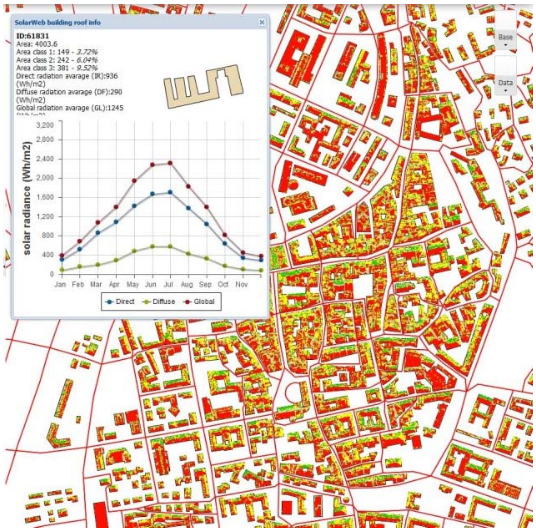


Figure 2-18 Screenshot from the WebGIS application with building footprints.

The roofs classified according to the incoming global solar irradiance (yearly average values), and monthly statistics.

Source. AGUGIARO et al. (2012).

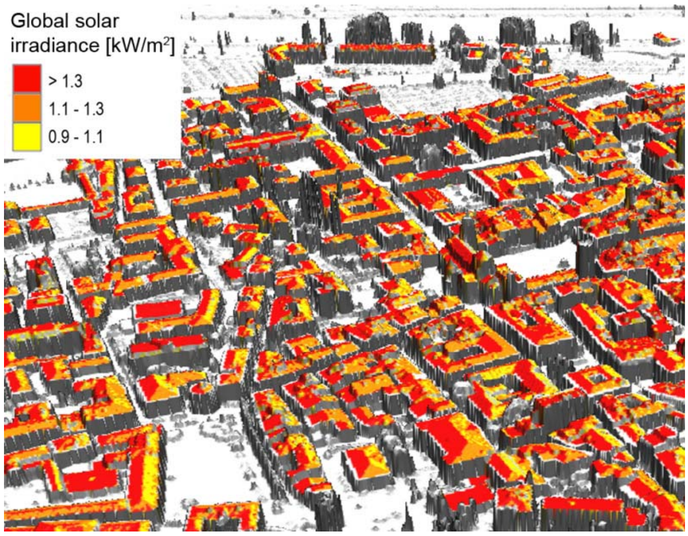


Figure 2-19 3D view of the study area showing the classified roofs mapped on top of the LiDAR DSM.

Source. AGUGIARO et al. (2012).

In addition to the abovementioned literature that reports solar analysis experiments, some authors were also considered for the research, among which: LEÓN-SÁNCHEZ (2013), KADEN et KOLBE (2014), WATE et COORS (2015). These authors explore the use of 3D city models for energy simulations in general, and their works were important to gain theoretical experience when it comes to handle data models and data structures.

Finally, it is also important to observe that attention has been paid to the documentation of each one of the solar analysis modules used in the present research, namely:

- OSGeo GRASS GIS – module r.sun
- ESGI ArcGIS – module area solar radiation
- EPFL CitySim Solver – KAEMCO CitySim PRO
- HFT Stuttgart SimStadt – solar potential analysis
- LadyBug Tools Ladybug – incident radiation
- Virtual City Systems – 3D Solar Potential Analysis

METHODOLOGY, EXPERIMENTAL DESIGN & DEVELOPMENT

3 METHODOLOGY, EXPERIMENTAL DESIGN & DEVELOPMENT

3.1 Method and Technical Pipeline

The research question of the present MSc thesis is: **“How far is it possible to perform solar analysis on buildings of favelas in São Paulo, with the goal of estimating PV Potential?**”

To respond the research question, two topics are defined, namely **roof mapping** and **solar irradiation modules**. The former deals with the feasibility of reconstructing building roofs in favelas from existing spatial datasets of São Paulo. The latter deals with the performance of existing solar irradiation modules when assessing buildings of favelas in São Paulo.

The implemented **method** is, in both topics, **comparison**, which is a strategy for assessing several workflows and their achieved results. In this sense, two main comparison strategies are established, as it can be observed in the chart below, which contains the general pipeline of this MSc thesis. Detailed information about each part of the pipeline is available in their corresponding sections.

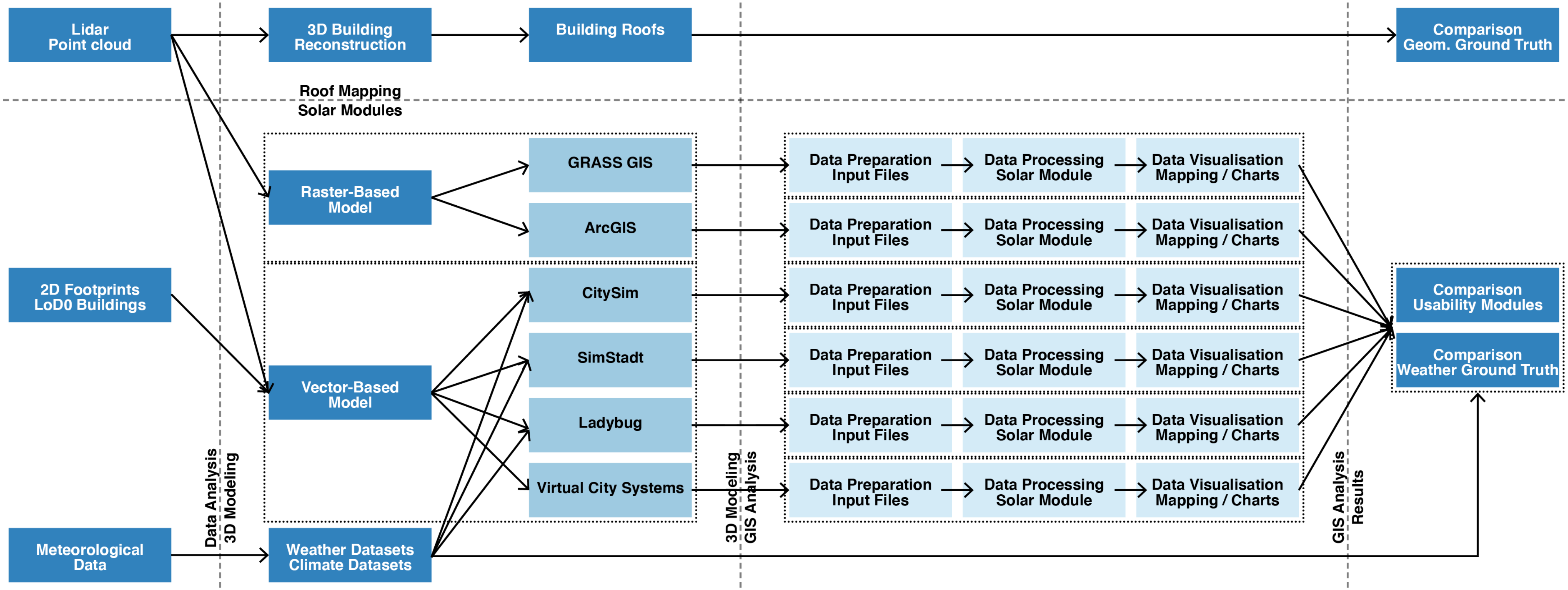


Figure 3-1 MSc thesis – general pipeline.

Firstly, regarding the **roof mapping**, it is possible to make an initial assumption that a pseudo LoD2 is sufficient to model the environment of a favela. In this sense, roofs are designed as flat surfaces, as in the case of LoD1 but their geometries are semantically enriched and classified as in LoD2. By projecting the output 3D roofs coming from a building reconstruction algorithm into the XY plane, the resulting 2D geometries can be compared to the official building footprints of São Paulo. Albeit most favelas in São Paulo lack cadastral data, some of these environments are already surveyed. Therefore, if an informal settlement with an available cadastre dataset is chosen, a first comparison can be established between what is understood as ground truth – the building footprints offered by the municipality – and the resulting geometries from the reconstruction algorithm.

Within the framework of this MSc thesis, the 3D Building reconstruction algorithm that is adopted is the one used for the 3D Building Cadastre of the Netherlands, implemented by the 3D Geoinformation Group[13]. This algorithm was slightly adapted by the original authors to accept the dataset of São Paulo. Details about this technical procedure are available in the roof mapping subsection.

When it comes to the **solar modules**, a broader comparison schema can be settled among the irradiation values presented by the modules themselves, but also with respect to values delivered by a meteorological station, which is in turn considered to be ground truth.

In this sense, the comparison schema consists of modelling two scenarios, i.e., one for a favela and another one for the area that hosts the meteorological station of São Paulo. Depending on the solar module under analysis, either a vector-based or a raster-based data city model is necessary, which means that, in total, four datasets need to be designed.

For illustrative purposes, the final version of these datasets are hereby presented, and detailed information can be found in the sub-sections dedicated to Brasilândia, Santana, CityGML and Raster data models, as well as the section in which their extensions are explained.

---

[13] 3D BAG – 3D Geoinformation Group at TU Delft

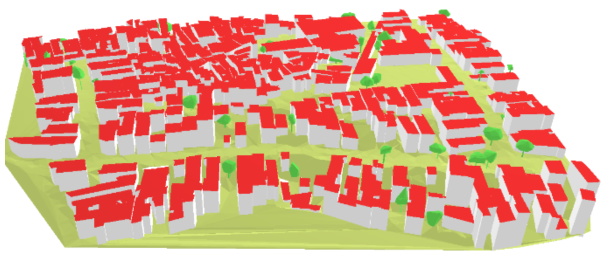


Figure 3-2 Favela scenario, vector-based model.

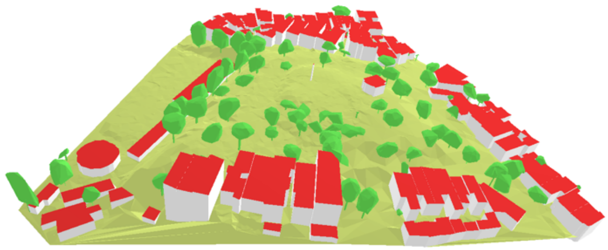


Figure 3-3 Meteorological station scenario, vector-based model.

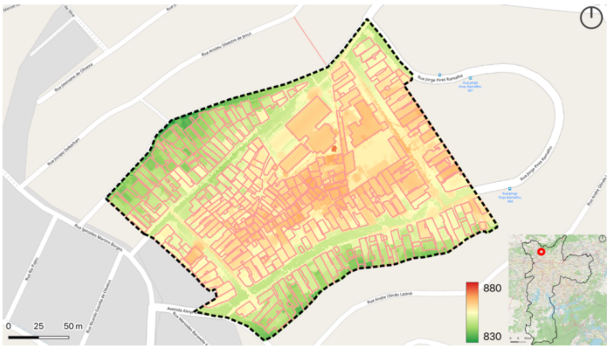


Figure 3-4 Favela scenario, raster-based model.

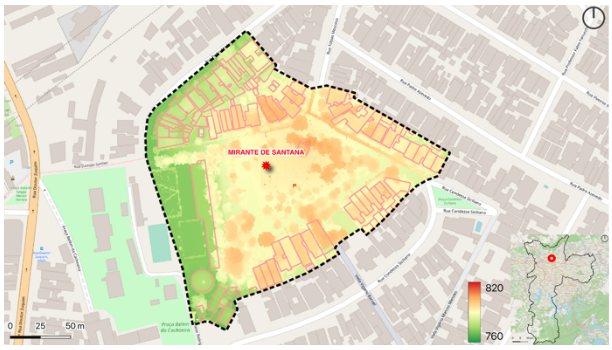


Figure 3-5 Meteorological station scenario, raster-based model.

With the meteorological station city models ready, for each solar module, a time series of irradiation values is extracted using the coordinates of this meteorological station. The monthly differences between ground truth and the simulation values coming from the software are computed, and later applied as correction factors for all irradiation values extracted from the favela city models. This procedure allows a comparison among the level of accuracy for each solar module when assessing solar irradiation on buildings of favelas. Moreover, by responding the qualitative sub-research questions of this topic, these solar modules are compared also in terms of their performance and usability in the data workflow.

It is important to stress that, during the research, the roof mapping topic and the solar irradiation topics are handled apart from each other. This principle is adopted so that the solar modules can be run and assessed while the investigation on roof mapping is not completed. Moreover, by following this strategy, the favela scenario is not affected by the quality of the reconstructed building roofs. In this sense, the favela scenario is modelled using the available building footprints from the city hall of São Paulo.

3.2 The Spatial Data Infrastructure of São Paulo

The starting point for designing the city models is retrieving the appropriate datasets. Within the framework of the spatial data infrastructure of São Paulo, the two main sources of data are the *Mapa Digital da Cidade de São Paulo* – Portuguese for Digital Map of the City of São Paulo – popularly known as GeoSampa, and the database of *Instituto Nacional de Meteorologia* – Portuguese for National Institute of Meteorology, also known as INMET.

Geosampa is the database and web-viewer from which all municipal publicly available geodata and metadata can be visualised, queried, and downloaded. It hosts a collection of 353 spatial datasets[14], which are clustered in thematic groups and sub-groups.

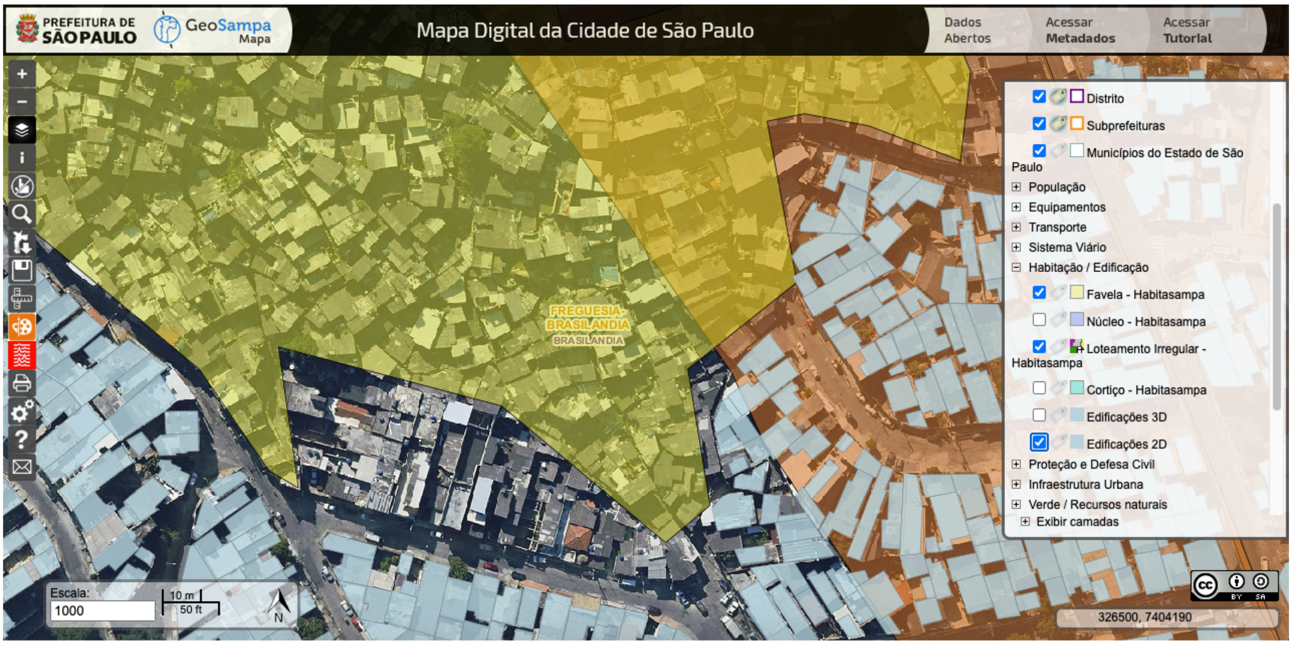


Figure 3-6 Digital map of the city of São Paulo – Geosampa.

Source: Geosampa.

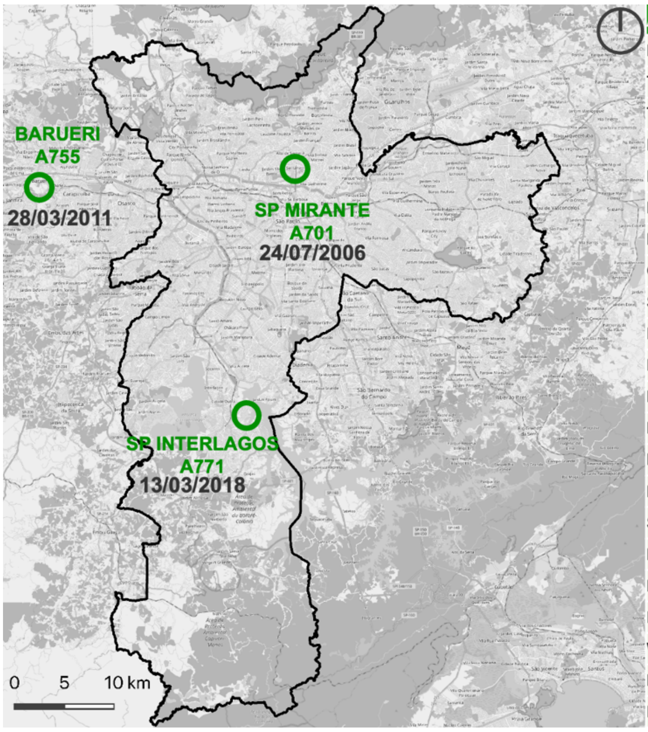
For this research, fragments of the following five datasets are used. The metadata regarding each one of these is present in the Appendix A of the research.

- Favelas (Habitasampa). A 2D polygon shapefile with attributes: id, name, address, number of households, total area, year of origin and land ownership.
- Irregular land parcelling (Habitasampa): A 2D polygon shapefile with attributes: id, name, number of plots, area, typology and other.

[14] As of September 7th, 2021. The current metadata list is constantly updated.

- Building 2D: A 2D polygon shapefile file with building attributes: id, area, and height. By analysis, it is possible to affirm that the height information is not always trustable. Some existing buildings have a height value of zero, while others have height values far away from reality.
- DSM: A LAZ-encoded point cloud file containing RGB info and ASPRS[15] lidar class values, namely: relief, vegetation, buildings, and other horizontal / vertical features.

The meteorological datasets, in turn, are retrieved from the database of INMET, a public body under the Ministry of Agriculture. The institution holds datasets for all 587 meteorological stations across Brazil, which are then subdivided into conventional stations and automatic stations. The following pictures situate these stations in the city of São Paulo and offer a preview of the data offered by these:



AUTOMATIC METEOROLOGICAL STATIONS - DATA
DATE (DD-MM-YYYY)
TIME UTC
TOTAL HOURLY PRECIPITATION (mm)
HOURLY ATMOSPHERIC PRESSURE AT STATION LEVEL (mB)
HOURLY ATMOSPHERIC PRESSURE AT SEA LEVEL AUT(mB)
MAXIMAL ATMOSPHERIC PRESSURE OF THE PREVIOUS HOUR (AUT)(mB)
MINIMAL ATMOSPHERIC PRESSURE OF THE PREVIOUS HOUR (AUT)(mB);
GLOBAL RADIATION (KJ/m²);
STATION CPU TEMPERATURE (ºC)
HOURLY DRY-BULB TEMPERATURE (ºC)
DEW POINT TEMPERATURE (ºC)
MAXIMUM AIR TEMPERATURE OF THE PREVIOUS HOUR (AUT)(ºC)
MINIMUM AIR TEMPERATURE OF THE PREVIOUS HOUR (AUT)(ºC)
MAXIMUM DEW POINT TEMPERATURE OF THE PREVIOUS HOUR (AUT)(ºC)
MINIMUM DEW POINT TEMPERATURE OF THE PREVIOUS HOUR (AUT)(ºC)
STATION'S BATTERY VOLTAGE (V)
MAXIMUM RELATIVE HUMIDITY OF THE PREVIOUS HOUR. (AUT)(%)
MINIMUM RELATIVE HUMIDITY OF THE PREVIOUS HOUR. (AUT)(%)
HOURLY RELATIVE HUMIDITY (%)
WIND DIRECTION (AZIMUTH º)
MAXIMUM WIND BLOW (m/s)
HOURLY WIND SPEED (m/s)

Figure 3-7 Automatic meteorological stations in the city of São Paulo and surroundings.

Values are available at hourly intervals, and the dates in the map represent the start of the records. Timestamps must consider the local time in Eastern Brazil, and therefore a reduction of three hours must be performed (UTC-3).

Data source: Geosampa, Open Street Map, INMET.

[15] American Society of Photogrammetry and Remote Sensing. More information regarding the Lidar point cloud classification can be found at The LAS 1.4 Specification manual.

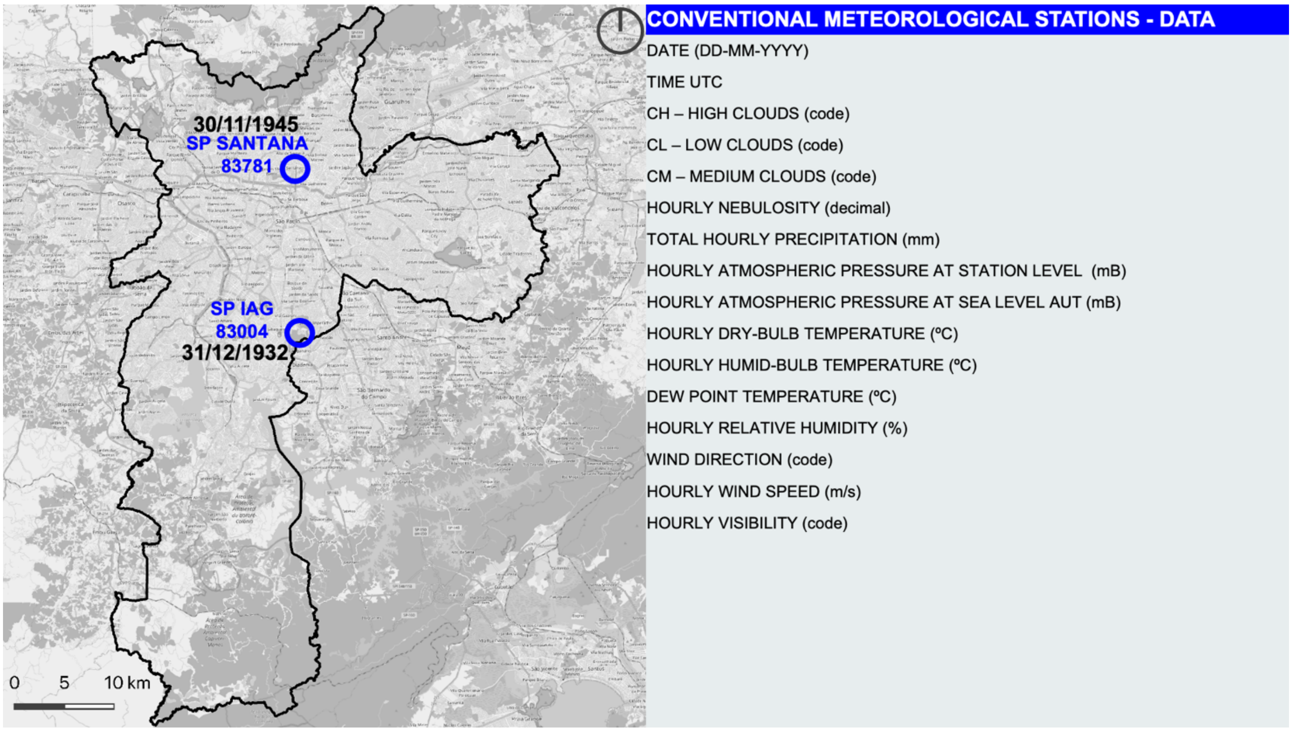


Figure 3-8 Conventional meteorological stations in the city of São Paulo.

Values are available three times per day, and the dates in the map represent the start of the records. Timestamps must consider the local time in Eastern Brazil, and therefore a reduction of three hours must be performed (UTC-3).

Data source: Geosampa, Open Street Map, INMET.

3.3 Scenario 1 – Brasilândia

Once the method, technical pipeline and datasets are defined, a successive task consists of defining the two scenarios that are modelled in this research, the first one of them being the favela scenario. In correspondence with a senior researcher[16] of the Faculty of Architecture and Urbanism – FAU – of the University of São Paulo – USP[17], some reference literature regarding environment and energy studies in favelas was offered as a starting point.

From GUSSON (2014), the first scenario for the present MSc thesis is selected. One of the sites that was assessed by this researcher is a small-scale city block located in the district of *Brasilândia*, north-western São Paulo. The area is bounded by four local streets, totalizing 21.794 square meters in moderate terrain slope, and it contains 168 buildings.

[16] Denise Duarte, full professor at FAU USP and former head of the Environment and Energy Studies Lab – LABAUT – within the same institution.

[17] The University of São Paulo is the alma mater of the author.

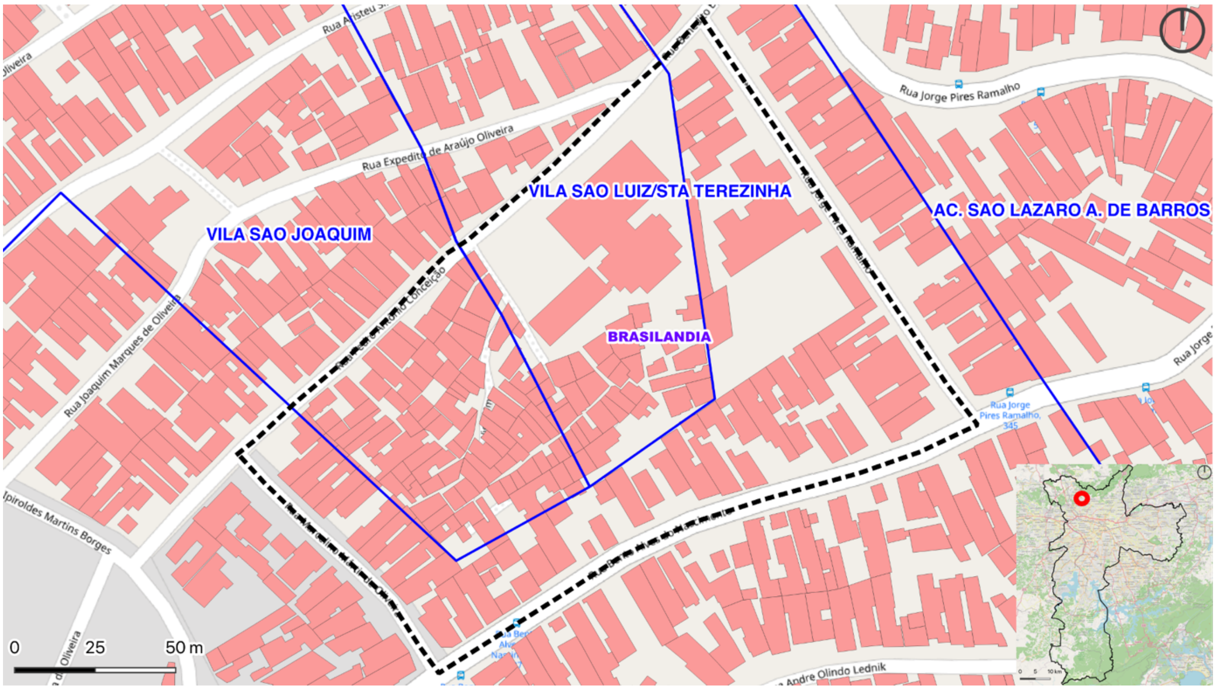


Figure 3-9 Scenario 1. The built environment and land ownership of Brasilândia city block in São Paulo.

The block perimeter is delimited by a black dot line, and the 2D footprints (light pink) contain building attributes, such as their ids. The irregular land parcelling perimeter (blue) of *Vila São Joaquim* intersects this city block, and the urban morphology of this site is typical of a favela.

Data source: Geosampa, Open Street Map.

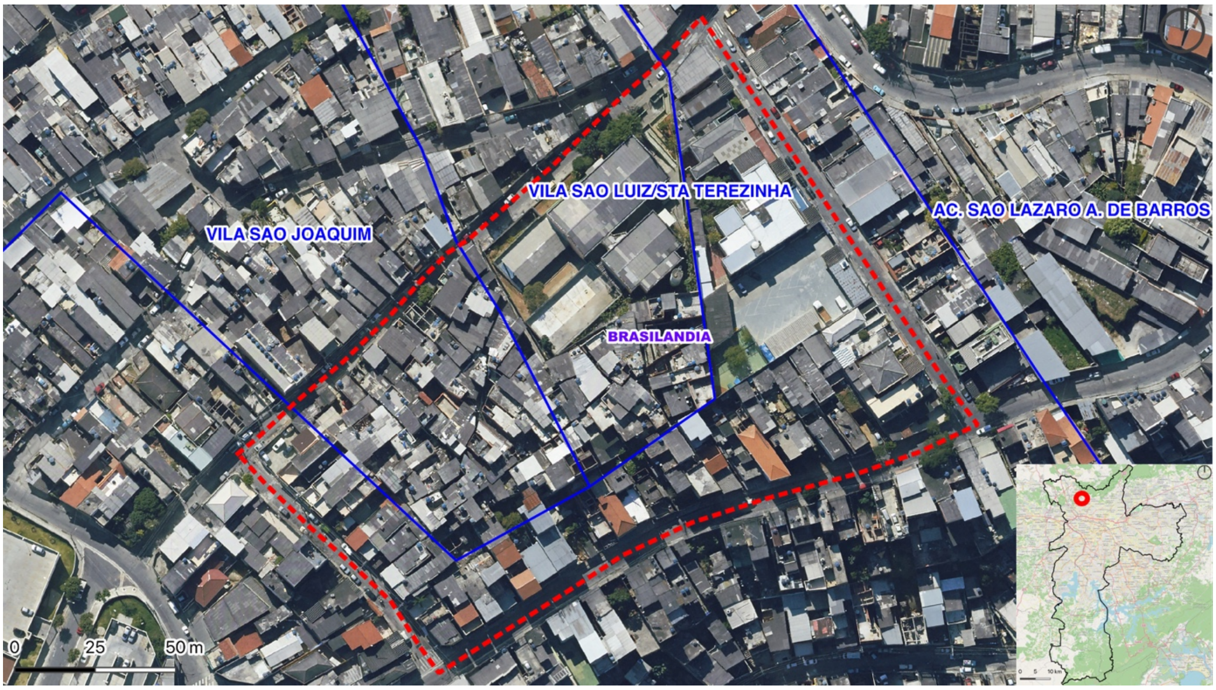


Figure 3-10 Scenario 1. Orthophoto of Brasilândia city block in São Paulo.

The visual interpretation of the orthophoto evidences a high density of buildings within the city block, the narrow and unorganized alleys in the urban fabric, unclear boundaries among dwellings and a sparse presence of high vegetation. All these characteristics are typical of favelas, despite this site is not officially considered to be one.

Data source: Geosampa, Open Street Map.

Figure 3-11 South-West corner of Brasilândia City Block.

Source: Google Street View.

Figure 3-12 South-East corner of Brasilândia City Block.

Source: Google Street View.

Figure 3-13 North-East corner of Brasilândia City Block.

Source: Google Street View.

Figure 3-14 North-West corner of Brasilândia City Block.

Source: Google Street View.

It is opportune to emphasise that this settlement is not classified as a favela from a legal perspective[18], and most of its extension is rather classified as an irregular land parcelling. Nevertheless, from a morphological point of view, this block presents characteristics that are typical of favelas, including: a high density of buildings within the city block, the narrow and unorganized alleys in the urban fabric, unclear boundaries among dwellings and a sparse presence of high vegetation. Therefore, it is possible to model this site as a scenario that represents a favela. The advantage of doing so is that, for this specific location, the 2D building footprint dataset is available, and so the methodological requirements for the research are met.

3.4 Scenario 2 – Santana

The second scenario that needs to be modelled is the one containing one of the meteorological stations within the city of São Paulo, so that a comparison with ground truth

[18] According to the Master Plan of the City of São Paulo (2014).

values can be established. The chosen site is the one in the surroundings of the meteorological station *Mirante de Santana*, located in the northern part of the city.

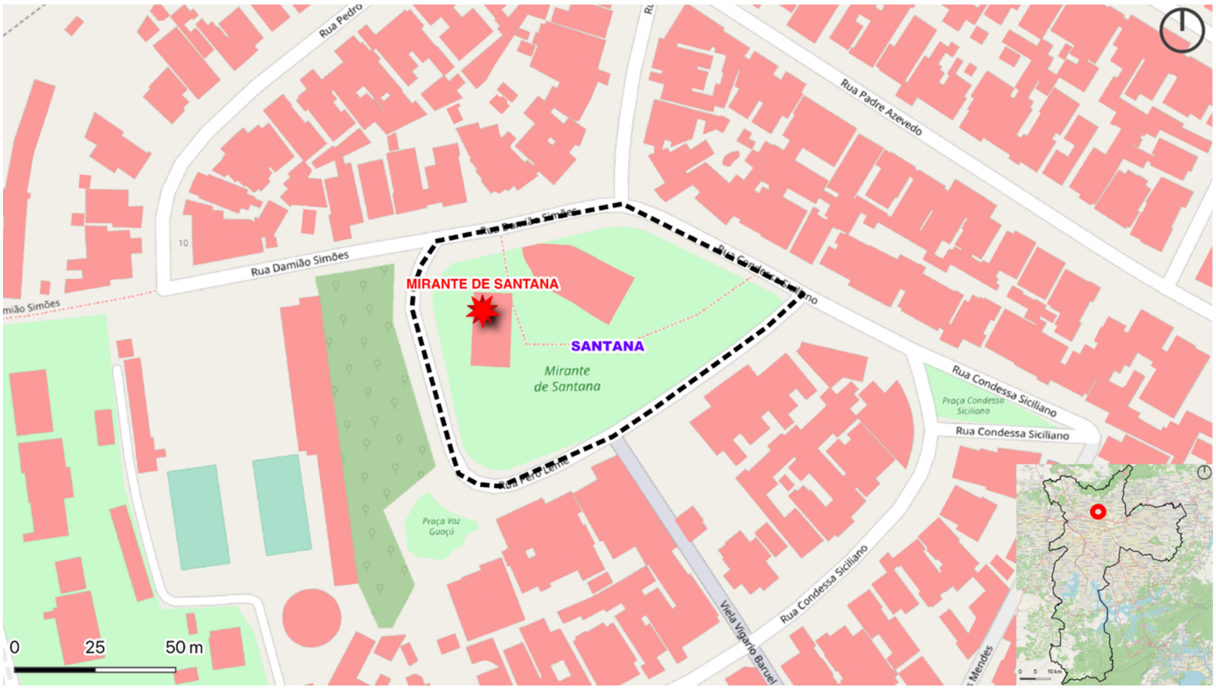


Figure 3-15 Scenario 2 The built environment of Santana city block in São Paulo.

The surrounding buildings are low and therefore do not produce significant shading to the meteorological station.

Data source: Geosampa, Open Street Map.

Figure 3-16 Scenario 2. Orthophoto of Santana city block in São Paulo.

The visual interpretation of the orthophoto evidences high vegetation in the surroundings, which must be considered in the solar irradiation analysis.

Data source: Geosampa, Open Street Map.

Selecting the meteorological station of Santana is positive for two main aspects:

- Firstly, the city block under analysis hosts not one but two meteorological stations, an automatic one and a conventional one. Albeit most necessary data for the present research come from the automatic meteorological station, cloud nebulosity data are only present in the conventional one. Therefore, a combination of data coming from these two stations makes the experiments feasible.
- Secondly, the city block of *Brasilândia* lies only nine kilometres away from *Mirante de Santana*, which is considered a short distance, given the scale of São Paulo. Therefore, applying a correction factor to the solar irradiance values in Brasilândia, based on the discrepancy between computed and ground truth values in *Santana*, is a safe assumption for the research.

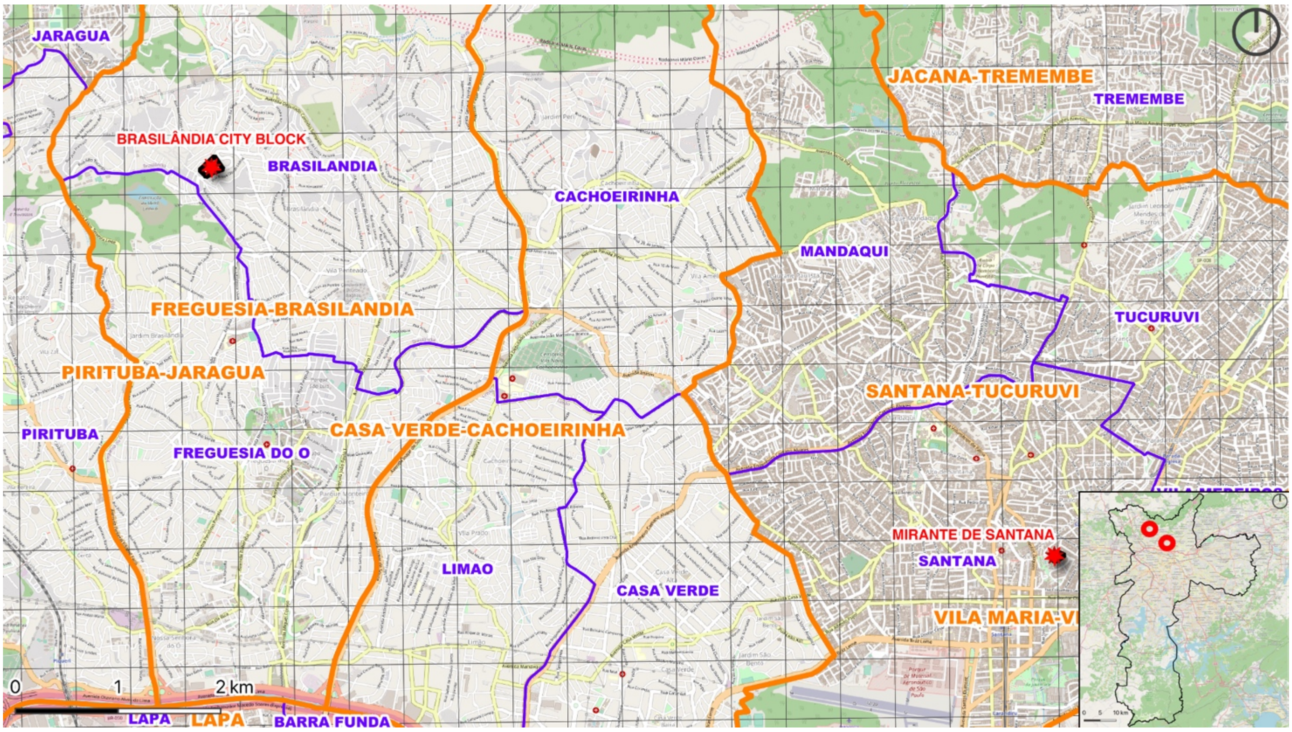


Figure 3-17 Distance between Scenario 1 Brasilândia and Scenario 2 Santana.

Administrative boroughs are represented in orange and districts in purple. The black rectangles correspond to the extensions of the point cloud tiles, approx. 500mx500m each. Both scenarios, represented with red dots, lie in the northern portion of São Paulo, in between the *Tietê* river basin (to the south) and the *Cantareira* hills (to the north).

Data source: Geosampa, Open Street Map.

3.5 Extending the scenarios

For the purpose of solar analysis, defining the geographical extensions of a city model is a two-way process, which requires an iteration until an optimal balance between accuracy and computing efficiency is achieved.

On the one hand, the larger the model is, the more realistic the shading effect will be, especially due to the relief. This is due to the digitalisation of buildings, trees, and relief elements of the surrounding environment, which will act as obstruction elements for the sun rays. In this sense, surfaces under analysis should not be considered by themselves, but also considering at least some context in which they belong to.

On the other hand, increasing the dimensions of the digital model results in increasing the number of geometries that are introduced in the model. Therefore, the computing efficiency will decrease, in a rate that varies according to the level of complexity of the algorithm itself. It is possible to argue, however, that simplifying the geometries may mitigate this downside effect. In addition, depending on the flexibility of the solar module, it is possible to segregate geometries between those which will be assessed and those which will only act as shading surfaces.

Still, striking a balance between the city model extension, the complexity of its geometries and its accuracy in terms of solar irradiation is only possible by a hands-on investigation. This process is somehow straightforward when one verifies the potential and limitations of one solar module, but much more complex when multiple modules, each one with its intrinsic demands, are confronted.

Therefore, defining the extension of the two scenarios – Brasilândia and Santana – is a dominant issue during the timeframe of this project.

In AGUGIARO et al. (2012), the authors establish that, considering the alpine environment, all DSM tiles within a buffer of 5 Km from the study area should be considered, even if in a lower spatial resolution. Albeit there are no mountain ranges in São Paulo, the proximity of both scenarios to the Cantareira hills had to be at least put into consideration. Therefore, a sequence of five buffers – one kilometre apart from each other – was set from scenario 1.

From this experiment, it was firstly imagined that a set of twenty-five tiles – 5 x 5 – would be a good compromise.

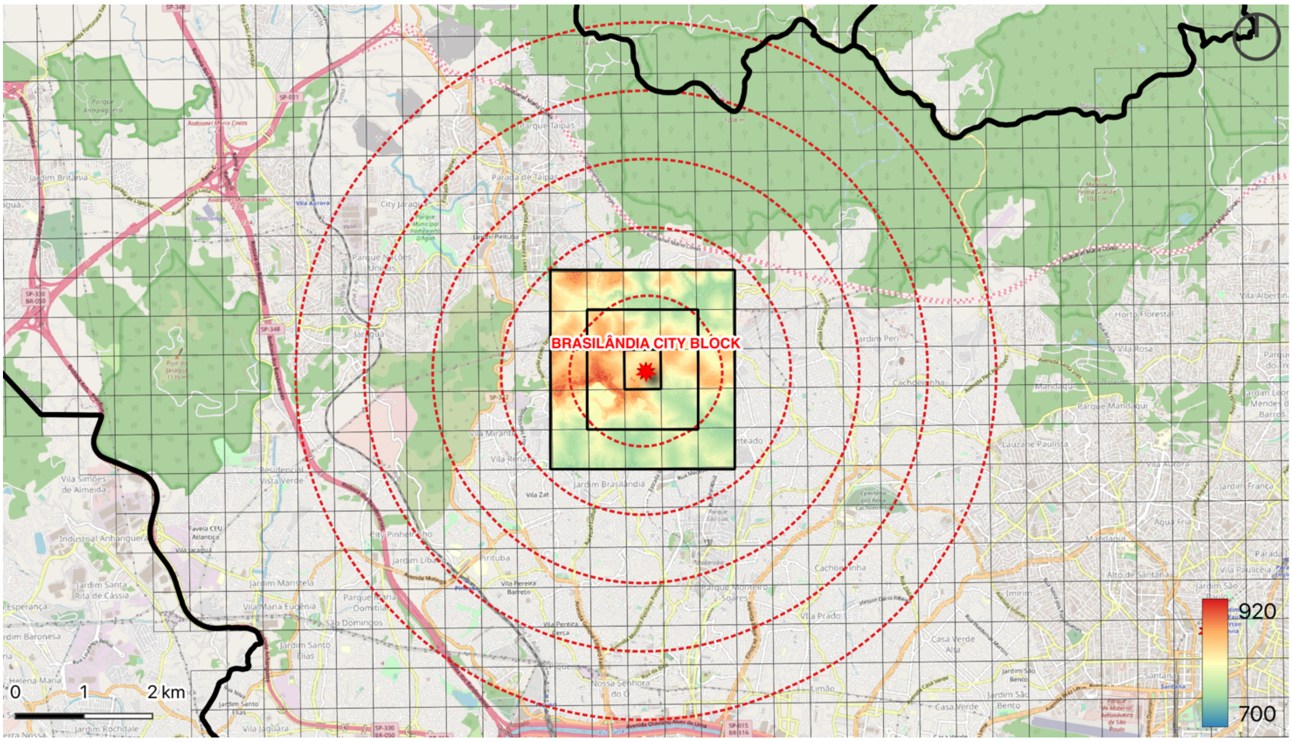


Figure 3-18 Buffer analysis for Scenario 1- Brasilândia.

A sequence of five buffers – one kilometre distant from each other – and the resulting DSM of a 5x5 scenario extent.

Data source: [Geosampa](), [Open Street Map]().

Despite computing a one-meter resolution 5x5 DSM successfully, this strategy results to be extremely computing inefficient. Further experiments run on a raster-model environment, such as ArcGIS, do not respond the solar analysis task. Even if it is possible to reduce the spatial resolution of the raster file, the same scaling effect would occur in a vector-based model.

In addition, for the purpose of methodological consistency, the same designing principles should apply to all solar modules under analysis, including the extensions of the scenarios. Therefore, an interactive process is conducted by testing the performances of these solar modules. In synthesis, it was decided to adopt a minimum extension criterion. With this strategy, the perimeters of Brasilândia and Santana are extended to incorporate the land plots that are adjacent to the streets that define such perimeters. In other words, the scenarios embrace what is on the other side of the streets.

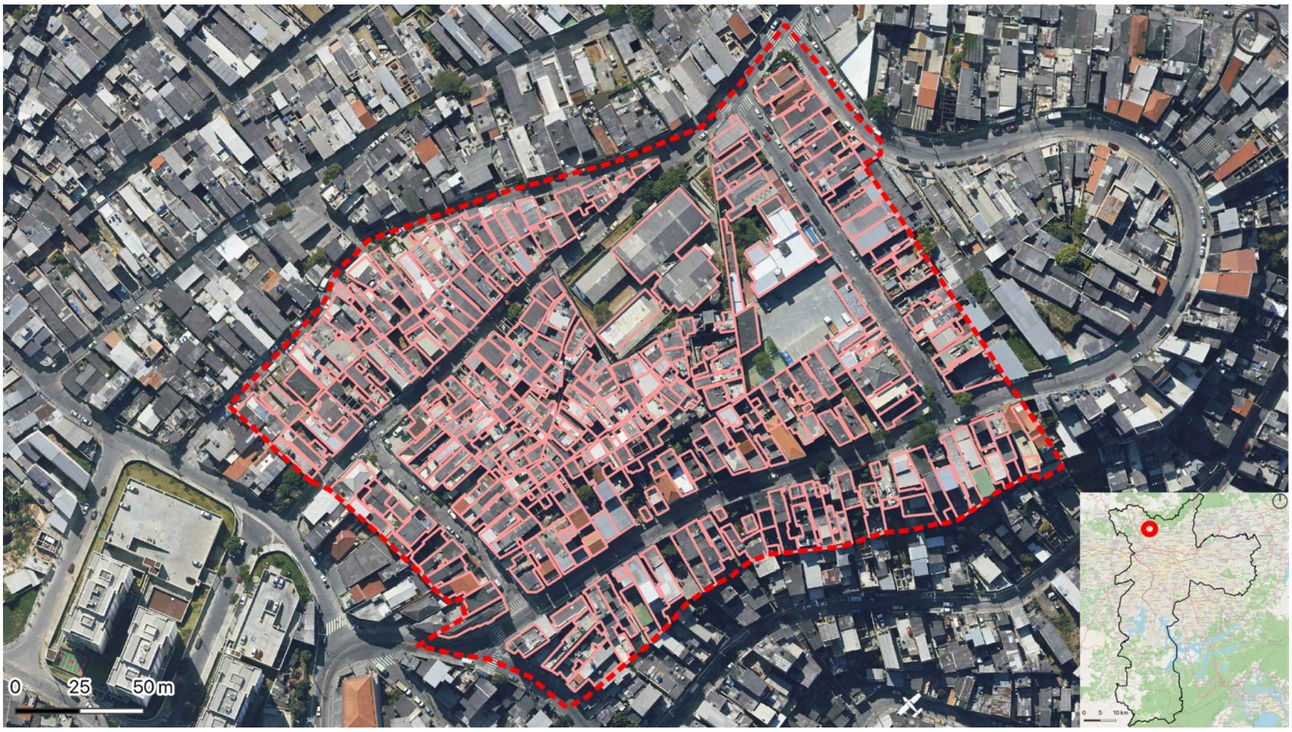


Figure 3-19 Scenario 1 - Brasilândia city block extended.

The extended scenario gathers 301 buildings within its perimeter, which results in an area of 43.883,58 m².

Data source: [Geosampa](), [Open Street Map]().

Figure 3-20 Scenario 2 - Santana city model extended.

The extended scenario gathers 100 buildings within its perimeter, which results in an area of 33.822,49 m².

Data source: [Geosampa](), [Open Street Map]().

The decision of limiting the extension of the scenarios decreases the level of accuracy achieved in the solar analysis until a certain point, but it also makes a fair comparison viable among all six solar irradiation modules. The only exceptions to this rule are the experiments conducted on the solar module of Virtual City Systems. In this case, the staff of the company adopts larger city models that were designed by the author prior to the establishment of the minimum extension criterion.

3.6 Vector Model – CityGML

Designing Santana and Brasilândia in a digital environment demands both a data model and a data structure that are compatible with the input demands of the software analysed in this research. According to ARROYO, LEDOUX et PETERS (2021, 2):

> "A **data model** is a high-level formalised way to structure information, generally using a set of abstract classes, relationships between them, and attributes to store information about them. In the context of geomatics, these classes are often spatial representations of real-world objects. (…) The typical examples of data models used in (…) geomatics literature are the raster and vector data models." ARROYO, LEDOUX et PETERS (2021, 2)

> "A **data structure** is a low-level description that specifies how to implement a data model, or occasionally a combination of multiple data models. Data structures are defined with little to no ambiguity, specifying features such as what sort of storage should be used for a given primitive (e.g. an array or a linked list). As opposed to a data model, creating a computer implementation of a data structure is thus relatively straightforward, and different people implementing the same data structure will end up with very similar implementations." ARROYO, LEDOUX et PETERS (2021, 2)

In the context of this MSc thesis, four out of six solar irradiation modules demand, as their geometrical input, a vector-based data model under the standard of CityGML 2.0, which is the most standardized and widely spread schema within the geo-information community. According to the Open Geospatial Consortium (OGC), an international consortium responsible for creating free-of-charge and publicly available open geospatial standards,

> “CityGML is an open data model and XML-based format for the storage and exchange of virtual 3D city models. It is an application schema for the Geography Markup Language version 3.1.1 (GML3), the extendible international standard for spatial data exchange issued by the Open Geospatial Consortium (OGC) and the ISO TC211. The aim of the development of CityGML is to reach a common definition of the basic entities, attributes, and relations of a 3D city model. This is especially important with respect to the cost-effective sustainable maintenance of 3D city models, allowing the reuse of the same data in different application fields.” (OGC, 2012, ix).

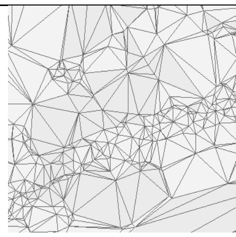
Therefore, another important task of this research is defining a data pipeline that creates XML-encoded CityGML models based on the input datasets of São Paulo, namely the building footprints and point cloud. The output of this pipeline results in well-structured files that can be automatically read by the software under analysis in this thesis.

The designing experiments are all conducted on a Safe Software FME Workbench, which is a software that allows extract, transform and load (ETL) operations with geo datasets. Without detailing the low-level implementation of the workbench proposed in this research, some high-level designing principles are worth of highlighting, especially for readers who are eventually not familiarised with the CityGML data model. These designing principles are grouped according to the three main feature types present in the output scenarios: relief, buildings, and trees.

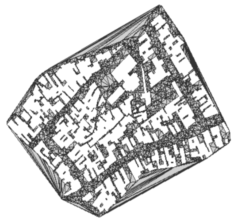
Relief

Starting from the original point cloud, this dataset is semantically filtered by the ground class, and spatially filtered to keep points that lie inside the scenario extents but outside any of the building footprints.

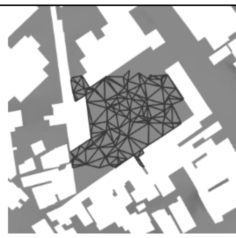
A first surface modeler defines an initial Triangular Irregular Network – TIN – and drapes the building footprints to this initial relief. A second surface modeler defines another TIN based on the points of the first TIN but then constrained by draped footprints that act as breaklines.



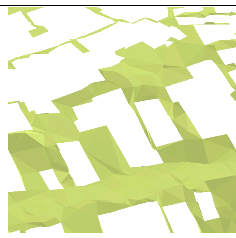
With the second modelled relief, its resulting surfaces are spatially filtered to keep only triangles that lie outside of the projected building footprints. The result is a set of triangles that represent the ground features of the model.



These ground features are enriched with some geometrical information, such as slope and azimuth angle, and then segregated into multiple tiles in a grid structure. The idea behind this strategy is to allow the importing / exporting of these tiles independently, reducing computing demand.



Each group of surfaces becomes one relief component, and these components become part of a relief feature. The CityGML features corresponding to the relief, relief components, triangulated surfaces and polygons are named accordingly, and stored in the CityGML model.



Buildings

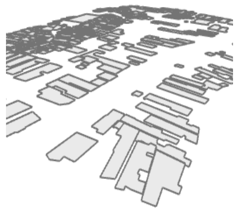
Starting from the original point cloud, this dataset is thematically filtered by the building class and its points are spatially filtered – and marked – according to the building footprint which they fall into.

A statistical analysis returns the median height value of the points belonging to the same building. Median is used since buildings in Brazil usually count on water tanks and other equipment in the roofs. Median is also preferred to avoid outliers, in this case wall points.

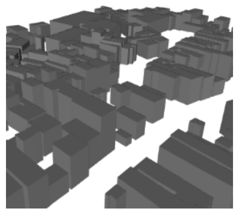
	gml_id	bld_gml_id	z.median
1	Building_1073...	Building_1073...	852.83
2	Building_1073...	Building_1073...	857.72
3	Building_1072...	Building_1072...	862.59
4	Building_1071...	Building_1071...	858.09
5	Building_1071...	Building_1071...	862.79
6	Building_1072...	Building_1072...	848.060000...
7	Building_1071...	Building_1071...	865.04
8	Building_1071...	Building_1071...	863.4
9	Building_1071...	Building_1071...	860.57
10	Building_1071...	Building_1071...	857.33
11	Building_1071...	Building_1071...	859.09
12	Building_1071...	Building_1071...	858.89

The polygonal footprints are elevated according to the retrieved roof height, which result in flat roofs. Likewise, the footprints are also passed as breaklines to conform the relief to their extents and dragged to the terrain. The lowest point minus 0.1m returns the height of the ground.

For each building, the difference between the flat roof and the flat ground returns the building height. If this height is lower than 3m, the roof is elevated from the flat ground up to three meters, thus resulting in a minimum three-meter height for all buildings.

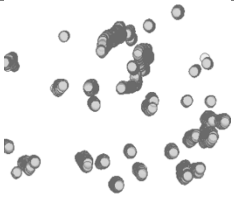
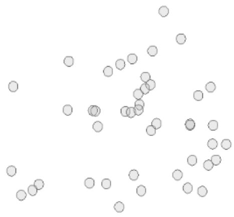
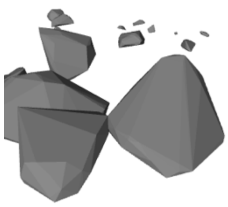
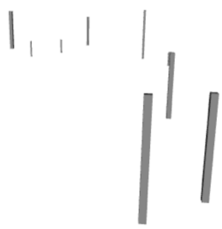


With all geometrical values retrieved, the flat grounds are extruded with the building height, and the resulting boundary representation surfaces are split semantically into ground, walls, and roofs, based on their slope values.



The CityGML features corresponding to the buildings and thematic surfaces are named accordingly – building, thematic surface, multi surface and polygon – and some computed geometrical attributes, such as area, volume, height, etc., are passed as CityGML generic attributes.

Trees

Starting from the original point cloud, this dataset is semantically filtered by the vegetation class and spatially filtered to keep only 3D points that lie inside the limits of the scenario but outside the projection of any building footprint into the XY plane.	
These several remaining points are buffered by a fixed five-meter searching radius and, for each one of these points, a list is created containing all the points that fall into their corresponding buffered circle.	
A Boolean test assesses, for each one of these points, if the point under analysis is the highest within its sub-group of points. If this is the case, the point is considered the uppermost part of the crown of a tree, and therefore individual vegetation objects – trees – are segregated.	
To eliminate vegetation points that belong to the trunks of these trees, a threshold is applied. For each tree, the height statistics of the clustered points eliminates any point that lies below their height average minus two times the standard deviation of the height of these points.	
The filtered groups of points are then turned into vector surfaces by defining a 3D convex hull of the remaining points. The mesh is re-ordered so that the normal vector of their surfaces points towards the outside, and the tree crown surfaces are marked as such.	
From the 3D centroid of the crown, the tree trunk is designed as a parallelepiped extruded towards the relief with a squared-horizontal section of 0.25m. A squared section is much preferred since it drastically reduces the number of surfaces in the trunk, which would have insignificant impact in the solar analysis.	
The CityGML features corresponding to the trees and tree parts are named accordingly – trees, multi surfaces, polygons – and some computed geometrical attributes, such as area, volume, height, etc. are passed as CityGML generic attributes.	

The resulting vector-based scenarios – Brasilândia and Santana – can be visually confronted with the original point cloud datasets.

Figure 3-21 Scenario 1 - Brasilândia - Input point cloud.

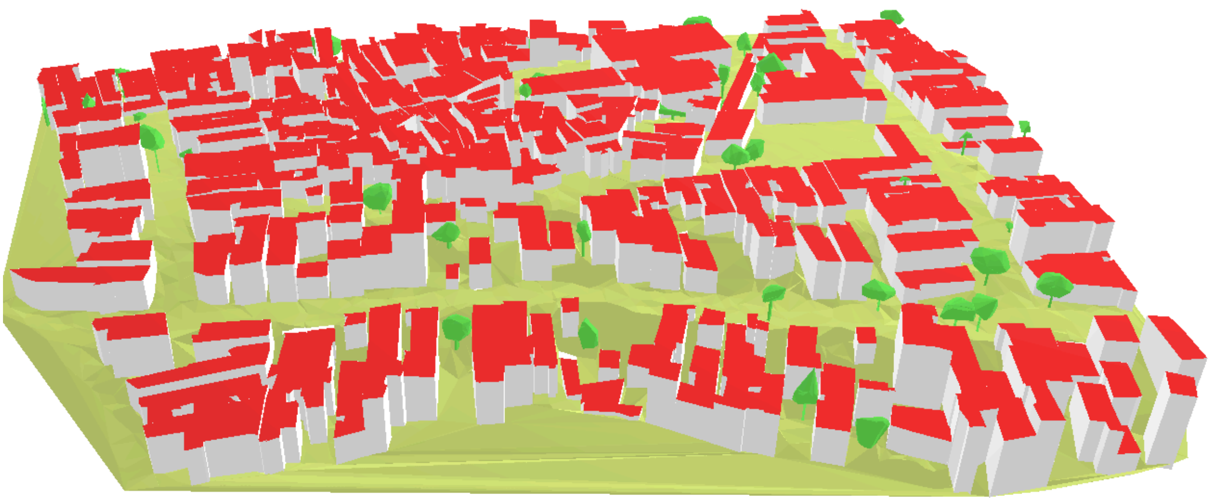


Figure 3-22 Scenario 1 - Brasilândia - CityGML model.

Figure 3-23 Figure 3 22 Scenario 2 - Santana - Input point cloud.

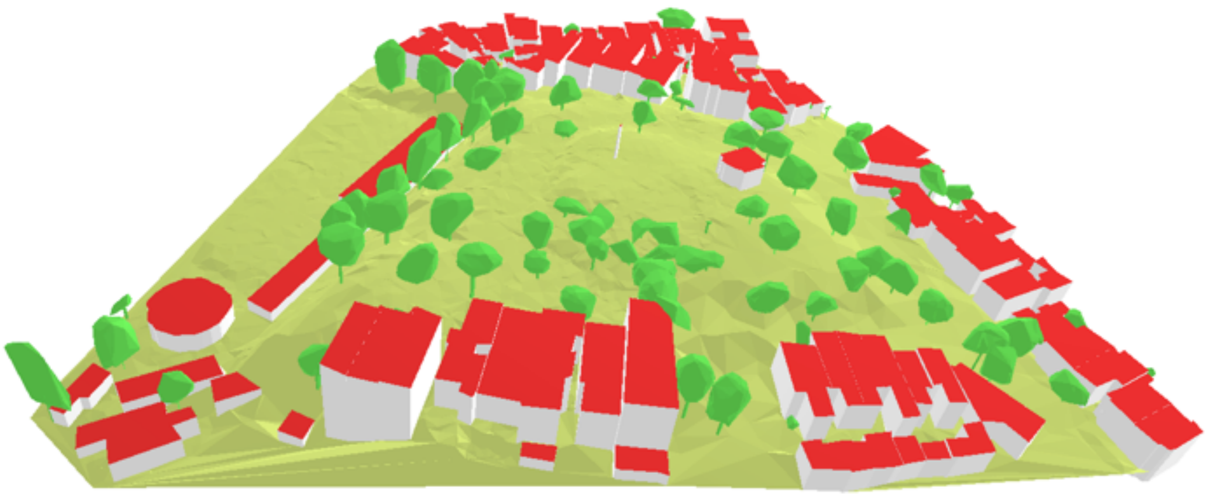


Figure 3-24 Figure 3 24 Scenario 2 - Santana - CityGML model.

In addition to all designing principles that already apply to both scenarios, another important aspect, considered in the case of Santana, is the digital reconstruction of the meteorological station. The tower hosting the metrological station is not a building, but rather an urban infrastructure, and therefore an alternative solution is required to incorporate this feature into the scenario. The workaround relies on creating a small artificial building footprint on the same position where the meteorological station lies, and then extrude this flat surface until the pseudo roof reaches the height of the meteorological station.

From the brief metadata available in the datasets of INMET, it is possible to retrieve the latitude and longitude of the automatic meteorological station, as well and its altitude with respect to sea level. No vertical datum is available for this height information. For this station, the declared height value of 785.64m is lower than the terrain value of 792.05m, which is retrieved by interpolating the relief surface created from the point cloud dataset – hereby considered as ground truth. Although this height difference might result from a datum discrepancy rather than an error per se, it is necessary to find a pragmatical solution for the issue, since the meteorological station lies significantly higher than ground, and therefore underestimating its height – in an arboreal environment – might compromise the accuracy of the model.

Therefore, a simple but efficient solution relies on using the point cloud as ground truth and extracting the maximum height value among all points that lie within the fenced perimeter where the meteorological station is. The images below correspond to the street view and the reconstructed 'artificial' building of Mirante de Santana. By a visual analysis of the street view image, it is possible to affirm that the assumption is coherent.

Figure 3-25 Street view image of the meteorological station of Santana.

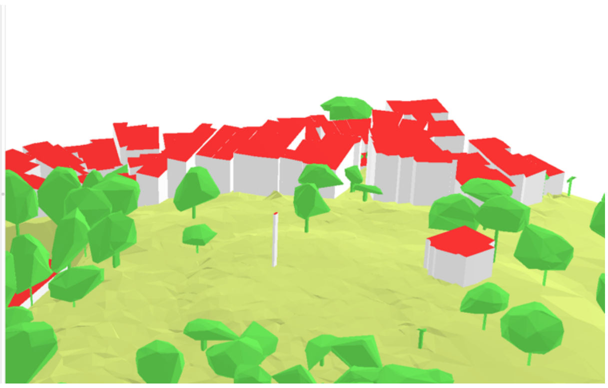


Figure 3-26 The reconstructed 'artificial' 3D building of Mirante de Santana.

The vertical elements in the image correspond to the automatic and conventional meteorological stations of Santana. Source: Google Street view.

The roof is set at 802.78m above sea level, and therefore 10.73m above ground level. By visual interpretation of the image on the left, this assumption seems plausible.

3.7 Raster Model

Among the six solar irradiation modules under analysis, the first two – namely the ones in GRASS GIS and ArcGIS – demand a raster-based data model as their input, which can be encoded in a GeoTiff extension. GeoTiff offers the advantage of being a geo-referenced data structure, i.e., the raster file will be correctly positioned in the GIS package based on a coordinate reference system.

Regarding the present research, a much more straightforward approach is established when designing the two scenarios in a raster-based model. The raster datasets are derived directly from the point cloud, with a spatial resolution of one meter. The rasterization process considers the maximum height value in the collection of points that lie within each resulting pixel of the grid.

However, a problem resulting from this decision is the noise generated by urban equipment that are present in the dataset, such as high-voltage power lines. These features may have an insignificant area or volume in 3D but, when present in a 2D raster, they act as barriers in the viewshed analysis. The following images present the issue.

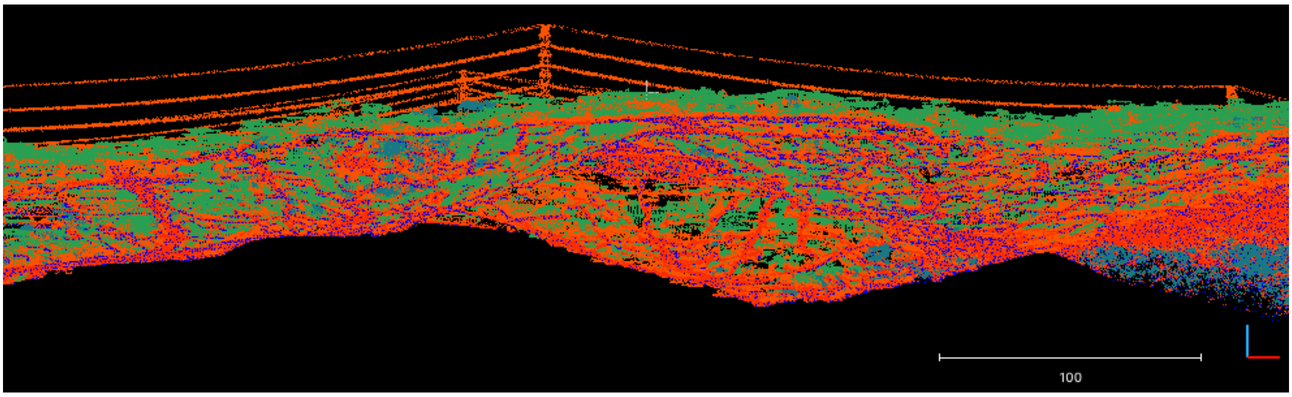


Figure 3-27 Example of a far field obstruction element in a point cloud dataset.

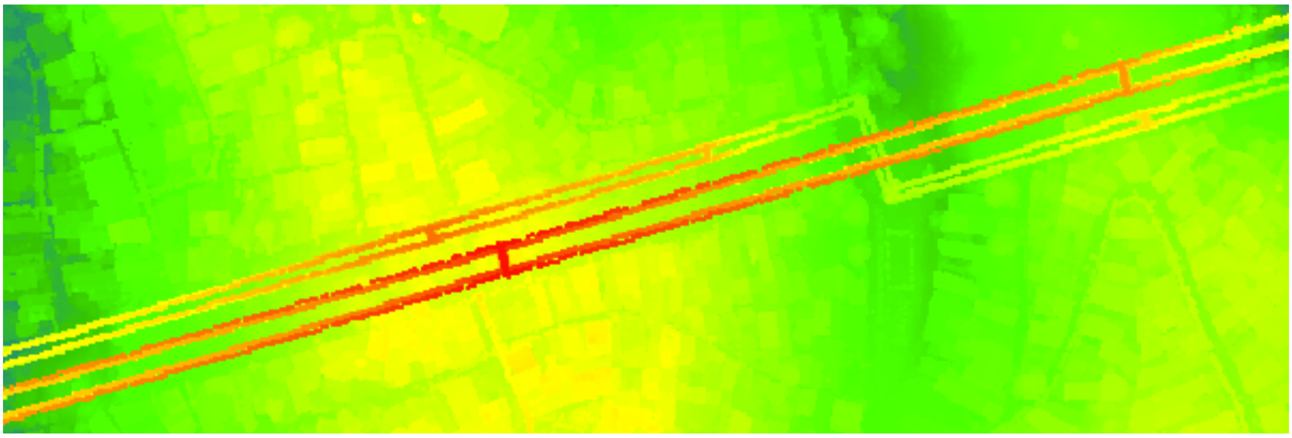


Figure 3-28 Example of a far field obstruction element when rasterised.

Therefore, the alternative is filtering the point cloud, thus keeping only points that represent relief, vegetation, and buildings. Filtering creates a few no-data values as some cells of the resulting raster remain empty. This does not pose a problem to the solar analysis, since most no-data values are relief cells and therefore lower than the building roofs.

Using [Cloud Compare](), a 3D point cloud (and triangular mesh) processing software, the input point cloud tiles are merged, and the resulting tile is filtered and rasterised. The final and most important step is modifying the value of the cell where the meteorological station lies, which is constrained to be equivalent as the height of the analogous roof in the CityGML model. The resulting datasets are hereby presented:

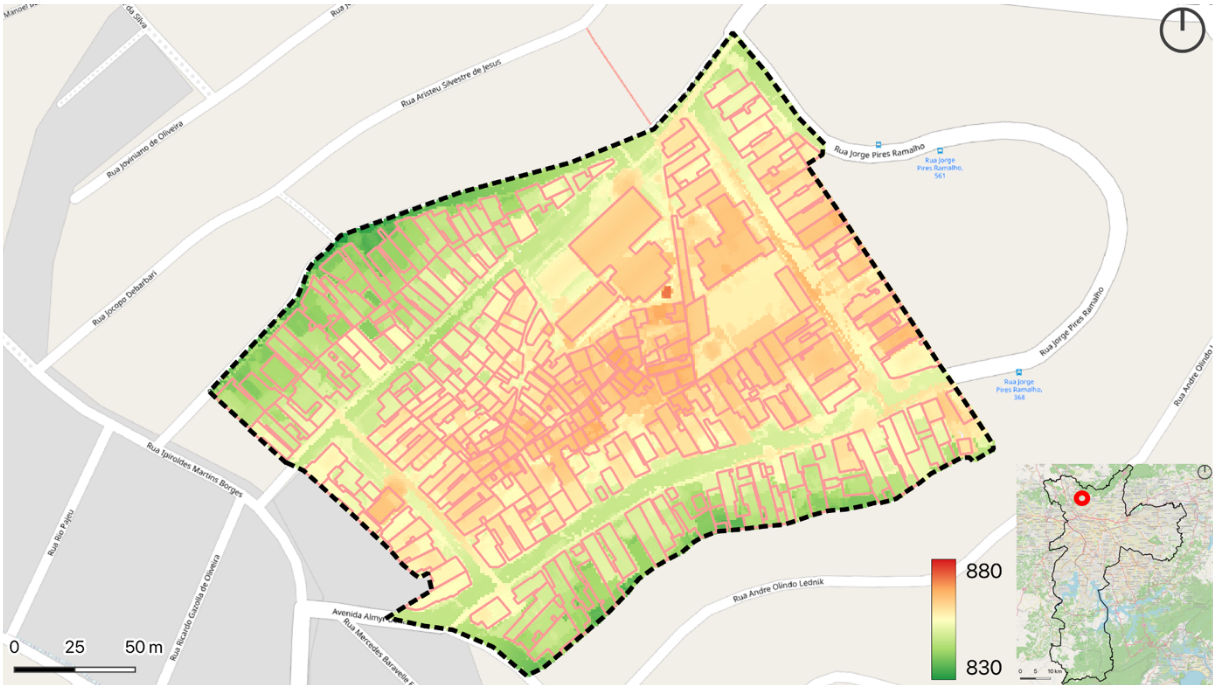


Figure 3-29 Scenario 1 - Brasilândia raster-based data model.

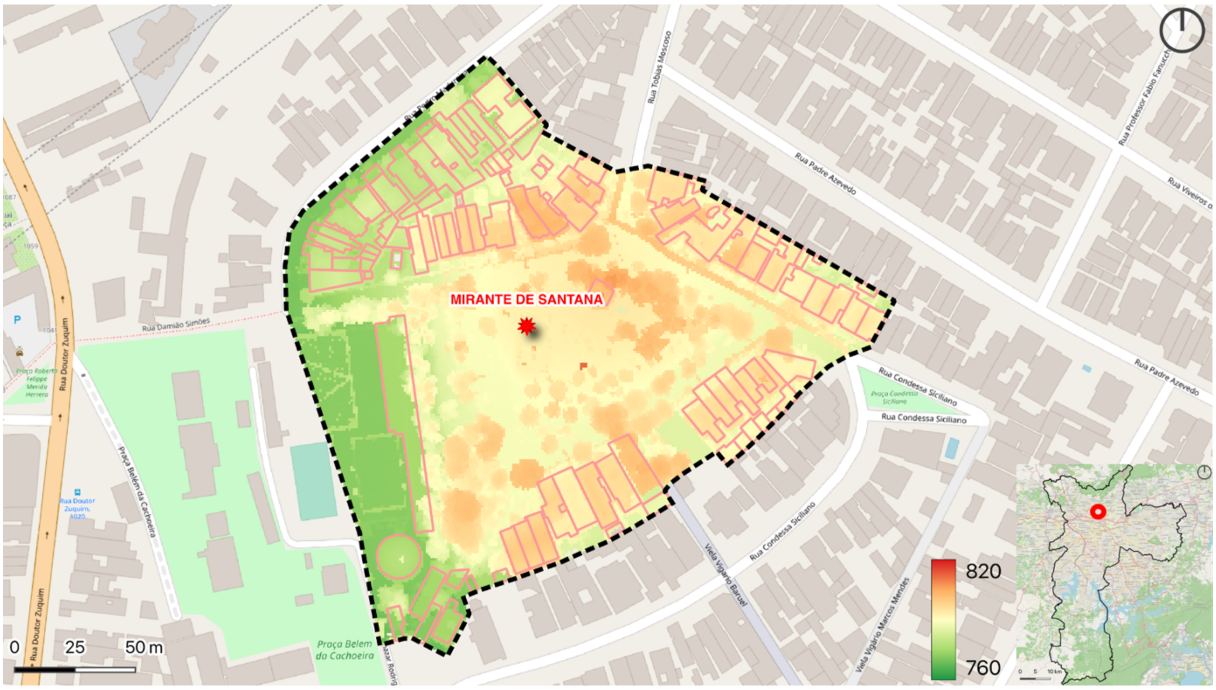


Figure 3-30 Scenario 2 - Santana raster-based data model.

The height value of the cell corresponding to the position of Mirante de Santana is changed to match the CityGML model.

3.8 Processing the Weather Dataset

Before running the solar modules, a last mandatory step is preparing the weather dataset for the comparison schema.

Both meteorological datasets – the automatic one and the conventional one – were downloaded with their records ranging from 24th of June 2006 – when the automatic station was launched – until the 12th of July 2021 – when these were queried from the governmental database. By making use of as many hourly records as possible in a long time series, the collection of solar irradiation values can be averaged into records that are representative of typical days of the year. This procedure mitigates the spiky characteristic of the curve and makes the ground truth dataset closer to reality.

Another important aspect to keep into consideration is the timestamp information present in the datasets from INMET, recorded as Coordinated Universal Time UTC. Since São Paulo lies in UTC - 3, the records are shifted accordingly. The following charts represent typical daily and monthly global solar irradiation values for the meteorological station of Mirante de Santana.

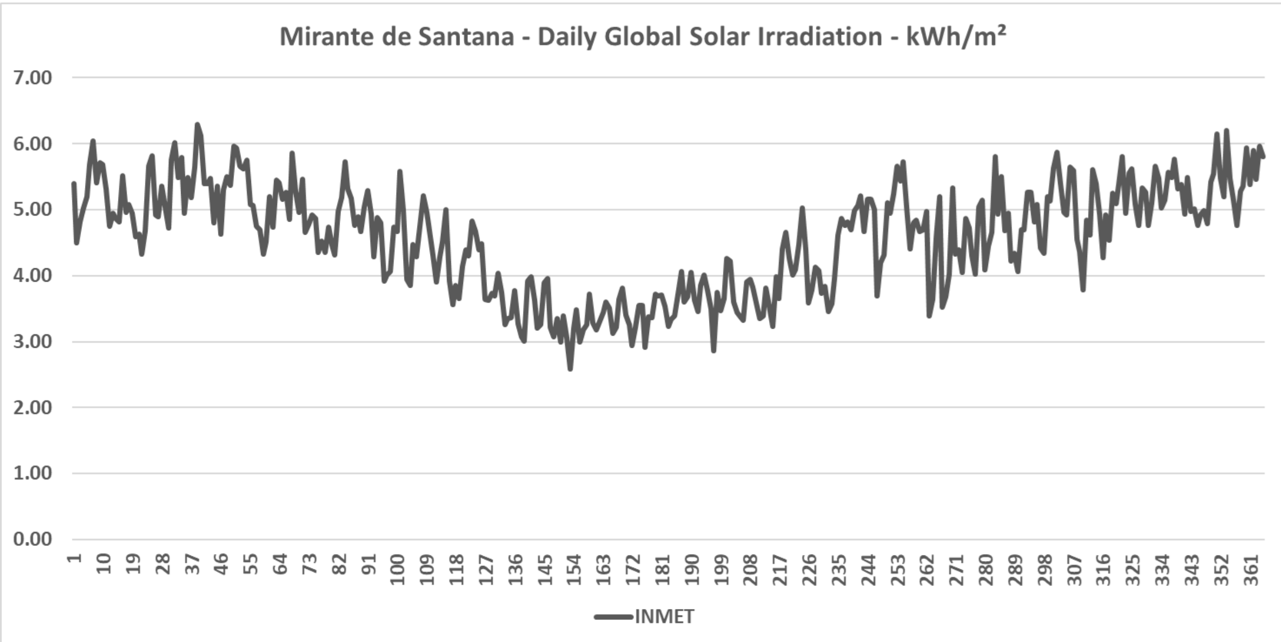


Figure 3-31 Mirante de Santana - Daily Global Solar Irradiation - kWh/m².

The city of São Paulo lies in the Southern Hemisphere, and therefore December is the month with highest records of solar irradiation, whereas June is the one with the lowest. Nevertheless, Winter – from June to September – is the season with clearest sky conditions, which balances the fact that the solar rays are obliquus at the latitude 23 degrees South during this period.

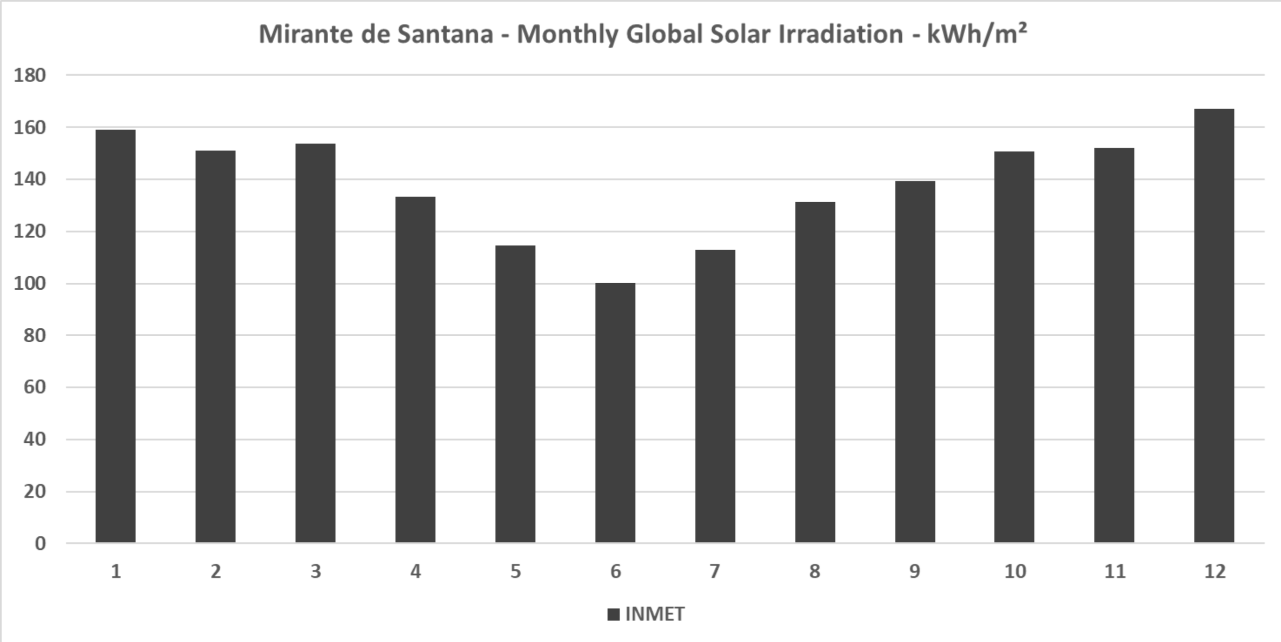


Figure 3-32 Mirante de Santana - Monthly Global Solar Irradiation - kWh/m².

As solar irradiation is a measure of energy per unit of area, the aggregation from hourly values into daily or monthly values correspond to the summation of all records that belong to the specific time interval. Averaging is only performed when confronting multiple hourly values that represent the same hour of the year (HOY).

3.9 Roof Mapping

Considering the recent efforts of the colleagues in the 3D geoinformation group, the present research analysis the results coming from their original 3D building reconstruction algorithm, which is then adapted to a hypothetical inexistence of building footprints in the city block of Brasilândia.

For such reconstruction task, the only input dataset is the lidar point cloud, which is clipped to the extents of the city block under analysis. According to the developers of the algorithm, this input point cloud is thematically filtered to preserve only points belonging to the building class. Successively, two main data pipelines are implemented, and their parameters are:

<u>Region growing</u>

- For each one of the remaining points, its five nearest neighbours are used to estimate the normal vector of the points.
- The dot product between the plane vector and the normal of the candidate point must be larger than 0.75.
- The distance between the plane and the candidate point must be smaller than 0.2m.

- The minimum number of points, so that a plane region is marked as valid, is 20.
- At every five new points that are added, the plane region is refitted to the inliers.

Alpha shape

- An alpha shape algorithm is executed to refine the resulting polygon, with an alpha value of 0.5m.

The image below represents the reconstructed building roofs, as output by the adapted 3D building reconstruction algorithm. The yellow lines correspond to the edges of the building roofs, whereas the white dots correspond to the input building points. It is possible to observe that, albeit some clusters of building points do not result in roof 3D surfaces, most of these points are contained by the roof planes.

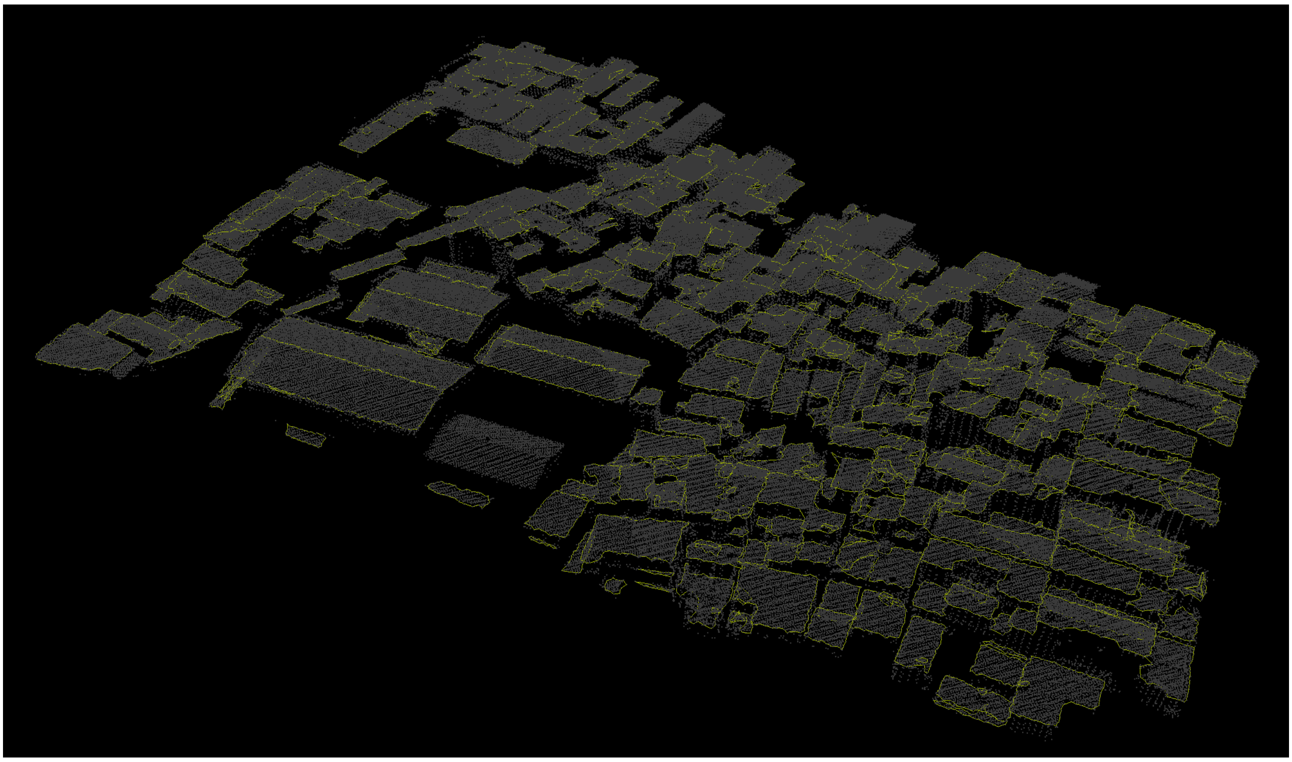


Figure 3-33 Reconstructed building roofs, as output by the adapted 3D building reconstruction algorithm.

Source: Ravi Peters.

In the last section of the following chapter, the reconstructed building roofs are confronted to the official 2D building footprints, and experiments are conducted to optimize their shapes.

IMPLEMENTATION & EXPERIMENTS

4 IMPLEMENTATION & EXPERIMENTS

In the present chapter, the experiments conducted in each one of the solar irradiation modules are systematically reported, with a focus on their data pipeline. The numeric results of each solar module in terms of irradiation values are presented in Chapter 5. The two raster-based modules are firstly presented, and sequentially the four vector-based ones. Finally, the experiments regarding roof mapping are reported.

All experiments are conducted on a MacBook Pro (13-inch, Mid 2012) equipped with a 2.5Ghz Dual-Core Intel Core i5 processor, 8Gb 1600 Mhz DDR3 of memory and a 1Tb SSD.

4.1 GRASS GIS

GRASS is an acronym for Geographic Resources Analysis Support System. This free-of-charge and open-license software is developed by a group of international scientists and developers from various fields[19], and it is a project member of the Open Source Geospatial Foundation (OsGeo). Running on Linux, Mac and Windows operating systems, GRASS offers an advantage when combining both a graphic user interface and a command line syntax, which allows a higher degree of flexibility for automation. For the present research, the GRASS GIS 7.9.dev (a7d1f4732) version is used on a macOS.

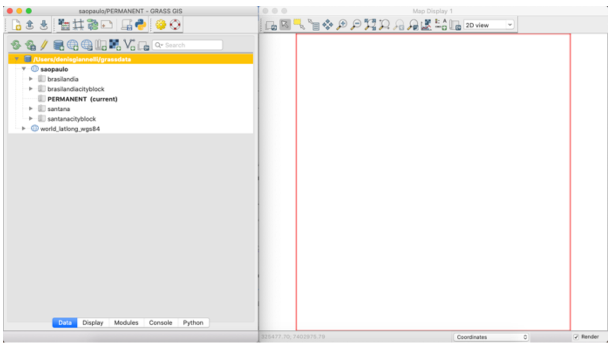


Figure 4-1 GRASS GIS – graphic user interface.

Figure 4-2 GRASS GIS – command line syntax

The software offers a portfolio of over five hundred algorithms, which are clustered according to the nature of the command: Display commands (d.*), Database commands (db.*),

[19] Overview of GRASS GIS, extracted from the official website.

General commands (g.*), Imagery commands (i.*), Miscellaneous commands (m.*), PostScript commands (ps.*), Raster commands (r.*), 3d raster commands (r3.*), Temporal commands (t.*), Test commands (test.*) and Vector commands (v.*)

In the case of solar irradiation analysis, the module is named **'r.sun'** (Solar irradiance and irradiation), which is therefore implemented with a raster-based algorithm. In the documentation of the module, a series of remarks from the authors[20] guide the data pipeline of the present research, and are hereby presented in topics to facilitate the reading:

> **"r.sun** computes **beam (direct), diffuse and ground reflected** solar irradiation raster maps for given day, latitude, surface and atmospheric conditions.
>
> (…)
>
> The shadowing effect of the **topography is incorporated by default**. This can be done either internally by calculation of the shadowing effect directly from the digital elevation model or by specifying raster maps of the horizon height which is much faster. These horizon raster maps can be calculated using **r.horizon**.
>
> (…)
>
> For latitude-longitude coordinates it requires that the **elevation map** is in meters. The rules are: lat/lon coordinates: elevation in meters; Other coordinates: elevation in the same unit as the easting-northing coordinates.
>
> (…)
>
> The model computes all three components of global radiation (beam, diffuse and reflected) **for the clear sky conditions**, i.e., not taking into consideration the spatial and temporal variation of clouds.
>
> (…)
>
> **r.sun** works in two modes: In the first mode it calculates for the set local time a solar incidence angle [degrees] and solar irradiance values [W.m-2**]. In the second mode daily sums of solar radiation [Wh.m-2.day-1]** are computed within a set day.
>
> (…)

[20] Jaroslav Hofierka, GeoModel, s.r.o. Bratislava, Slovakia. Marcel Suri, GeoModel, s.r.o. Bratislava, Slovakia. Thomas Huld, JRC, Italy.

> The solar incidence angle raster map *incidout* is computed specifying elevation raster **map *elevation***, aspect raster **map *aspect***, slope steepness raster **map *slope*,** given the day ***day*** and local time ***time***.
>
> (…)
>
> The specified day *day* is the number of the day of the **general year** where January 1 is day no.1 and December 31 is 365. Time ***time* must be a local (solar) time**
>
> (…)
>
> The program uses the Linke atmosphere turbidity factor and ground albedo coefficient. A **default, single value of Linke factor is *lin*=3.0** and is near the annual average for rural-city areas.
>
> (…)
>
> **The solar radiation maps for a given day are computed by integrating the relevant irradiance between sunrise and sunset times for that day**. The user can set a finer or coarser time step used for all-day radiation calculations with the *step* option. **The default value of *step* is 0.5 hour**.
>
> (…)
>
> The output units are in Wh per square meter per given day **[Wh/(m²)/day]**.
>
> (…)
>
> The **solar geometry** of the model is based on the works of Krcho (1990), later improved by Jenco (1992). The equations describing Sun -- Earth position as well as an interaction of the solar radiation with atmosphere were originally based on the formulas suggested by Kitler and Mikler (1986). This component was considerably updated by the results and suggestions of the working group co-ordinated by Scharmer and Greif (2000)
>
> (…)
>
> The **clear-sky solar radiation model** applied in the r.sun is based on the work undertaken for development of European Solar Radiation Atlas (Scharmer and Greif 2000, Page et al. 2001, Rigollier 2001). The clear sky model estimates the global radiation from the sum of its beam, diffuse and reflected components."

As the meteorological dataset of São Paulo only offers global solar irradiation values, i.e., beam (direct), diffuse and reflected irradiations combined, the sum of these three is the only metric that is extracted when running the simulations. Based on the documentation and considering the improvement of computing efficiency when providing *r.sun* with horizon maps, a data pipeline is adopted from the raster-based digital models:

Figure 4-3 GRASS GIS – data pipeline.

Within the GRASS GIS database, a new location named *saopaulo* is created with the appropriate coordinate referencing system of the input datasets, EPSG 31983. For each one of the scenarios – Brasilândia and Santana –, a new map set is created, thus dividing the data accordingly. The following pages describe the pipeline for Brasilândia, which is the same as the one that is conducted for Santana.

Firstly, the raster-based model is imported with *r.import* and the region of the current project is set with *g.region* to match the extents of this raster dataset. This second step is important in a GRASS GIS environment to run the algorithms according to the spatial resolution and extension of the raster dataset. The following pieces of command describe the procedures, resulting in the figure below:

```
(Fri Sep 10 07:52:35 2021)
r.import input=(…)/brasilandia_block_V8_DSM.tif output=brasilandia_block_V8_DSM
Importing raster map <brasilandia_block_V8_DSM>...
(Fri Sep 10 07:52:36 2021) Command finished (0 sec)
(Fri Sep 10 07:55:00 2021)
g.region raster=brasilandia_block_V8_DSM@brasilandiacityblock
(Fri Sep 10 07:55:00 2021) Command finished (0 sec)
```

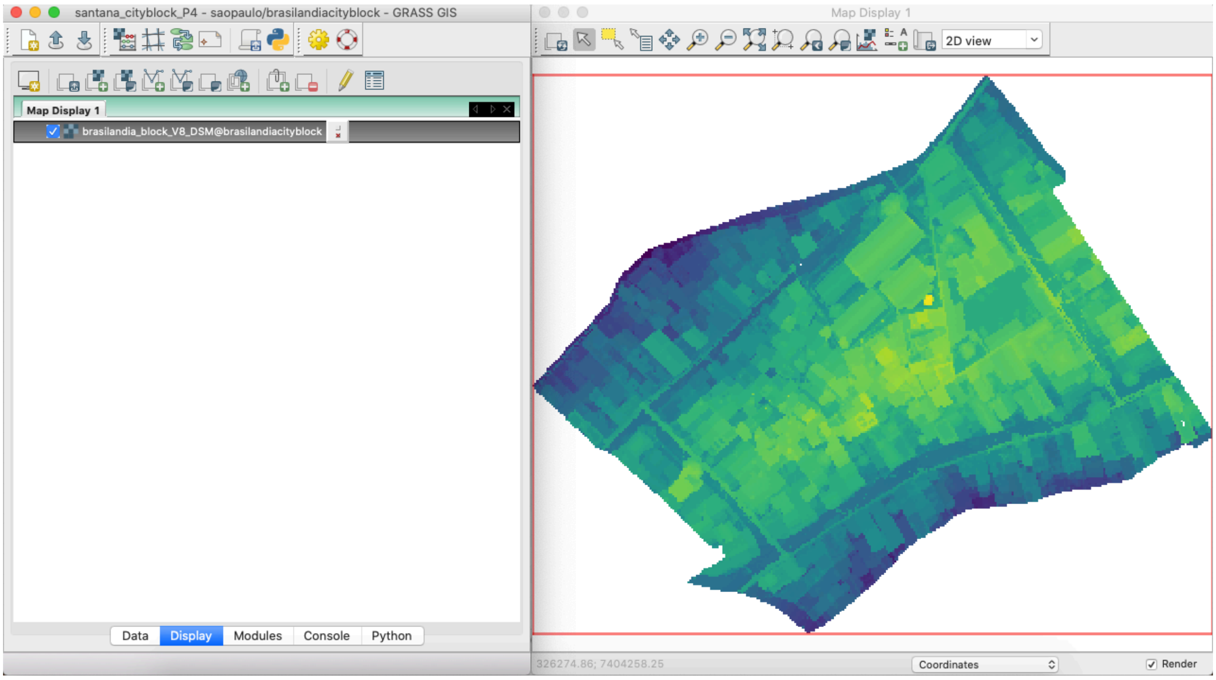


Figure 4-4 GRASS GIS – r.import and g.region.

Secondly, from the imported digital surface model, the r.slope.aspect module is run to extract the horizon and vertical angles (in degrees) of the surfaces in the model. This is a necessary step to incorporate the effect of the relief in the final solar irradiation simulation.

```
(Fri Sep 10 07:57:04 2021)
r.slope.aspect -e --overwrite elevation=brasilandia_block_V8_DSM@brasilandiacityblock slope=brasilandia_block_V8_slope aspect=brasilandia_block_V8_aspect
Aspect raster map <brasilandia_block_V8_aspect> complete
Slope raster map <brasilandia_block_V8_slope> complete
(Fri Sep 10 07:57:04 2021) Command finished (0 sec)
```

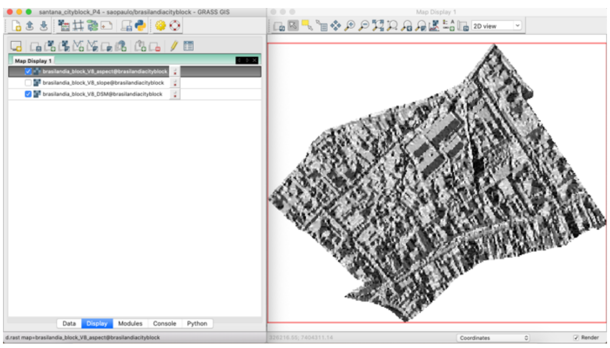


Figure 4-5 GRASS GIS – aspect map.

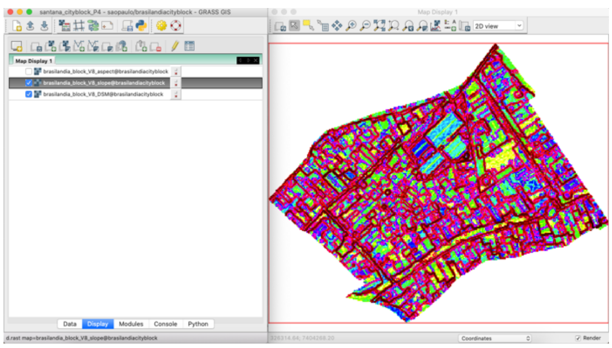


Figure 4-6 GRASS GIS – slope map.

Thirdly, the r.horizon module extracts sequential maps of inclination angles (in degrees), which in turn correspond to the maximum line of sight from a given position in the scenario. These horizon maps are computed at every one step degree of horizon angle, i.e., from 0 (east) to 360 (east). The following pictures correspond to the four cardinal directions, but such maps are computed multiple times in a loop iteration.

```
(Fri Sep 10 07:59:42 2021)
r.horizon -d --overwrite elevation=brasilandia_block_V8_DSM@brasilandiacityblock direction=0 step=1 output=brasilandia_block_V8_horizon
Calculating map 1 of 360 (angle 0.00, raster map <brasilandia_block_V8_horizon_000>)
 (…)
Calculating map 360 of 360 (angle 359.00, raster map <brasilandia_block_V8_horizon_359>)
(Fri Sep 10 08:01:04 2021) Command finished (1 min 21 sec)
```

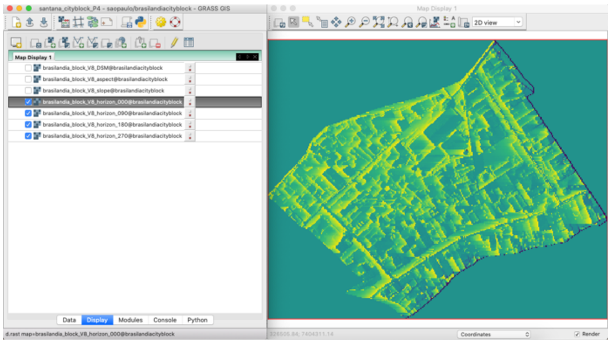


Figure 4-7 GRASS GIS – horizon map angle 000 (east).

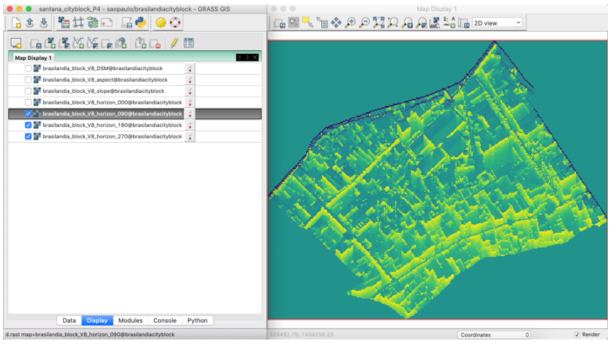


Figure 4-8 GRASS GIS – horizon map angle 090 (north).

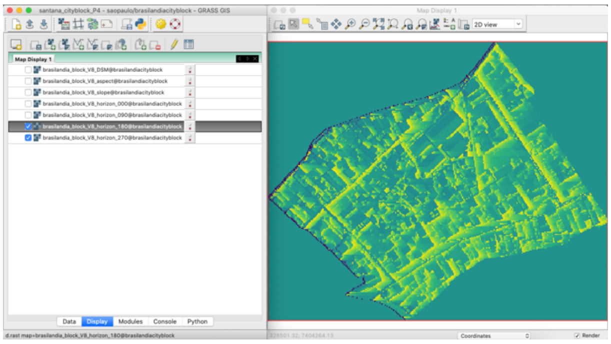


Figure 4-9 GRASS GIS – horizon map angle 180 (west).

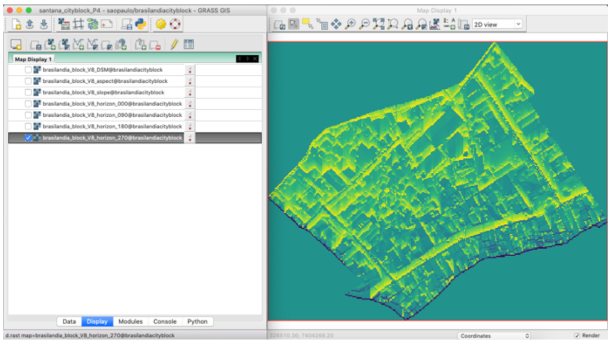


Figure 4-10 GRASS GIS – horizon map angle 270 (south).

Sequentially, the r.sun module is run according to the instructions in the documentation. In a python code editor, a for loop statement iterates over all days of the year [1,365] and runs the second module of r.sun, thus extracting daily summation of solar irradiation values from the DSM, aspect, slope, and horizon maps.

```python
def compute_r_sun(i):
    day_value = str(i).zfill(3)
    glob_rad_value = 'brasilandia_block_V8_glob_irrad_' + day_value
    gscript.run_command('r.sun',
                        overwrite=True,
                        elevation='brasilandia_block_V8_DSM',
                        aspect='brasilandia_block_V8_aspect',
                        slope='brasilandia_block_V8_slope',
                        horizon_basename='brasilandia_block_V8_horizon',
                        horizon_step=3,
                        glob_rad=glob_rad_value,
                        day=i,
                        step=0.5)
if __name__ == '__main__':
    for i in range(1,366):
       compute_r_sun(i)
```

Some details are worth of explanation:

- In the absence of Linke coefficients, r.sun adopts the generic value of 3.0. This parameter relates to the atmospheric turbidity and varies according to the environment under analysis – rural, urban, industrial areas – and to the month of the year. In the documentation of GRASS GIS, Linke coefficients are given for a mild climate in the northern hemisphere, which is not the case of São Paulo. Since the author did not encounter any Linke maps for São Paulo, it was decided to set it as default, which is, according to the developers, “near the annual average for rural-city areas”. Further experiments that adopt higher Linke values can be carried out to understand the impact in the overall irradiation results.
- Analogously, in the absence of ground Albedo coefficients, GRASS GIS defines the constant of 0.2. This parameter relates to the proportion of incident radiation that is reflected by the ground surface, and therefore varies according to the multiple materials that are present in the scenario, such as asphalt, concrete, bare ground, grass, etc. In the case of São Paulo, a land cover map would be required to define specific albedo parameters at every pixel of the study area, according to the material. Nevertheless, this dataset is not available in a spatial resolution that is compatible to the analysis that is carried out. Therefore, the author decided to adopt the standard value of 0.2
- Due to a limitation of GRASS GIS, not all 360 horizon maps can be open at the same time. Therefore, by testing how many horizon maps GRASS GIS could open simultaneously, it was concluded that only one third of these – one at every three-

degree step – could be used in the solar analysis. Nevertheless, this limitation only impacts the extension of the scenario itself since the analysis conducted in GRASS GIS – as it is the case of all other simulations for this research – only considers the relief within the scenario. In practical terms, this means that all features within the scenario will be incorporated in the shading effect of the r.sun module, regardless of whether they are considered in the horizon maps or not.

```
(Fri Sep 10 08:06:36 2021)
/Users/denisgiannelli/grassdata/saopaulo/santanacityblock/.tmp/Deniss-MacBook-Pro.local/74503.0.py
Number of threads <1>
Mode 2: integrated daily irradiation for a given day of the year
Using Linke constant: 3.000000
Using albedo constant: 0.200000
Using slope map <brasilandia_block_V8_slope>
Using aspect map <brasilandia_block_V8_aspect>
(…)
Number of threads <1>
Mode 2: integrated daily irradiation for a given day of the year
Using Linke constant: 3.000000
Using albedo constant: 0.200000
Using slope map <brasilandia_block_V8_slope>
Using aspect map <brasilandia_block_V8_aspect>
(Fri Sep 10 08:14:02 2021) Command finished (7 min 26 sec)
```

Exclusively for visualisation purposes, when all daily irradiation maps are computed for both scenarios, another loop statement computes the summation of global irradiation values for each month of the year, based on the appropriate daily interval of the corresponding month. The maximum and minimum values among all these maps are extracted, and all maps are coloured with the same hue ramp. The following images illustrate the datasets for the months of June (winter) and December (summer).

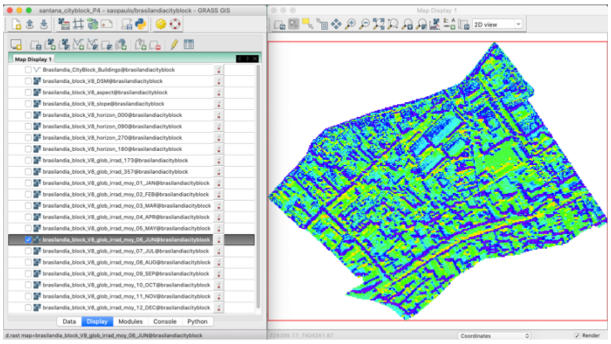


Figure 4-11 GRASS GIS – solar irradiation values – JUN.

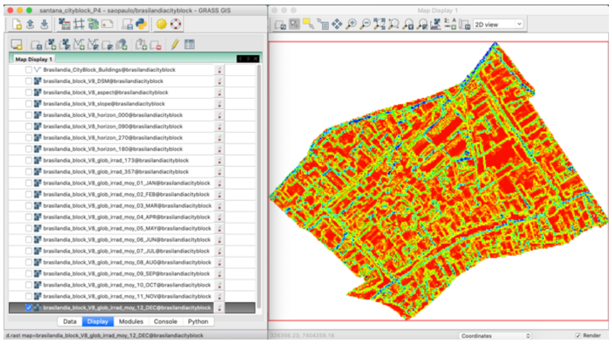


Figure 4-12 GRASS GIS – solar irradiation values – DEC.

Lastly, the computed daily solar irradiation values are extracted from GRASS GIS:

In the case of Santana, for each computed daily map, the only value of interest is the one associated to the raster cell where the meteorological station lies. Therefore, the *r.what* module is used to extract multiple irradiation values across all daily maps, based on the same coordinates of the meteorological station. By hands-on experience, it can be affirmed that this module results extremely sensitive to the input coordinates. The most verisimilar and consistent set of values is derived from the coordinates of the centre of the pixel, and not the coordinate of the meteorological station, although these differ by a couple of centimetres. The *r.what* command is also limited to a maximum number of maps, and therefore these are divided into four groups of days throughout the year.

```
def extract_info_mirante():
    temp_maps = ''
    # (1, 91) (91,181) (181,271) (271,366)
    for i in range(1,91):
        temp_maps += 'santana_block_V8_glob_irrad_'+str(i).zfill(3)+','
    daily_maps = temp_maps[:len(temp_maps)-1]
    mirante  =  'mirante_santana'
    gscript.run_command('r.what',
                        flags = 'n',
                        map = daily_maps,
                        coordinates = '334568.4892302206,7400606.521734362’)
if __name__ == '__main__':
    extract_info_mirante()
```

In the case of Brasilândia, for each computed daily map, an averaging is performed to extract a representative solar irradiation value for each roof in the scenario. Averaging is requested since irradiation is a measure of energy per square meter, and each roof contains multiple pixels of one square meter each. Therefore, the *v.rast.stats* module is used to compute the statistics of each building roof across all daily maps, based on the original building footprints.

```
def extract_info_brasilandia():
    for i in range(1,366):
        image = 'brasilandia_block_V8_glob_irrad_' +str(i).zfill(3)
        day = 'd' + str(i).zfill(3)
        gscript.run_command('v.rast.stats',
                                                   flags = 'c',
                                                   map =
'Brasilandia_CityBlock_Buildings',
                                                   raster= image,
                                                   column_prefix= day,
                                                   method='average',
                                                   percentile=100)
if __name__ == '__main__':
    extract_info_brasilandia()
```

It is important to notice that, until this point, the solar irradiation analysis in GRASS GIS does not make use of the cadastral dataset, which is a positive quality considering the mapping of solar irradiation in favelas. Nevertheless, a single value of irradiation per building can only be extracted if geometrical data of such buildings are available. Visual interpretation of the values is possible, but time consuming and subjective to the human sight.

The extracted daily irradiation values are converted from Wh/m^2 to kWh/m^2 and aggregated into monthly and yearly values. With the results available, it is possible to respond the sub-research questions regarding GRASS GIS:

- **What are the minimum data requirements to run the solar module?**
  Strictly speaking, only the raster-based digital surface model of the scenario is required. In combination with slope and aspect maps, the module considers the relief. If the horizon maps are computed beforehand, the process runs much faster. No weather dataset is mandatory, as the module computes irradiance values internally based on the position of the scenario, and clear sky conditions are assumed.

- **How automatized is the solar irradiation pipeline? From a user perspective, how complex is setting up the simulation?**

Considering the optimal dataflow, multiple intermediate steps are necessary, namely: computing of aspect, slope, and horizon maps. Nevertheless, these can be scripted and therefore effortlessly automatized.

- **Is the solar irradiation module self-contained, or is it part of a greater package of solutions?**
  The r.sun module is self-contained, delivering only irradiation results.

- **What is the computational running time of the module if this metric is retrievable?**
  For the scenario of Brasilândia, computing all intermediate and final steps results in a running time of HH:MM:SS = 00:19:01.

- **What kind of urban features – among buildings, vegetation, and relief – does the module support, and how do these features participate in the simulation?**
  All urban features that are previously incorporated in the DSM will participate in the simulation as raster cells, which store horizontal global irradiation values.

- **What is the degree of flexibility offered by the solar module in terms of selecting which features of the city model will be assessed and which will take part only as shading elements?**
  There is no flexibility, all pixels that are assigned with a DSM value will be assessed, either in the first mode, which calculates solar incidence angle and solar irradiance, or in the second one, which calculates daily sums of solar irradiation for a specified day. There is not direct option of assessing only selected portions of the DSM. A mask layer could have been adopted to verify if a selection of pixels is possible, but this is out of the scope of this research.

- **How is the output data delivered by the module in terms of time granularity and data model?**
  The module delivers daily values of global solar irradiation in a raster model. If the first mode of r.sun is implemented, hourly values are also possible.

4.2 ArcGIS

ArcGIS is a commercial suite of services developed by the US company ESRI – an acronym for Environmental Systems Research Institute[21]. Originally created as ARC/INFO, this command line solution for GIS analysis was later merged with its guided user interface and adopted the name ArcDesktop. In the present time, ArcDesktop and ArcGIS Pro are offered by the company, with the latter being the successor of the former. ArcGIS is a windows-based suite of services, which require a disk partition or a virtual machine in case of other operating systems. For the present research, ArcDesktop 10.8 was adopted, considering the academic license available for students and staff of Delft University of Technology. For methodological consistency of the research, ArcGIS was run on a bootcamp partition of the same computer that ran other software.

Within ArcMap – the guided user interface of ArcGIS –, the user may access all ArcToolboxes that are available under the commercial license that they signed for, one of these being the spatial analyst. This toolbox is an extension that operates with raster and vector datasets and hosts a collection of toolsets grouped by functional categories.

In the case of solar irradiation analysis, the toolset is named 'Area Solar Radiation', which is implemented with a raster-based algorithm. In the documentation of the module, a series of remarks from ESRI guide the choice of parameters for the simulation, and the most important ones are hereby presented in topics to facilitate the reading:

[21] ESRI's official website

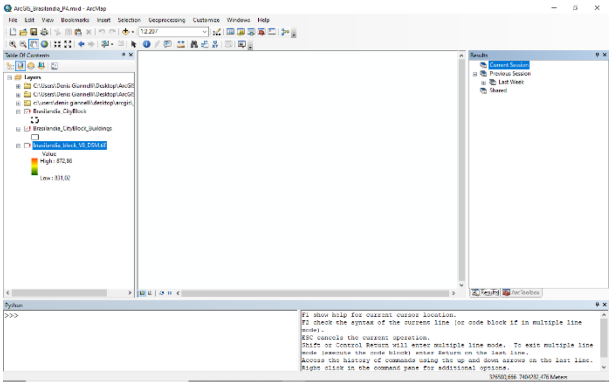


Figure 4-13 ArcGIS – Guided user interface overview.

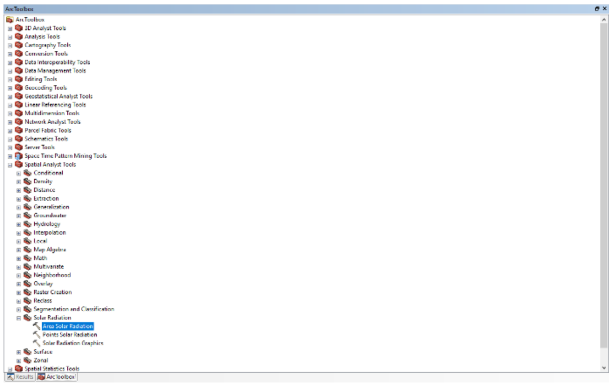


Figure 4-14 ArcGIS – ArcToolboxes and Spatial Analyst Tools.

“The **output** radiation rasters will always be floating-point type and have units of **watt hours per square meter (Wh/m²)**. The direct duration raster output will be integer with unit hours.

(…)

The **latitude for the site area** (units: decimal degree, positive for the northern hemisphere and negative for the southern hemisphere) is used in calculations such as solar declination and solar position (…) Because the solar analysis is designed for landscape scales and local scales, it is acceptable to use one latitude value for the whole DEM. (…) For input surface rasters containing a spatial reference**, the mean latitude is automatically calculated**.

(…)

Sky size is the resolution of the viewshed, sky map, and sun map rasters that are used in the radiation calculations (units: cells per side). These are upward-looking, hemispherical raster representations of the sky and do not have a geographic coordinate system. These rasters are square (equal number of rows and columns). Increasing the sky size increases calculation accuracy but also increases calculation time considerably. (…) **A value of 200 is default and is sufficient for whole DEMs with large day intervals** (for example, > 14 days).

(…)

For **multiday time configurations**, the maximum range of days is a total of **one year** (365 days, or 366 days for leap years). If the start day is greater than the end day, the time calculations will proceed into the following year. (…) **The start day and end day cannot be equal.**

(…)

The year value for time configuration is used to determine a leap year. It does not have any other influence on the solar radiation analysis as the calculations are a function of the time period determined by Julian days.

(…)

The **Create outputs for each interval** check box provides the flexibility to calculate insolation integrated over a specified time period or insolation for each interval in a time series. For example, for the within-day time period with an hour interval of one, checking this box will create hourly insolation values; otherwise, insolation integrated for the entire day is calculated.

(…)

The use of a **z-factor** is essential for correcting calculations when the surface z units are expressed in units different from the ground x,y units. To get accurate results, the z units should be the same as the x,y ground units. If the units are not the same, use a z-factor to convert z units to x,y units.

(…)

How slope and aspect information are derived for analysis: **FROM_DEM** — The slope and aspect rasters are calculated from the input surface raster. This is the default. FLAT_SURFACE — Constant values of zero are used for slope and aspect.

(…)

Because the viewshed calculation can be highly intensive, **horizon angles** are only traced for the number of calculation directions specified. Valid values must be multiples of 8 (8, 16, 24, 32, and so on). Typically, a value of 8 or 16 is adequate for areas with gentle topography, whereas a value of 32 is adequate for complex topography. The default value is 32

(…)

zenith_divisions (Optional): The number of divisions used to create sky sectors in the sky map. The default is eight divisions (relative to zenith). Values must be greater than zero and less than half the sky size value.

(…)

azimuth_divisions (Optional): The number of divisions used to create sky sectors in the sky map. The default is eight divisions (relative to north). Valid values must be multiples of 8. Values must be greater than zero and less than 160.

(…)

Type of **diffuse radiation model: UNIFORM_SKY** — Uniform diffuse model. The incoming diffuse radiation is the same from all sky directions. This is the default.

(…)

The diffuse proportion is the fraction of global normal radiation flux that is diffuse. Values range from 0 to 1. This value should be set according to atmospheric conditions. Typical values are 0.2 for very clear sky conditions and 0.3 for generally clear sky conditions.

(…)

Transmittivity is a property of the atmosphere that is expressed as the ratio of the energy (averaged overall wavelengths) reaching the earth's surface to that which is received at the upper limit of the atmosphere (extraterrestrial). Values range from 0 (no transmission) to 1 (complete transmission). Typically observed values are 0.6 or 0.7 for very clear sky conditions and 0.5 for only a generally clear sky."

Again, as the meteorological dataset of São Paulo only offers global solar irradiation values, i.e., beam (direct), diffuse and reflected irradiations combined, the sum of these three is the only metric that is extracted when running the simulation.

It is important to observe that, unlike GRASS GIS, the solar module of ArcGIS does not allow the user to specify aspect and slope values from maps that are already computed. Therefore, the user either chooses to compute slope and aspect values each time that the Area Solar Radiation module runs, or they adopt a zero-value resolution. The same applies for the horizon maps, as the only option for the user is to provide the module with an angular step for horizon calculations, multiple of eight.

In this sense, the data pipeline in ArcGIS is more straightforward than the one in GRASS GIS, considering that the simulation runs directly with the Area Solar Radiation module. In a python code editor, a for loop statement iterates over all days of the year, except the last [1,364], and this remaining one is run separately. The reason behind this issue relies on the fact that the start and ending day of the simulation cannot be the same, i.e., the 1st of January.

```
#AreaSolarRadiation(in_surface_raster, {latitude}, {sky_size},
{time_configuration}, {day_interval}, {hour_interval},
{each_interval}, {z_factor}, {slope_aspect_input_type},
{calculation_directions}, {zenith_divisions}, {azimuth_divisions},
{diffuse_model_type}, {diffuse_proportion}, {transmittivity},
{out_direct_radiation_raster}, {out_diffuse_radiation_raster},
{out_direct_duration_raster})

for i in range(1,365):
    dsm     = "(…)/brasilandia_block_V8_DSM.tif"
    fout    = "(…)/d"+str(i).zfill(3)
    multi   = "MultiDays   2021    "+str(i)+"    "+str(i+1)
    arcpy.gp.AreaSolarRadiation_sa(dsm, fout, "-23,4973947988913", "200", multi,
"1", "0,5", "INTERVAL", "1", "FROM_DEM", "32", "8", "8", "UNIFORM_SKY", "0,3",
"0,5", "", "", "")
```

```
dsm      = "(…)/brasilandia_block_V8_DSM.tif"
fout_365     = "(..)/d"+str(365).zfill(3)
multi_365    = "MultiDays   2021    "+str(365)+"    "+str(1)
arcpy.gp.AreaSolarRadiation_sa(dsm, fout_365, "-23,4973947988913", "200",
multi_365, "1", "0,5", "INTERVAL", "1", "FROM_DEM", "32", "8", "8",
"UNIFORM_SKY", "0,3", "0,5", "", "", "")
```

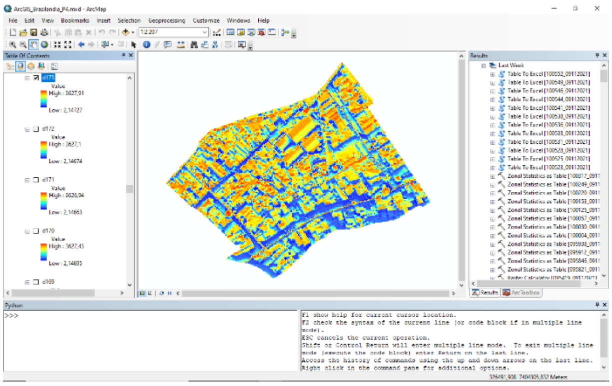


Figure 4-15 ArcGIS – solar irradiation values – June 22nd.

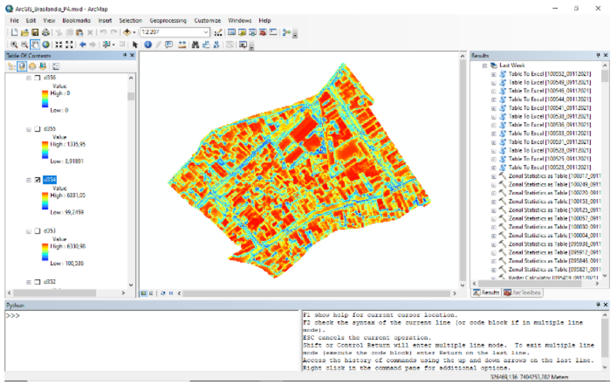


Figure 4-16 ArcGIS – solar irradiation values – Dec.20th..

Analogously to the experiments conducted on GRASS GIS, the final step is extracting the irradiation values for Brasilândia and Santana, each one with different implementations.

In the case of Santana, the irradiation values are retrieved by the module ExtractMultiValuestoPoints. A loop statement iterates over all irradiation maps and queries the values corresponding to the pixel overlapping the meteorological station of Mirante de Santana, which is input as a shapefile.

```
mirante_santana = "(…)/mirante_santana.shp"
file = ''
for i in range(1,366):
    file += "'d" + str(i).zfill(3)+"' d"+str(i).zfill(3)+";"
file = file[:len(file)-1]
arcpy.gp.ExtractMultiValuesToPoints_sa(mirante_santana, file, "NONE")
```

Regarding Brasilândia, the daily values are summated into monthly values of solar irradiation with the RasterCalculator tool. These monthly raster values are in turn averaged according to the spatial distribution of the building roofs, which are imported as polygon shapefiles. The resulting values are exported in xls files for further processing.

```
month_dict = {}
month_dict[1] = range(1,32)
month_dict[2] = range(32,60)
```

```python
month_dict[3] = range(60,91)
month_dict[4] = range(91,121)
month_dict[5] = range(121,152)
month_dict[6] = range(152,182)
month_dict[7] = range(182,213)
month_dict[8] = range(213,244)
month_dict[9] = range(244,274)
month_dict[10] = range(274,305)
month_dict[11] = range(305,335)
month_dict[12] = range(335,366)

for i in range(1,13):
    maps = ''
    for j in month_dict[i]:
        maps += '"d' + str(j).zfill(3) + '"' + ' + '
    maps = maps[:len(maps)-3]
    fout = "(…)/m"+str(i).zfill(2)
    arcpy.gp.RasterCalculator_sa(maps, fout)
```

```python
brasilandia_buildings = "(…)/Brasilandia_CityBlock_Buildings.shp"

for i in range(1,13):
    raster = "(...)/m"+str(i).zfill(2)
    fout = "(...))/m"+str(i).zfill(2)+"_average"
    arcpy.gp.ZonalStatisticsAsTable_sa(brasilandia_buildings, "FID", raster,
fout, "DATA", "MEAN")
```

```python
for i in range(1,13):
    fin = "(…)/m"+str(i).zfill(2)+"_average"
    fout = "(…)/m"+str(i).zfill(2)+"_average.xls"
    arcpy.TableToExcel_conversion(Input_Table=fin, Output_Excel_File=fout,
Use_field_alias_as_column_header="NAME",
Use_domain_and_subtype_description="CODE")
```

The extracted daily irradiation values are converted from Wh/m^2 to kWh/m^2 and aggregated into monthly and yearly values. With the results available, it is possible to respond the sub-research questions regarding ArcGIS:

- **What are the minimum data requirements to run the solar irradiation module?** Only the digital surface model (DSM) is requested to run the Area Solar Radiation module of ArcGIS.

- **How automatized is the solar irradiation pipeline? From a user perspective, how complex is setting up the simulation?**

Only a single step is necessary, and multiple days of the year can be computed in a loop statement. These may also be computed in a single run of the module but, due to memory issues, it was preferred to treat them independently.

- **Is the solar irradiation module self-contained, or is it part of a greater package of solutions?**
  The 'Area Solar Radiation' module is self-contained and exports only global irradiation values, optionally beam (direct), diffuse, and reflected ones as well.

- **What is the computational running time of the module if this metric is retrievable?**
  For the scenario of Brasilândia, computing all iterations in the for loop statement results in a running time of HH:MM:SS = 04:43:41.

- **What kind of urban features – among buildings, vegetation, and relief – does the module support, and how do these features participate in the simulation?**
  By using the Area Solar Radiation tool, all urban features that are previously incorporated in the DSM will participate in the simulation as raster cells, which store horizontal global irradiation values.

- **What is the degree of flexibility offered by the solar module in terms of selecting which features of the city model will be assessed and which will take part only as shading elements?**
  There is no flexibility, all pixels that are assigned with a DSM data value will be assessed. If the Point Solar Radiation tool is used instead, individual locations (pixels) of the DSM will be assessed. Nevertheless, this is done in a point-based approach, and therefore multiple portions of the DSM would be analysed with a high number of points.

- **How is the output data delivered by the module in terms of time granularity and data model?**
  The module delivers daily values of global solar irradiation in a raster model. If the 'within-day' time period is chosen, a nested loop could retrieve hourly values.

4.3 CitySim

CitySim Solver is a command-line integrated solver for simulating the energy demand and supply of buildings for space conditioning[22], developed at the Solar Energy and Building Physics Laboratory of the Swiss Federal Institute of Technology Lausanne. According to the developers, "the software CitySim is aiming to provide a decision support for urban energy planners and stakeholders to minimize the net use of non-renewable energy sources as well as the associated emissions of greenhouse gases". (EPFL)

Built on top of CitySim Solver, CitySim PRO is a Graphical User Interface developed by the company KAEMCO, "aiming at the simulation and optimisation of the sustainability of urban settlements"[23]. This Guided User Interface is free-of-charge for academic purposes and runs exclusively on a Windows Operating System. Therefore, analogously to the use of ArcGIS, the author uses CitySim on a bootcamp partition of the same computer for the methodological consistency. The version that is used in the present research is the 64-bit one that was built on the 6th of August 2021.

Unlike the two previous software, which offer several independent modules (GRASS GIS) or tools (ArcGIS) in the same GUI, CitySim PRO presents a unique running process for several applications that are handled by CitySim Solver. Their resulting attributes are:

- Short-wave irradiation (Wh/m^2) at surface resolution.
- Long-wave net irradiation (Wh/m^2) at surface resolution.
- Surface temperature (°C) at surface resolution.
- Photovoltaic production (Wh) at building resolution.
- Solar thermal production (Wh) at building resolution.
- Sky view factor [0,1] at surface resolution.
- Heating demand (Wh/m^3) at building (thermal zone) resolution.
- Cooling demand (Wh/m^3) at building (thermal zone) resolution.
- Indoor temperature (°C) at building (thermal zone) resolution.

[22] City Sim Software website.

[23] KAEMCO website.

In the case of the present research, the focus relies exclusively on short-wave irradiation, as this variable relates to the solar energy received by buildings and other features in the scenario.

CitySim Pro is also advantageous for allowing different vector-based data structures as the possible geometrical input for the simulations, such as DXF files from the designing domain or xml-encoded CityGML 2.0 files from the geo-information domain. For this research, the CityGML files of Santana and Brasilândia are imported and then converted by the software into the native data model of CitySim, which is also an xml-encoded model.

The following figure introduces the data pipeline for the solar irradiation analysis with CitySim. The experiments conducted with Santana and Brasilândia have identical processes.

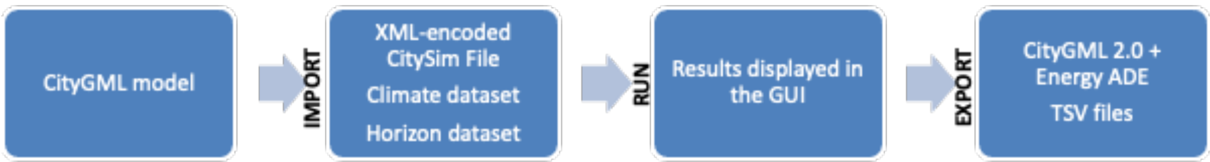


Figure 4-17 CitySim – data pipeline.

Firstly, by clicking on importing CityGML 2.0, the scenario is imported directly into the GUI of CitySim. In the present version of the software, buildings encoded in LoD2/LoD3 (buildings), solitary vegetation objects (trees) and relief components (relief) of the scenario are correctly imported and converted into the xml encoded CitySim file. This is possible due to the close contact that the research has had with the developer of CitySim throughout the timeline of their MSc thesis. The developer received feedback from the author and their supervisors and improved the importing mechanism to offer an efficient data workflow from CityGML to CitySim.

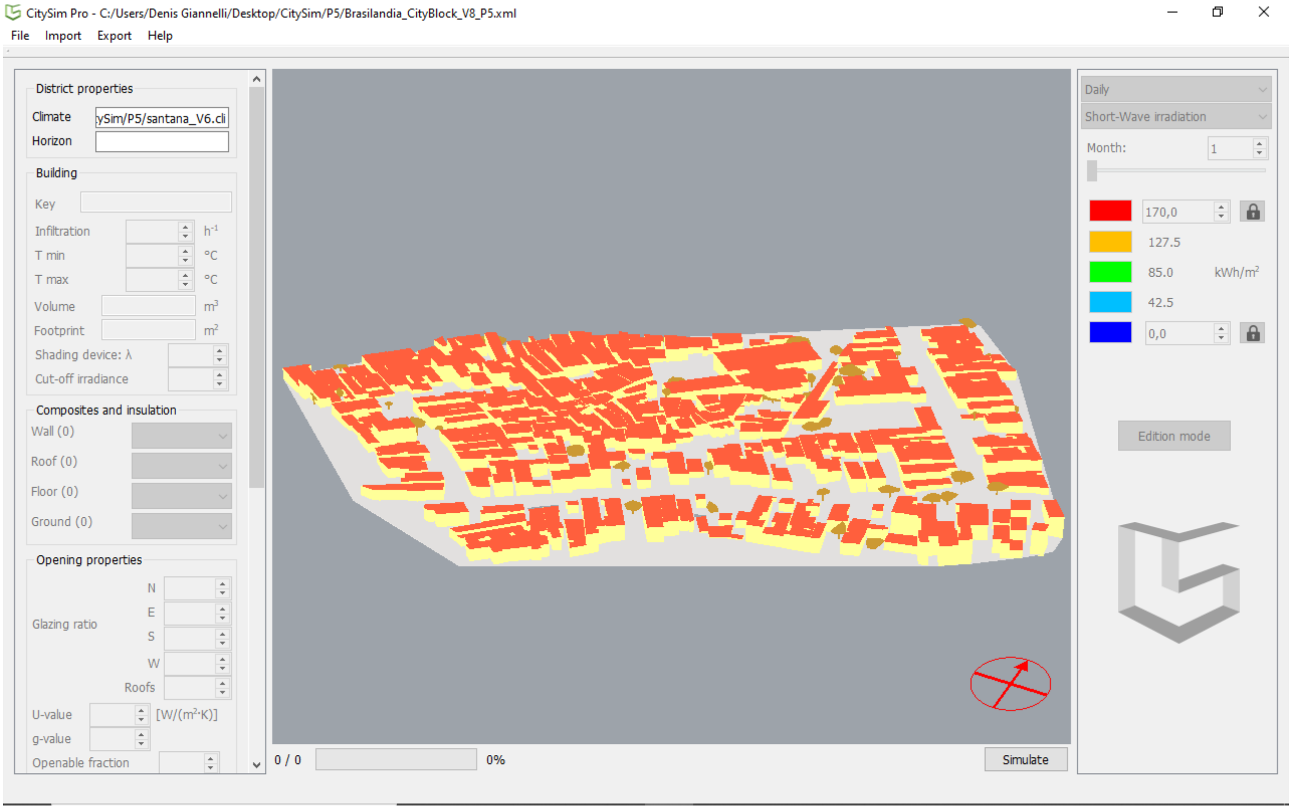


Figure 4-18 CitySim – importing CityGML buildings, trees and relief features.

In parallel with the geometrical input dataset, CitySim also requests two other input datasets, namely a climate file (.CLI) and a horizon file (.HOR).

The climate file is a TAB-separated-value data structure that encodes location-based climate parameters for every hour in the length of a typical year, and therefore it encodes 8760 records. This sort of file can be collected via commercial software, such as Meteonorm, but at the expense of high fees even with a student license. Therefore, the author creates their own climate file based on the two meteorological datasets of Mirante de Santana. The following principles are adopted:

- The climate values of specific timestamps throughout the original timeseries of the meteorological datasets are aggregated according to the hour of the year (HOY, from 0 to 8759) that they correspond. This is performed by averaging multiple records across different years that are available in the dataset.
- In the case of wind velocity and direction, these attributes can be understood as the magnitude and direction of a wind vector, and therefore averaging their values is performed in terms of a vectorial calculus.

- Ground temperature data, which are mandatory in a climate file, are not available in the datasets of São Paulo. Therefore, the same value of hourly average air temperature is used for both air temperature (Ta) and surface temperature (Ts) in the climate file.
- Cloud nebulosity data, also mandatory in a climate file, are available only for three timestamps per day. Therefore, a python script was designed to perform a series of vario-length but linear interpolations between existing values in the original dataset. Although this solution is not ideal, it is preferred than replacing null values with the general average.

The following table presents the conversion between the variable names and their unit of measurements, in the original datasets of São Paulo, and their counterparts in the designed climate file. Additionally, the conversion factor is present when this is necessary.

São Paulo dataset			Climate file	
Variable	**U.O.M.**	**Factor**	**Variable**	**U.O.M.**
Data Medição	DD/MM/YYYY	-3 hours (DD)	dm	Integer [1,31]
Data Medição	DD/MM/YYYY	-3 hours (MM)	m	Integer [1,12]
Hora Medição	HH*100	-3 hours/100	h	Integer [1,24]
Rad. Global	[Kj/m²]	*3600/1000	G_h	Float [W/ m²]
Temp. ar bulbo seco	[ºC]	-	Ta	Float [ºC]
Temp. ar bulbo seco	[ºC]	-	Ts	Float [ºC]
Vento velocidade	[m/s]	vector average	FF	Float [m/s]
Vento direção	[º]	vector average	DD	Integer [º]
Umidade relativa	[%]	-	RH	Integer [%]
Precipitação total	[mm]		RR	Float [mm]
Nebulosidade	[decimal]	*8/10	N	Integer [Octas]

```
Sao Paulo - Mirante de Santana
-23.49629154,-46.62009244,802.78,-3

dm  m   h   G_h Ta  Ts  FF  DD  RH  RR  N
1   1   1   0.0 21.9    21.9    0.5 73  78  0.0 6
1   1   2   0.0 21.8    21.8    0.6 45  79  0.0 6
1   1   3   0.0 21.5    21.5    0.3 64  81  0.1 6
1   1   4   0.0 21.4    21.4    0.5 101 81  0.1 6
1   1   5   0.0 21.2    21.2    0.5 81  81  0.0 6
1   1   6   0.0 21.1    21.1    0.2 96  82  0.1 6
1   1   7   20.1    21.5    21.5    0.6 78  79  0.0 6
1   1   8   136.8   22.3    22.3    0.4 116 76  0.0 6
1   1   9   316.8   23.2    23.2    1.3 117 73  0.1 6
1   1   10  502.1   24.2    24.2    1.3 117 70  0.0 6
1   1   11  611.7   25.0    25.0    1.8 138 66  0.2 6
1   1   12  638.7   26.2    26.2    1.8 134 63  0.6 6
```

Figure 4-19 CitySim – a fragment of the climate file.

The second line of the file stores the latitude, longitude, altitude (m) and the meridian of the site. Records range from the hour of the year (HOY) 0 (1 1 1) to 8759 (31 12 24).

The horizon file, in turn, is a simpler TAB-separated-value data structure that "investigates the far field obstruction (as the mountains) of the skyline." (MUNTANI et al., 2018). It encodes only two variables, namely the phi and theta angles, both in degrees. The former defines the horizontal angle of the line of sight, and it is encoded as anti-azimuth (-180° = NORTH, -90° = EAST, 0° = SOUTH, +90° = WEST, +180° = NORTH). The latter evaluates the vertical angle that a far field obstruction produces in the sky dome.

```
-144    5.49
-143    7.59
-142    8.21
-141    8.29
-140    7.85
-139    7.85
-138    7.64
-137    7.51
-136    7.51
-135    5.05
```

Figure 4-20 CitySim – A portion of an horizon file that contains phi (left) and theta (right) angles.

In the case of the solar analysis conducted in São Paulo, both for Santana and Brasilândia, there are no **far field** obstructions since the existing obstructions are produced by the geometries within the model, i.e., **near field** obstructions. Analogously to the horizon maps in GRASS GIS, the far field obstructions caused by the surrounding relief could have been considered but, if this was case, the replicability of the same input data would not have been possible among all software.

Considering that each one of the two scenarios of the present research relates to a different urban context, the horizon files must be considered independently. These are computed based on the raster-based DSM files with the support of the r.horizon module of GRASS GIS, i.e., the same module that was used in the analysis with GRASS.

In the case of Santana, the analysis of phi and theta values indicates that only two features in the scenario produce minimal obstruction to the meteorological station: the first one is a collection of trees towards the phi angle of -112, which lie behind the administrative building and whose crowns are indeed slightly higher than the meteorological station itself; the second one is a nearby streetlamp post towards the phi angle of -54. It is important to emphasise, however, that most of these near field obstructions (buildings, trees and relief) are already incorporated within the geometrical model itself.

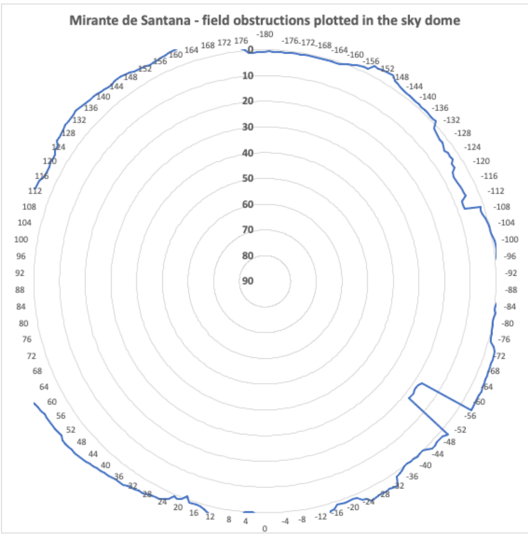


Figure 4-21 Santana – field obstructions in the sky dome.

Figure 4-22 Santana – field obstructions in ortophoto.

In the case of Brasilândia, however, producing a horizon file based on near field obstructions within the model is not reasonable. For the purposes of the present MSc thesis, each building roof in the city model of Brasilândia would require at least its own horizon file, since these near field obstructions change significantly according to the position and shape of the neighbouring buildings. Therefore, albeit it is possible to produce a horizon file based on the centroid of the CityGML model of Brasilândia, for instance, in practical terms the resulting phi and theta angles are not applicable to all buildings in the model, especially those in the borders of the model. The simulation in Brasilândia runs therefore without a horizon file.

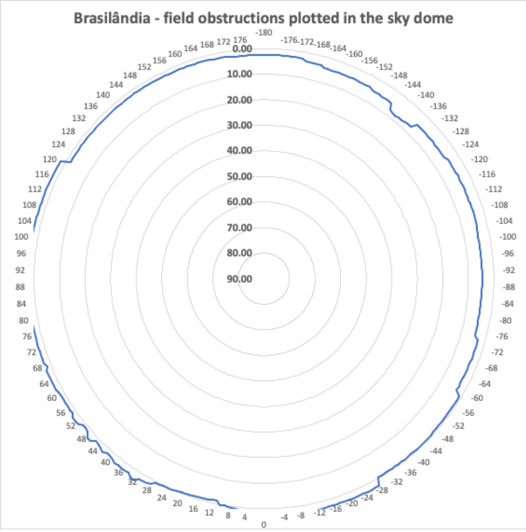


Figure 4-23 Brasilândia – phi and theta angles for the centroid of the city model.

Such horizon data, based on near field obstructions, are exclusively for the point that generates them, and therefore are not considered in the simulation with CitySim.

Figure 4-24 Brasilândia – centroid of the city model.

The horizon data (on the left) is only applicable to the point under consideration (on the right) since the neighbouring buildings procude different results at every position.

Once the three required datasets are produced, the simulation process is run in CitySim. The process is triggered on a one-click basis and the user is allowed to keep track of the computation steps and the estimated computing time. When concluded, the results of solar irradiation (along with other variables) can be visualised in the guided user interface according to the desired time granularity, i.e., hourly, daily, monthly, or yearly.

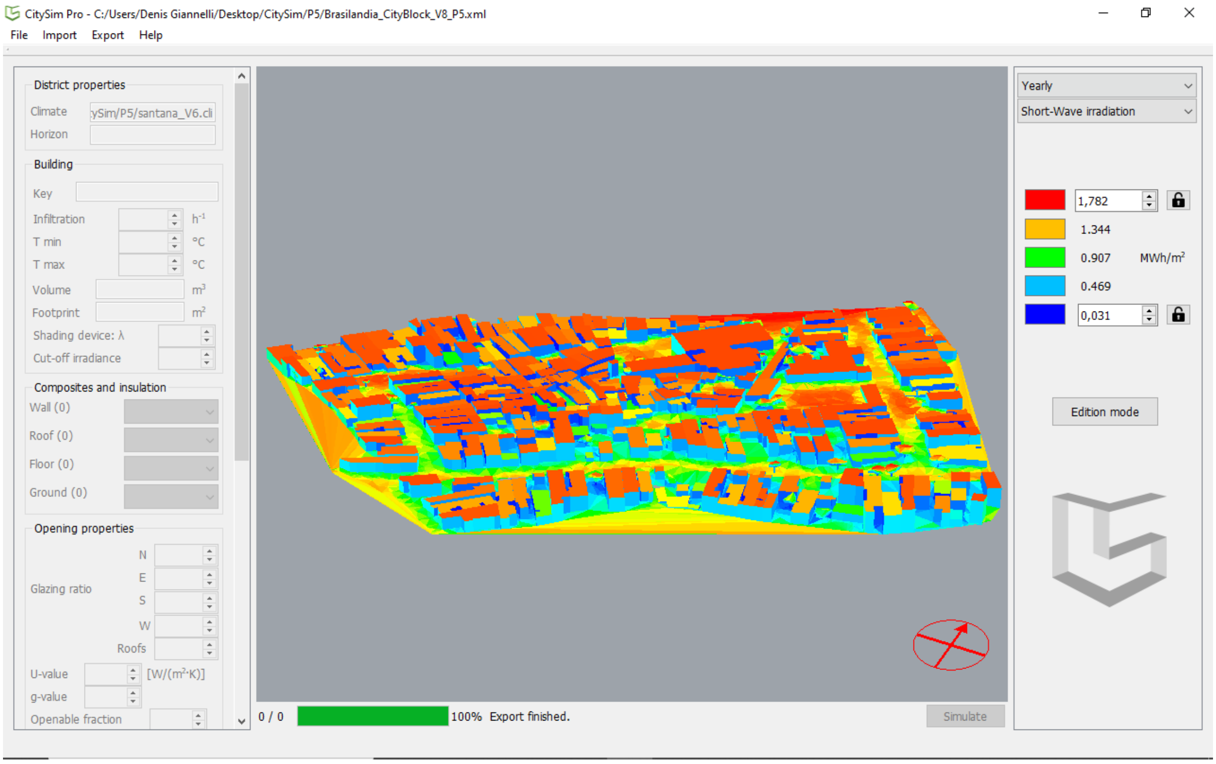


Figure 4-25 CitySim – annual solar irradiation values.

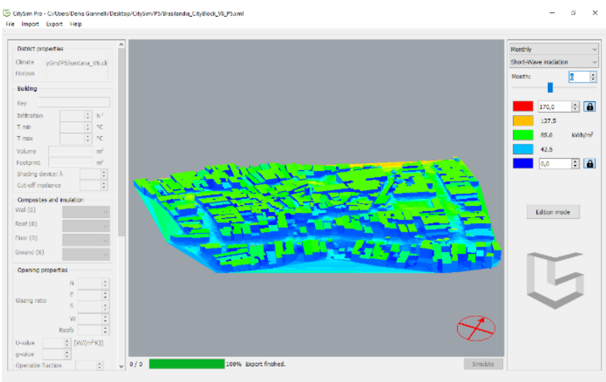


Figure 4-26 CitySim – solar irradiation values – JUN.

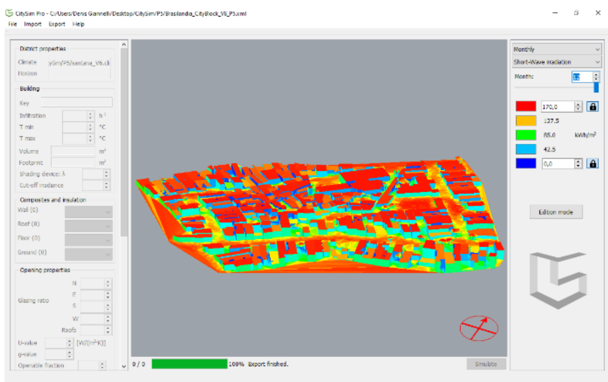


Figure 4-27 CitySim – solar irradiation values – DEC.

The results are collectively exported in tab-separated-value files, one of which (SW) storing hourly values of daily irradiance in W/m² per surface of the model – including terrain, tree and building surfaces. Considering that, within the time interval of one hour, the value of solar irradiance in W/m² is equivalent to that of solar irradiation in Wh/m², the hourly records are summated into daily, monthly, and yearly values of energy per square meter. With the results available, it is possible to respond the sub-research questions regarding CitySim:

- **What are the minimum data requirements to run the solar irradiation module?**

CitySim requires an xml-encoded CitySim file, which can be automatically derived from a CityGML file. In addition, it also demands a climate file and allows the use of a horizon file.

- **How automatized is the solar irradiation pipeline? From a user perspective, how complex is setting up the simulation?**
  The software runs on a one-click basis if all datasets are already available. The creation of climate and horizon files is performed outside of CitySim.

- **Is the solar irradiation module self-contained, or is it part of a greater package of solutions?**
  The solar irradiation module is part of a greater package of solutions present in CitySim Solver. In the current version of CitySim, it is not possible to opt for running only the solar irradiation simulation.

- **What is the computational running time of the module if this metric is retrievable?**
  CitySim PRO presents an estimating computing time by the beginning on the simulation, which is not preserved afterwards. For the scenario of Brasilândia, the simulation ran in approximately three hours, excluding the exporting process. Part of this amount of time is explained since CitySim processes all nine variables in just one workflow. Additionally, in the case of Santana, it took seven and a half hours to run the whole model. This difference is due to a greater number of tree surfaces in the model of Santana.

- **What kind of urban features – among buildings, vegetation, and relief – does the module support, and how do these features participate in the simulation?**
  In the present version of CitySim, the importing mechanism supports buildings, solitary vegetation objects (trees) and relief components encoded in CityGML. In the case of solar irradiation, all imported surfaces have their solar irradiation values calculated.

- **What is the degree of flexibility offered by the solar module in terms of selecting which features of the city model will be assessed and which will take part only as shading elements?**
  It is not possible to distinguish between surfaces for which irradiation values will be calculated and surfaces that will only act as far field obstructions. The horizon file creates obstruction angles on the sky dome, but for a general position and considering far field elements.

- **How is the output data delivered by the module in terms of time granularity and data model?**
  Hourly solar irradiation values are delivered in a table-separated-value dataset.

4.4 SimStadt

SimStadt is a free-of-charge and open-license urban simulation software developed at the German Stuttgart Technology University of Applied Sciences, and a homonymous project that was conducted by the same institution and concluded in 2015. According to the developers, "the application scenarios range from high-resolution simulations of building heating requirements and potential studies for photovoltaics to the simulation of building refurbishment and renewable energy supply scenarios[24]". Despite delivering versions for Windows, Linux and macOS, SimStadt requires a couple of additional software to be installed beforehand, and some of these are Windows-based only. Therefore, the experiments of the present research are also conducted in a bootcamp partition of the same computer for the methodological consistency. The version that is used in the present research is SimStadt 0.10.0-SNAPSHOT (refactor, rev. 385eb0c, 20210506).

The software offers a Guided User Interface, in which the user creates a unique data repository, their projects and all data workflows that are associated to a specific project. The repository structure is presented by the developers in the following SimStadt schema. The algorithms, nevertheless, are run in a command-line interface, from which the user may track the steps of the workflow that is being run.

[24] Documentation of SimStadt.

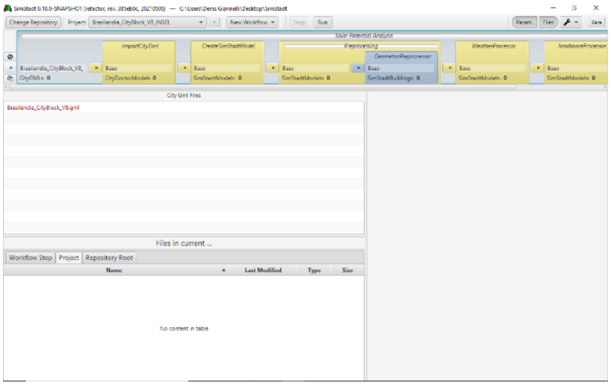


Figure 4-28 SimStadt – guided user interface.

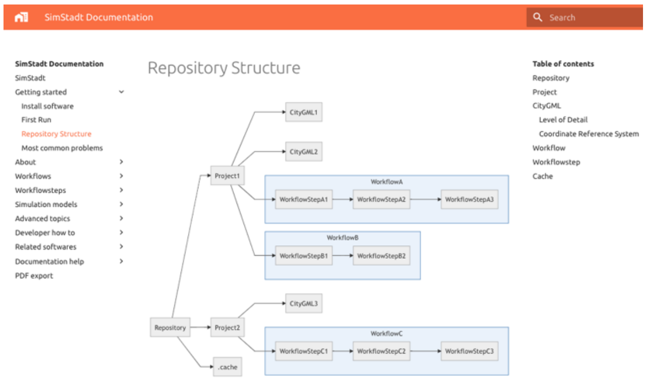


Figure 4-29 SimStadt – repository structure.

Source: Simstadt documentation.

Therefore, unlike CitySim Pro, which runs all simulations of CitySim Solver combined, SimStadt defines specific workflows for different analyses, namely: Biomass analysis, District heating network analysis, Environmental analysis with refurbishment strategy, Environmental analysis, Heat demand analysis with EnergyADE writer, Heat demand analysis with refurbishment strategy, Heat demand analysis, Hourly heat demand analysis with energy system, Hourly heat demand analysis, Photovoltaic potential analysis, Photovoltaic potential financial analysis and Solar potential analysis.

Although a photovoltaic analysis is available on SimStadt, the present research focuses on the Solar Potential Analysis workflow, since this one “calculates global solar irradiance on every [building] surface”[25], and therefore the results coming from this software can be compared to the results coming from the other solar modules. The steps of the Solar Potential Analysis workflow are illustrated below and described sequentially for the scenario of Brasilândia. The same strategies are adopted for the one of Santana.

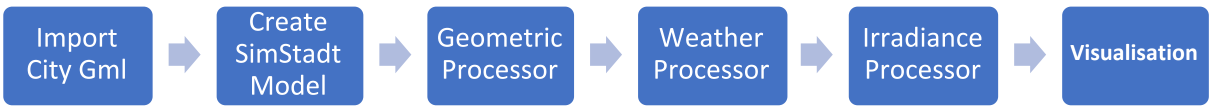


Figure 4-30 SimStadt – Solar Potential Analysis data pipeline.

Firstly, the CityGML file of Brasilândia is read and imported to SimStadt. Since all buildings of the CityGML file are written in LoD2, these are imported as such and validated against

[25] SimStadt Solar Potential Analysis documentation

the building schema of CityGML. In this step, it is already noticeable that only buildings are imported to SimStadt. Solitary vegetation objects (trees) and Relief components (relief) present in CityGML are neglected.

Secondly, a SimStadt model is created, with the option of maintaining building parts – buildings that are composed of multiple sub parts – or keeping just the buildings. In the case of Santana and Brasilândia, there are no building parts, so just buildings are kept.

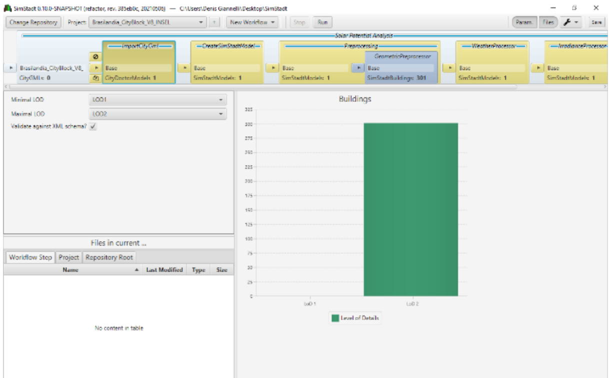


Figure 4-31 SimStadt – importing CityGML.

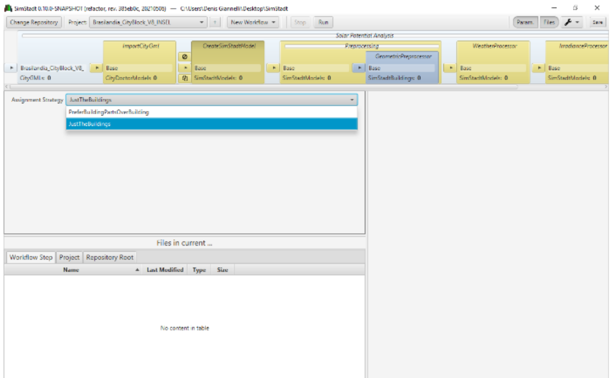


Figure 4-32 SimStadt – create SimStadt model.

Thirdly, the geometric processor retrieves important information from the buildings, such as the orientation of their surfaces, their height, area, volume, and roof type. It is important to emphasise that, until this point, no impacting decision needs to be taken for the purpose of solar irradiation analysis.

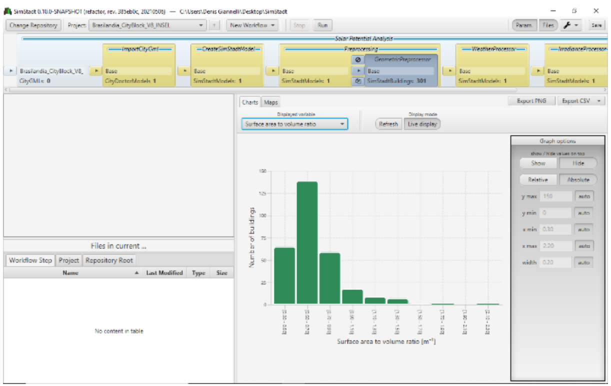


Figure 4-33 SimStadt – geometric processor – area/volume ratio.

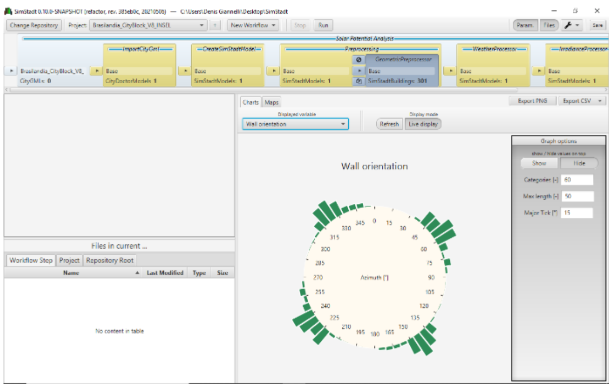


Figure 4-34 SimStadt – geometric processor – wall orientation.

Sequentially, in the weather processor workstep, SimStadt retrieves weather data for the location of the model and creates synthetic hourly values from monthly (or hourly) means. The user needs to choose the weather data source, and four main options are possible:

- Monthly global horizontal irradiance and ambient temperature from the PVGIS online database.
- Monthly global horizontal irradiance and ambient temperature from the INSEL offline database.
- Hourly global horizontal irradiance and ambient temperature from a local TMY3 file.
- Monthly global horizontal irradiance and ambient temperature from a local TMY3 file.

In the case of PVGIS, the workstep runs successfully for the scenario of Brasilândia, but not for the one of Santana. The exception raised in the command line is that "monthly weather values don't seem to be plausible for this location. Cannot generate hourly values". In previous experiments conducted by the author that made use of larger CityGML models of Santana, the PVGIS weather data option was possible. Nevertheless, when the final CityGML model of Santana was input, this error is raised, and the simulation is terminated. It is possible to conclude, therefore, that the spatial query performed in the PVGIS database does not retrieve weather data for the reduced extents of the final CityGML model of Santana, but only for its larger counterpart, which incorporates adjacent areas of the city of São Paulo. Therefore, the option of this weather database is excluded since the PVGIS values of Santana cannot be confronted with the ground truth ones from the INMET station.

The INSEL offline database is the one present in INSEL8, i.e., one of the software that needs to be installed before SimStadt. This is a software for simulation, monitoring, and visualization of energy systems. This option handles the analyses in both Santana and Brasilândia. However, for the purpose of this research, adopting an external weather database – while it is possible to design a TMY3 file from ground truth data – compromises the methodological consistency of the research. The input irradiance values of INSEL would differ from the ones of INMET, as it is already the case of the experiments conducted on GRASS GIS and ArcGIS, which make use of their own weather databases. Nevertheless, with the possibility of inputting user-defined weather data, the author understands that the comparison between SimStadt and ground truth values becomes unbiased.

Lastly, there is the possibility of extracting hourly or monthly values from local TMY3 files. Analogously to the climate files (.CLI) for CitySim, hourly TMY3 files (.TMY3) are structured in hourly records containing weather attributes. However, this sort of file is much more complex than a climate file requested for CitySim, as it demands a total of 68 parameters, among weather data, data flags and degrees of certainty of the weather data.

By hands-on experience, debugging of test TMY3 files and analysis of TMY3 files available for other cities, it is concluded that:

- Apart from the date (MM/DD/YYYY) and time columns (HH:MM), other essential attributes for which values need to be incorporated in the TMY3 file are: 'Global horizontal irradiance', 'Direct normal irradiance', 'Diffuse horizontal irradiance' and 'Dry-bulb temperature'. The meteorological station of São Paulo does not offer direct and diffuse horizontal irradiance values, but rather only global horizontal ones. In this case, a workaround needs to be made:
    - From the Energy Plus website, it is possible to download an EPW file designed with weather data collected from the same meteorological station (837810 – INMET) and based on averaged values of records from 2001 to 2010, i.e., the previous decade with respect to the ground truth dataset adopted in this research.
    - This EPW file contains typical hourly records of global, direct, and diffuse horizontal irradiance. The summations of direct and diffuse values, however, do not reach the summation of global irradiance, respectively 35.98% and 51.47% of this last one, and therefore these sub-divided irradiance components correspond to only 87.45% of the global irradiance one.
    - In order to adapt the ground truth dataset of the present research to the requirements of a TMY3 file, an assumption is made to derive direct and diffuse values but respecting the referential EPW file. In this case, for every typical hour of the year (HOY) record present the newly generated EPW file, the global irradiance value comes directly from the ground truth dataset of the research, whereas the direct and diffuse irradiance values are defined as this

global irradiance value multiplied by the hourly proportion of these parameters in the referential EPW file.

- By proceeding with this assumption, the resulting TMY3 file stores the exact same global irradiance values as the ones present in the ground truth dataset and also in the Climate file prepared for CitySim. The summations of derived direct and diffuse irradiance values result in 26.21% and 60.87% of the summation of global irradiance values, respectively, which combined reach 87.08% of the amount of global irradiance.
- In practical terms, SimStadt is input with a TMY3 file based on an assumption that, albeit not ideal, is the best compromise between the ground truth dataset and the weather data requirements of SimStadt. It would have been much preferred to adopt only global irradiance values – as in the case of CitySim, which handles simulations when diffuse and direct parameters are unknown – , but this is not possible in SimStadt or, if it is, the documentation of the software does not address the issue.

- Apart from the direct and diffuse irradiance values, which had to be derived with the aforementioned assumption, all the other attributes present in the meteorological dataset of INMET which are also requested in the TMY3 file were written accordingly, adjusting the unit of measurement if this was the case.

- For the remaining attributes requested in the TMY3 schema that are not present in the meteorological dataset of INMET, the author writes their corresponding values as -9990, i.e., no data values. This is possible since the weather processor of SimStadt does not need all the attributes contained in a TMY3 file.

```
,Sao Paulo,,-3,-23.49629154,-46.62009244,802.78,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,
Date (MM/DD/YYYY),Time (HH:MM),ETR (W/m^2),ETRN (W/m^2),GHI (W/m^2),GHI source,GHI uncert (%),DNI (W/m^2),DNI sourc
01/01/2021,01:00,0,0,0.0,3,0,0.0,3,0,0.0,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,21.9,3,0,1
01/01/2021,02:00,0,0,0.0,3,0,0.0,3,0,0.0,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,21.8,3,0,1
01/01/2021,03:00,0,0,0.0,3,0,0.0,3,0,0.0,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,21.5,3,0,1
01/01/2021,04:00,0,0,0.0,3,0,0.0,3,0,0.0,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,21.4,3,0,1
01/01/2021,05:00,0,0,0.0,3,0,0.0,3,0,0.0,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,21.2,3,0,1
01/01/2021,06:00,224,35,0.0,3,0,0.0,3,0,0.0,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,21.1,3,
01/01/2021,07:00,531,200,20.1,3,0,0.2,3,0,19.7,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,21.5
01/01/2021,08:00,818,474,136.8,3,0,8.1,3,0,122.7,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,22
01/01/2021,09:00,1065,802,316.8,3,0,101.5,3,0,182.4,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0
01/01/2021,10:00,1254,1112,502.1,3,0,289.6,3,0,175.6,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,
01/01/2021,11:00,1372,1333,611.7,3,0,404.0,3,0,195.7,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,
01/01/2021,12:00,1413,1413,638.7,3,0,467.3,3,0,171.4,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,
01/01/2021,13:00,1372,1333,800.5,3,0,564.9,3,0,218.9,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,0,-9990,3,
```

Figure 4-35 SimStadt – Fragment of a TMY3 file designed with hourly records.

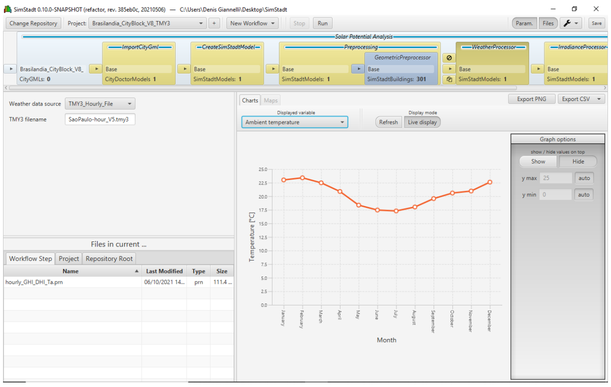


Figure 4-36 SimStadt – TMY3 – ambient temperature.

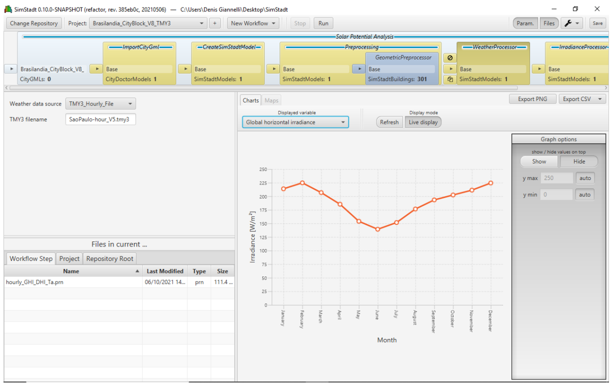


Figure 4-37 SimStadt – TMY3 – global hor. Irradiance.

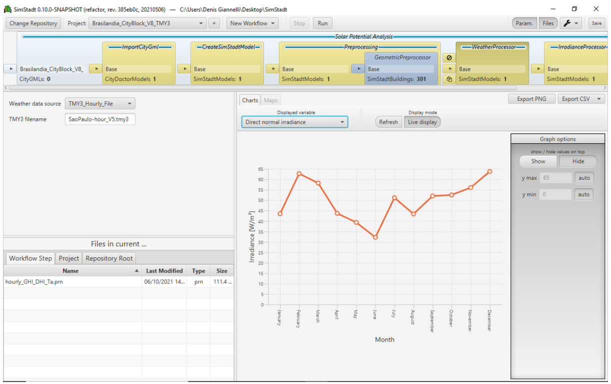


Figure 4-38 SimStadt – TMY3 – direct hor. irradiance.

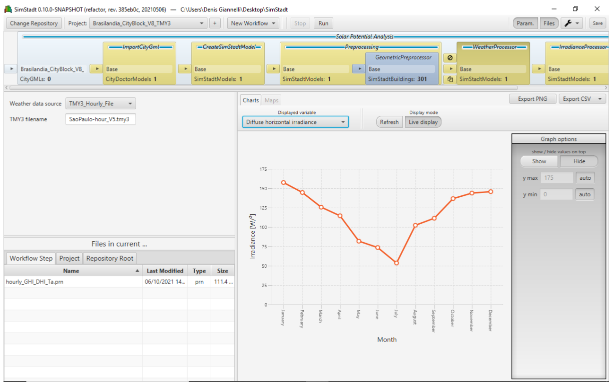


Figure 4-39 SimStadt – TMY3 – diffuse hor. irradiance.

Finally, the irradiance processor calculates hourly solar irradiance values for all wall and roof surfaces of the LoD2 buildings in the model. According to the documentation, there are five radiation models to be chosen, namely:

- Hay sky model, calculated with INSEL: fast, without shadows
- Perez sky model, calculated with INSEL: fast, more detailed, and possibly more accurate than Hay, without shadows
- Perez sky model, calculated with Simplified Radiosity Algorithm: can be very slow, should not be used for larger models, with shadows
- Perez sky model, calculated with Simplified Radiosity Algorithm on tiles: can be used for larger models (which will be split in smaller tiles), with shadows
- Perez sky model, calculated with Simplified Radiosity Algorithm on a geodesic dome: fast, can be used for comparisons with other SRA calculations, without shadows

The Hay and Perez models that are calculated with INSEL have the disadvantage of calculating, for each timestamp under analysis, the same irradiation value for all building roofs, since the shading effect that one building produces into another is not taken into consideration. Moreover, these do not offer the possibility of retrieving hourly irradiance values that are kept into cache files, unlike the Simplified Radiosity Algorithm (SRA).

The SRA is another software that needs to be installed prior to the installation of SimStadt. It creates a large file with hourly irradiance for every surface. This is possible by checking the ‘Keep hourly cache files’ in the GUI. Moreover, the ‘with shadows’ options allow the algorithm to consider the shading effect of one building into another. Since both Santana and Brasilândia models are relatively small, these can be run without tiling. Therefore, the option of Perez calculated with the Simplified Radiosity Algorithm (SRA) is chosen.

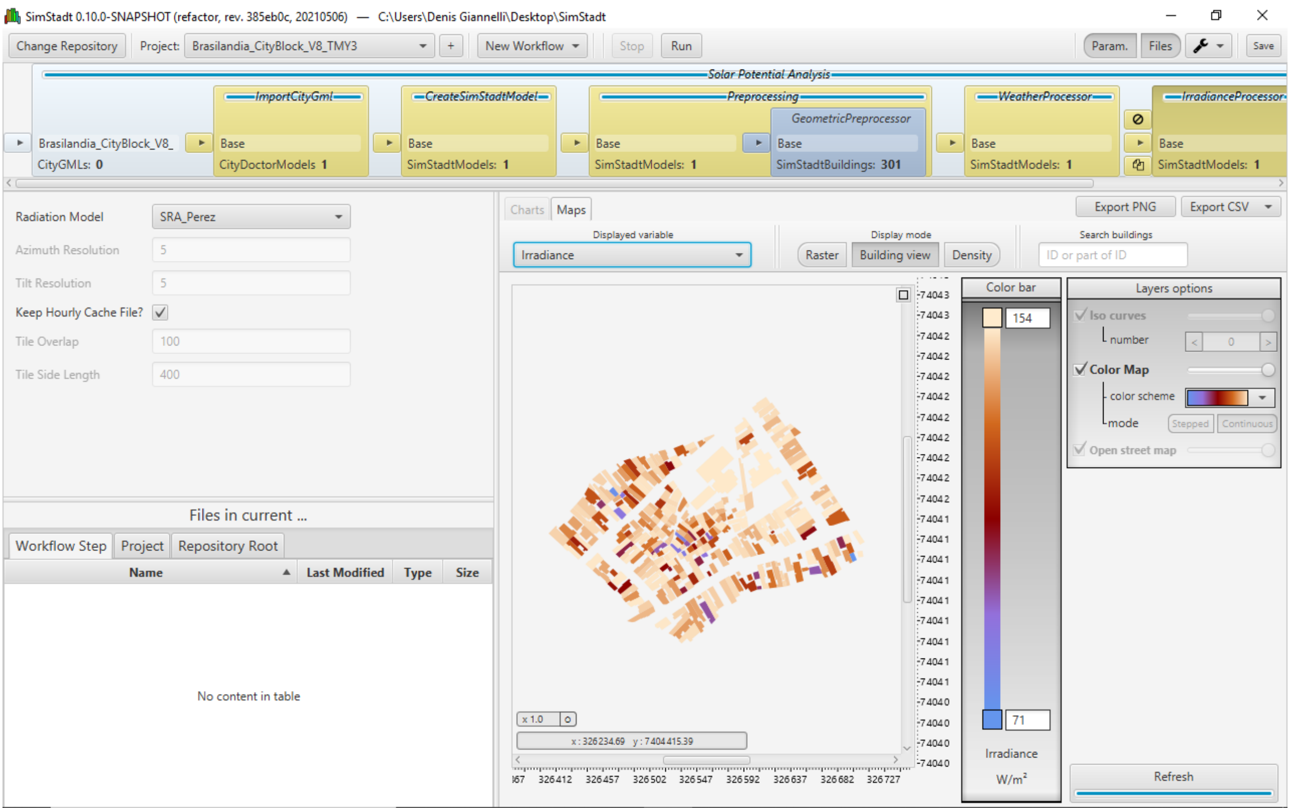
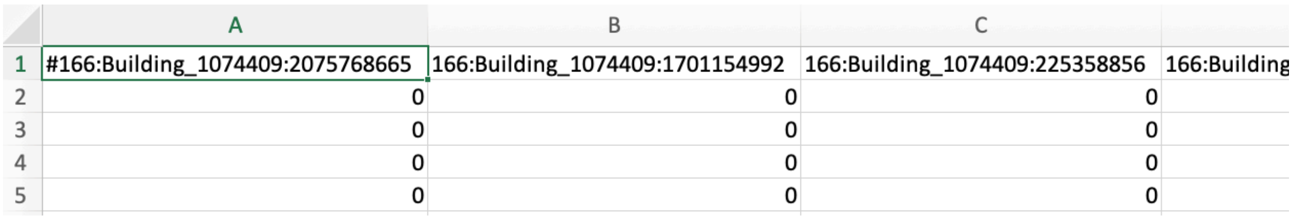


Figure 4-40 SimStadt – irradiance processor – the values differ for each building roof.

After running the algorithms on the command line, the guided user interface of the software displays only yearly averaged values of solar irradiance per building roof. The collection of hourly values of solar irradiance, in W/m², are retrievable in the cache files.

Nevertheless, a major problem encountered is that, in the internal process of SimStadt, the original CityGML ids of the polygons are not preserved. Therefore, for the purposes of this research, an assumption is made in the sense that, among all polygons belonging to the same Building GML id, the one that has the greater summation of solar irradiation is the roof of the building. For sites located in low latitudes, such as the city of São Paulo, this assumption is possible since the solar rays are not as obliquus as they are in temperate surfaces of the planet.



	A	B	C	
1	#166:Building_1074409:2075768665	166:Building_1074409:1701154992	166:Building_1074409:225358856	166:Building
2	0	0	0	
3	0	0	0	
4	0	0	0	
5	0	0	0	

Figure 4-41 SimStadt – the structure of the .OUT cache files with hourly irradiance values.

It is possible to observe that the building GML id is present, whereas the GML id of the polygon or multisurface are absent.

With the results available, it is possible to respond the sub-research questions regarding SimStadt.

- **What are the minimum data requirements to run the solar irradiation module?**
  Strictly speaking, only the CityGML file of the scenario, with buildings modelled in LoD1 and/or LoD2. A user-defined weather dataset is not necessary since SimStadt offers the databases of PVGIS and INSEL. Nevertheless, a TMY3 file can be designed, which allows the user to keep control of the weather data that is input for the simulation. In this case, the irradiance data need to be segmented into hourly or monthly global, direct, and diffuse horizontal irradiance values.

- **How automatized is the solar irradiation pipeline? From a user perspective, how complex is setting up the simulation?**
  After configuring all parameters in the Solar Potential Analysis steps, SimStadt runs the workflow on a one-click basis and displays the results in the GUI. The hourly cache files are also stored automatically in the chosen repository for SimStadt.

- **Is the solar irradiation module self-contained, or is it part of a greater package of solutions?**

The Solar Potential Analysis workflow is self-contained, delivering only irradiance values. The intermediate worksteps can be re-used in other workflows.

- **What is the computational running time of the module if this metric is retrievable?**
  For the scenario of Brasilândia, computing the whole workflow results in a running time of HH:MM:SS = 00:08:06.

- **What kind of urban features – among buildings, vegetation, and relief – does the module support, and how do these features participate in the simulation?**
  Only buildings are supported by the software. If the Perez Radiation model is chosen in combination with the Simple Radiosity Algorithm, the building surfaces can act as both obstruction elements and as surfaces that are assessed.

- **What is the degree of flexibility offered by the solar module in terms of selecting which features of the city model will be assessed and which will take part only as shading elements?**
  There is no degree of flexibility in terms of distinguishing imported features. These will either be processed with shading effect or without it.

- **How is the output data delivered by the module in terms of time granularity and data model?**
  SimStadt delivers a .CSV file with annual averaged solar irradiance (W/m²) values per surface of the scenario, and a .OUT file with hourly values of solar irradiance. In the case of the hourly values, these do not have information regarding the GML id of the surface that they correspond to, which is a drawback of the software.

4.5 Ladybug

Ladybug Tools is a free-of-charge and open license collection of computer applications with the goal of supporting environmental design and education. These tools are implemented in Python programming language and maintained by a team of developers in a community-supporting approach, with the forum of discussion being the most resourceful documentation of their tools.

This collection of tools is composed by four main applications, namely: Ladybug, Honeybee, Bufferfly and Dragonfly. Among these, Ladybug – the plugin – is the one adopted for the present research. According to the developers, Ladybug “performs detailed analysis of climate data to produce customized, interactive visualizations for environmentally informed design”, such as radiation studies that quantifies the amount of energy incident at surfaces of the model given a time interval.

Ladybug and all other tools are conceived as a plugin for Grasshopper 3D, which is “a graphical algorithm editor tightly integrated with Rhino’s 3-D modelling tools[26].” Rhino – short for Rhinoceros 3D[27] – is a commercial 3D computer graphic and 3D computer-aided design (CAD) software developed by Robert McNell & Associates. The CAD environment of Rhino combined with the parametric design capabilities of Grasshopper makes these platforms attractive for professional from the design domain – such as Architects, Urban Planners / Designer and Landscape Architects – who seek strong data visualisation tools with minimal programming effort.

In short, for the solar irradiation analysis with Ladybug, the present research makes use of a student license of Rhinoceros 7.10.21256.17002 2021-09-13, combined with Grasshopper 1.0.0007 2021-09-13 and Ladybug 1.2.0. All experiments are conducted on a macOS.

---

[26] Grasshoppper 3D website.

[27] Rhinoceros website.

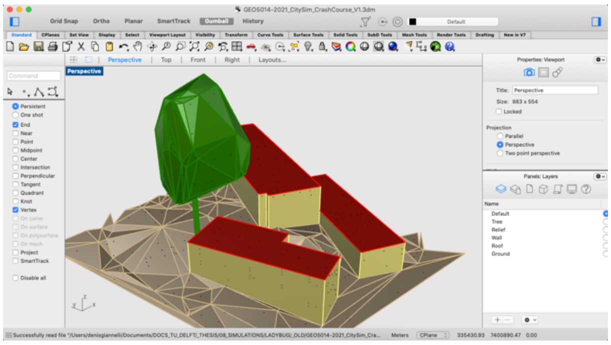


Figure 4-42 Ladybug – A simple example of a computer-aided design model in Rhino.

The geometries can be parametrically produced in Grasshopper – such as these ones – and then 'baked' in Rhino. When baked, these can also serve as geometrical input for environmental analysis in Grasshopper, creating therefore a two-way data pipeline.

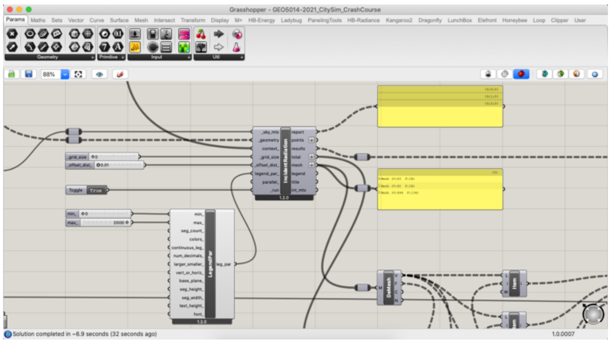


Figure 4-43 Ladybug – The workbench of Grasshopper with some Ladybug components.

Grasshopper operates in a workbench with several components ('boxes') and connectors ('cables') among these, which define the data pipeline. At every step, the geometrical results of individual components can be visualised in the Rhino environment.

For the experiments conducted in Rhino/Grasshopper/Ladybug, the main component adopted in the present research is *Incident Radiation*. According to the documentation contained in its Python script, this component calculates the incident radiation on geometry using a sky matrix. Therefore, to make use of this component, a double task is necessary: providing *Incident Radiation* with Rhino CAD geometries, and with a proper Sky Matrix. The workbench adopted in this MSc thesis is adapted from the work conducted in DOAN (2021), with the noticeable difference that the present research imports Rhino CAD geometries via an FME workbench instead of doing so via 3DCityDB. The main data pipeline is summarised in the following process diagram.

Figure 4-44 Ladybug – geometrical data pipeline.

As Rhino is a CAD environment, the first necessary step is importing all geometries and their GML ids, coming from the xml-encoded CityGML model, to a rhino model. A pragmatical solution consists of defining an FME workbench that reads all CityGML geometries and their attributes, and converts them into an organised DWG file, which is a native data structure for Rhino.

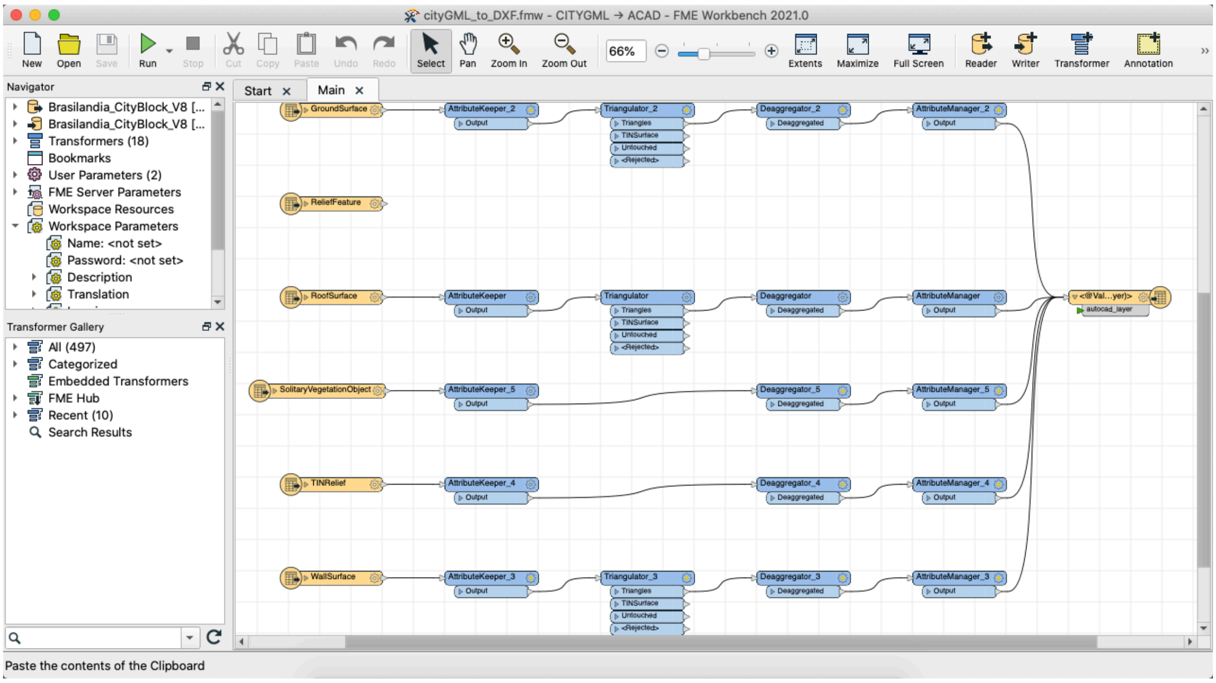


Figure 4-45 Ladybug - An FME workbench that converts CityGML to DWG.

With this approach, three important aspects need to be considered:

- The Solar Incident Radiation of Ladybug accepts both B-REP (boundary representation) and Meshes as its geometrical input. If the latter is chosen, the algorithm will output a single value of radiation per face of the mesh, which is positive since it avoids further geometrical aggregation, as it is the case when a B-REP is chosen and these boundaries are split into a grid structure.
- By hands-on experience, it is concluded that the Ladybug algorithm runs correctly only if the mesh is composed by triangles. Otherwise, these original polygons – such as roofs – are further split and the correspondence between ids and geometries is compromised. Therefore, the FME workbench also conducts a triangulation of the ground, wall, and roof thematic surfaces before these are imported into Rhino.

Figure 4-46 Ladybug – roof pre triangulation.

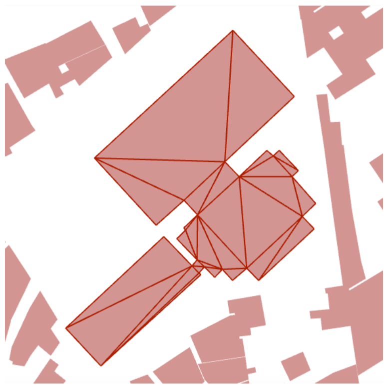


Figure 4-47 Ladybug – roof post triangulation.

- In the process of converting a CityGML to DWG, the sematic information regarding the id of the thematic surface or polygon needs to be preserved. This is possible since the GML ids of the thematic surfaces are passed in the DWG file as the name of the layer in which these geometries are stored. In the case of tree and/or relief polygons, these are stored with the id of the solitary vegetation object and relief component, respectively.

By the end of this process, all geometries and GML ids, which were originally present in the CityGML model, are stored in the Rhino CAD environment.

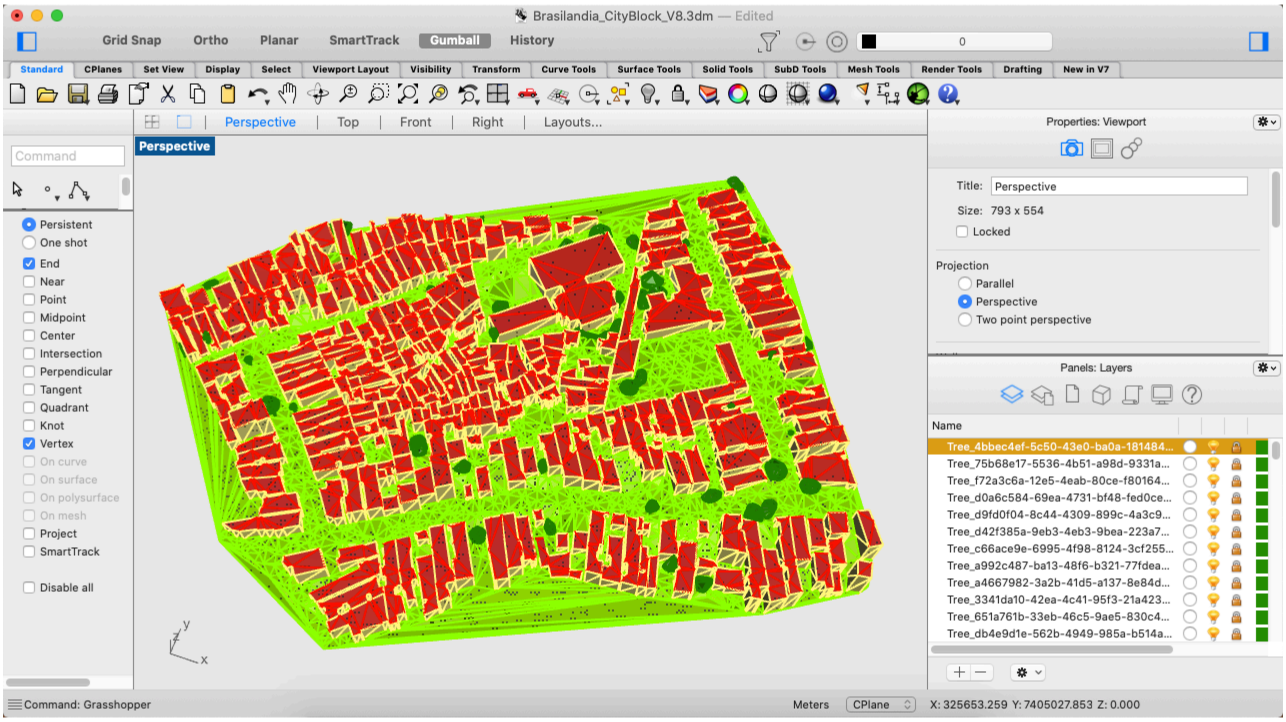


Figure 4-48 Ladybug – Geometries and attributes stored in Rhino.

Once these geometries are stored into Rhino, the following step is associating them with mesh components into Grasshopper. By associating, it is meant that the geodata will still be stored exclusively in the Rhino file, but the Grasshopper workbench will recognise the geometrical and attribute properties of the features when performing solar analysis via Ladybug.

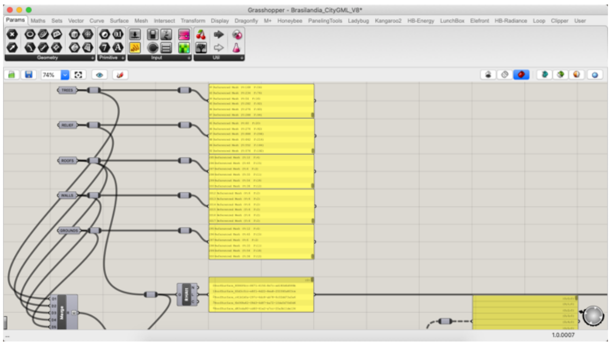


Figure 4-49 Ladybug – Rhino geometries and attributes referenced in Grasshopper according to their feature type.

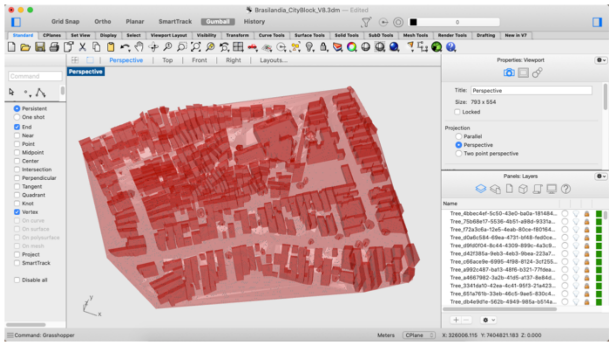


Figure 4-50 Ladybug - Grasshopper referenced meshes appear in the Rhino environment as 'virtual' geometries.

Regarding the geometrical input of Ladybug Incident Radiation, a positive aspect can be mentioned about its flexibility. The component allows the user to distinguish between features that will be considered as 'geometries', i.e., "Rhino Breps and/or Rhino Meshes for which incident radiation analysis will be conducted", and 'context', i.e., "Rhino Breps and/or Rhino Meshes representing context geometry that can block solar radiation to the test geometry."

This is the only software, among all under study in this research, in which the user may select exactly which geometries will be assessed and which will not, thus reducing unnecessary computation time for features that do not matter. In practical terms, for the case of Santana, only the single roof of the pseudo building – representing the meteorological station – will be assessed, and all other geometries in the model will act as obstruction elements. In the case of Brasilândia, only the roofs of all buildings will be assessed, since the values obtained from these are the only ones that can be compared to the ones coming from other software.

After providing the Ladybug Incident Radiation component with Rhino geometries, the second task is providing it with a so-called Sky Matrix, which stores direct normal and diffuse horizontal radiation values coming from each patch of a sky dome. The weather data pipeline for ladybug is represented below:

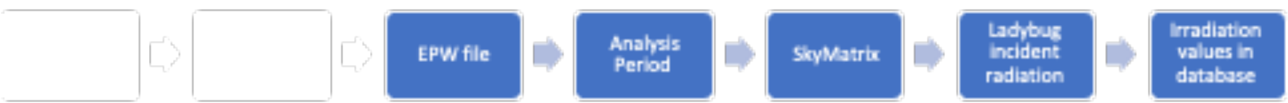


Figure 4-51 Ladybug – Weather data pipeline

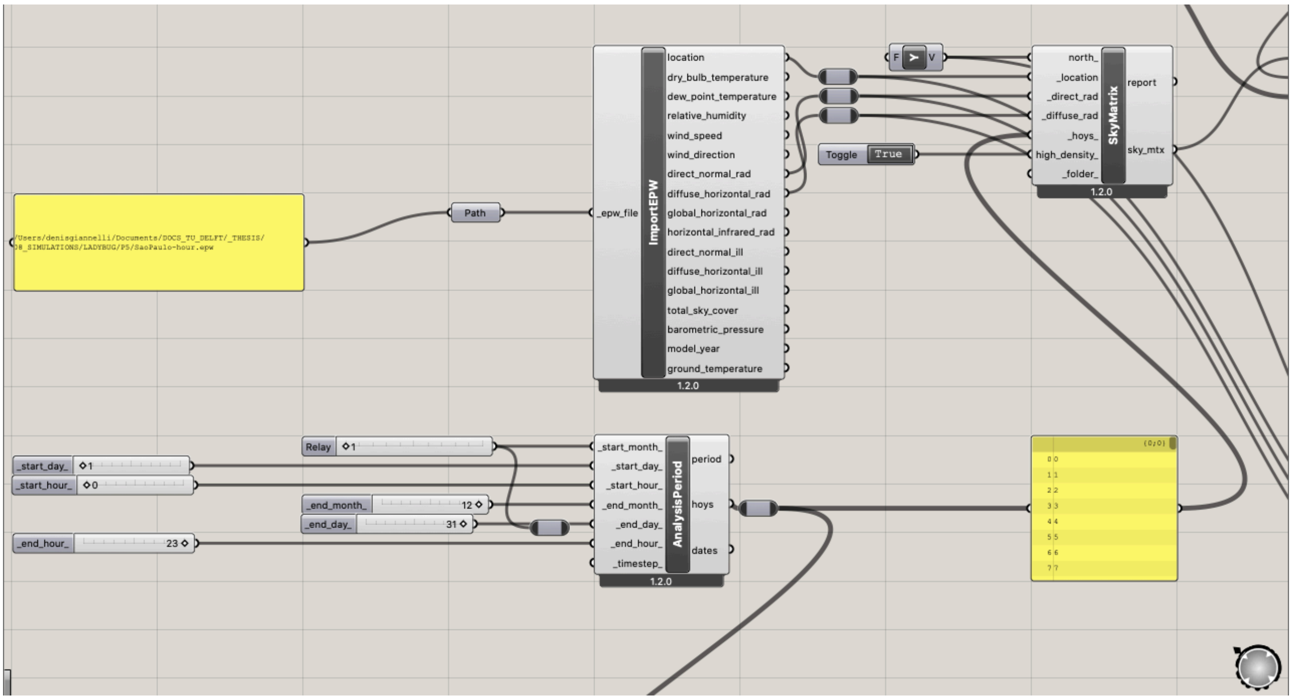


Figure 4-52 Ladybug – A portion of the workbench that handles weather data.

The Sky Matrix can be constructed with another Ladybug component named SkyMatrix which, according to the developers of Ladybug, "uses Radiance's gendaymtx function to calculate the radiation for each patch of the sky. Gendaymtx is written by Ian Ashdown and Greg Ward[28].

SkyMatrix requires location, direct normal radiation, and diffuse horizontal radiation data coming from a so-called EPW file. This last is a structured weather file with hourly records within a typical year, containing values from multiple attributes, analogously to the case of a TMY3 file required by SimStadt. For the present research, the author designed its own EPW file based on the ground truth dataset from INMET. Just as in the case of the TMY3 file, the EPW file also demands the distinction among global, direct, and diffuse horizon irradiation (Wh/m2), and therefore the proportional distribution of global irradiation values between direct and diffuse values was adopted, as it was carried out for the TMY3 file for SimStadt.

28 Documentation of gendaymtx.

```
LOCATION,Sao Paulo,SP,BRA,,,-23.49629154,-46.62009244,-3,802.78
DESIGN CONDITIONS,1,2017 ASHRAE Handbook -- Fundamentals - Chapter 14 Climatic Design Information,,Heating,7,9,10.2
TYPICAL/EXTREME PERIODS,6,Summer - Week Nearest Max Temperature For Period,Extreme,2/10,2/16,Summer - Week Nearest
GROUND TEMPERATURES,3,.5,,,,22.10,21.61,20.68,19.80,18.15,17.29,17.06,17.51,18.54,19.82,21.06,21.88,2,,,,21.45,21.3
HOLIDAYS/DAYLIGHT SAVINGS,No,0,0,0
COMMENTS 1,""
COMMENTS 2,""
DATA PERIODS,1,1,Data,Sunday, 1/ 1,12/31
2021,1,1,1,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,21.9,17.8,78,92400,0,0,386,0.0,0.0,0.0,999,999,999,
2021,1,1,2,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,21.8,17.8,79,92300,0,0,384,0.0,0.0,0.0,999,999,999,
2021,1,1,3,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,21.5,17.9,81,92300,0,0,382,0.0,0.0,0.0,999,999,999,
2021,1,1,4,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,21.4,17.8,81,92300,0,0,381,0.0,0.0,0.0,999,999,999,
2021,1,1,5,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,21.2,17.8,81,92300,0,0,380,0.0,0.0,0.0,999,999,999,
2021,1,1,6,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,21.1,17.8,82,92300,224,35,381,0.0,0.0,0.0,999,999,9
2021,1,1,7,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,21.5,17.7,79,92300,531,200,384,20.1,0.2,19.7,999,99
2021,1,1,8,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,22.3,17.8,76,92400,818,474,389,136.8,8.1,122.7,999,
2021,1,1,9,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,23.2,18.0,73,92400,1065,802,400,316.8,101.5,182.4,9
2021,1,1,10,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,24.2,18.2,70,92400,1254,1112,408,502.1,289.6,175.6
2021,1,1,11,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,25.0,18.0,66,92400,1372,1333,415,611.7,404.0,195.7
2021,1,1,12,0,?9?9?9?9E0?9?9?9*9*9?9?9?9?9?9?9?9?9?9?9?9?9C9?9?9,26.2,18.1,63,92400,1413,1413,423,638.7,467.3,171.4
```

Figure 4-53 Ladybug – the structure of a EPW file.

The difference between adopting a TMY3 file in SimStadt and a EPW file in Ladybug, however, consists precisely in whether the “Global horizontal irradiance” (TMY3) and the “Global Horizontal Radiation” (EPW) values are considered in the simulation or not. Regarding SimStadt, it is not possible to derive neither from the guided user interface nor from the documentation if the Simplified Radiosity Algorithm takes the global horizontal irradiance attribute into account. When it comes to Ladybug, however, the workbench components and connectors make clear that only Direct Normal Radiation [Wh/m2] and Diffuse Horizontal Radiation [Wh/m2] will be passed into the SkyMatrix component. In practical terms, this means that the Ladybug Incident Radiation component will be based on a lower amount of solar radiation with respect to the ground truth dataset.

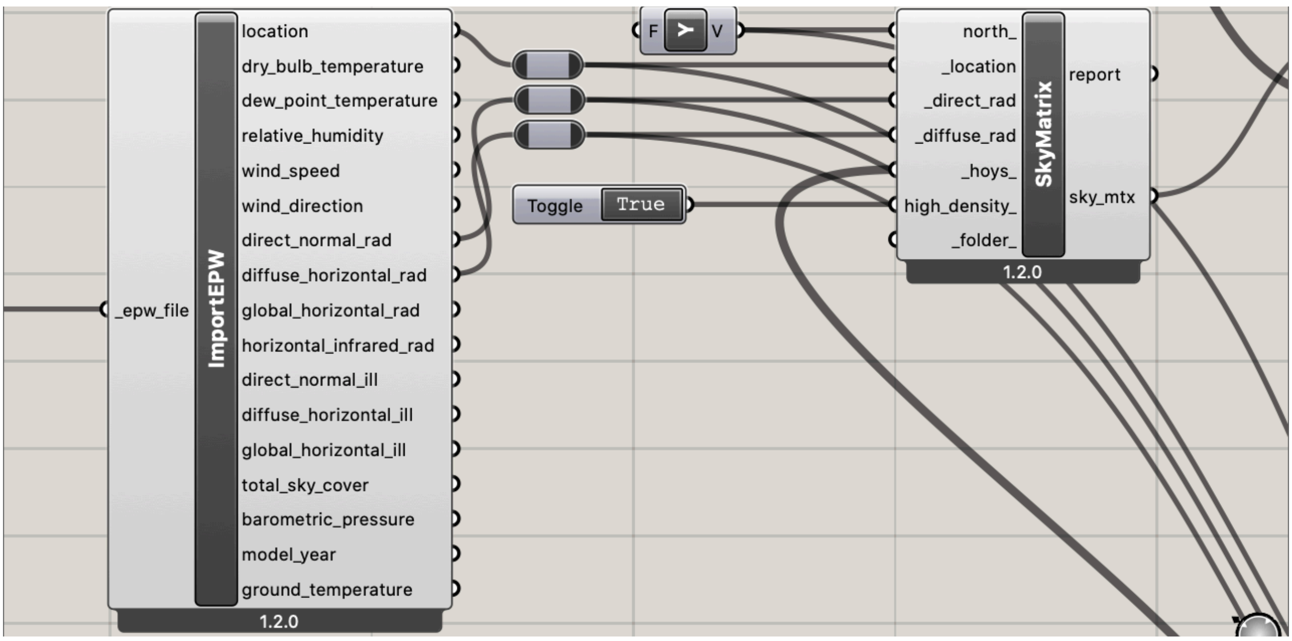


Figure 4-54 Ladybug - the SkyMatrix component does not take global horizontal radiation values into account.

Another input dataset requested by the SkyMatrix component is the HOYs, i.e., “a number or list of numbers between 0 and 8760 that represent the hour(s) of the year for which to generate the sky matrix. (…) By default, the matrix will be for the entire year.”

In the case of Brasilândia, as the present research only handles monthly irradiation values (but for multiple building roofs simultaneously), the grasshopper workbench is designed such that, for every month of the year, the HOYs list is manually adapted at each iteration. This manual procedure prevents an unnecessary and more complex data tree, in which the irradiation values would be structured according to the HOY (first level), building (second level) and surface/triangle (third level).

In the case of Santana, however, it is important to extract hourly or at least daily solar irradiation values, and this task requires an automatic pipeline, since it is feasible to manually export 12 monthly values, but neither 365 daily values nor 8760 hourly values. Nevertheless, the data tree does not consider the building data tree level, since there are only two surfaces that are assessed, i.e., the two triangles that represent the roof level of the artificial building of the meteorological station.

In this case, the workaround consists of implementing the workbench with an auxiliary component, the Ladybug Real Time Incident Radiation. This component takes into consideration two different inputs: the first is _int_mtx, i.e., “a Geometry/Sky Intersection Matrix from the ‘LB Incident Radiation’ component”, which “contains the relationship between each point of the analysed geometry and each patch of the sky.”. The second one is _sky_mtx, i.e., “a Sky Matrix from the ‘LB Cumulative Sky Matrix’ component, which describes the radiation coming from the various patches of the sky.’. In other words, the Ladybug Real Time Incident Radiation component takes the _int_mtx, which stores the whole sky matrix, and, for each sub-list of HOYs, it will query the irradiation values that relate to the specific time interval under consideration. To check if this workaround produces valid results, the author confronted the summation of multiple hourly irradiation values with the single irradiation value for a whole year, and the difference resulted in an insignificant value of 0.0026 kWh/m² for the whole year.

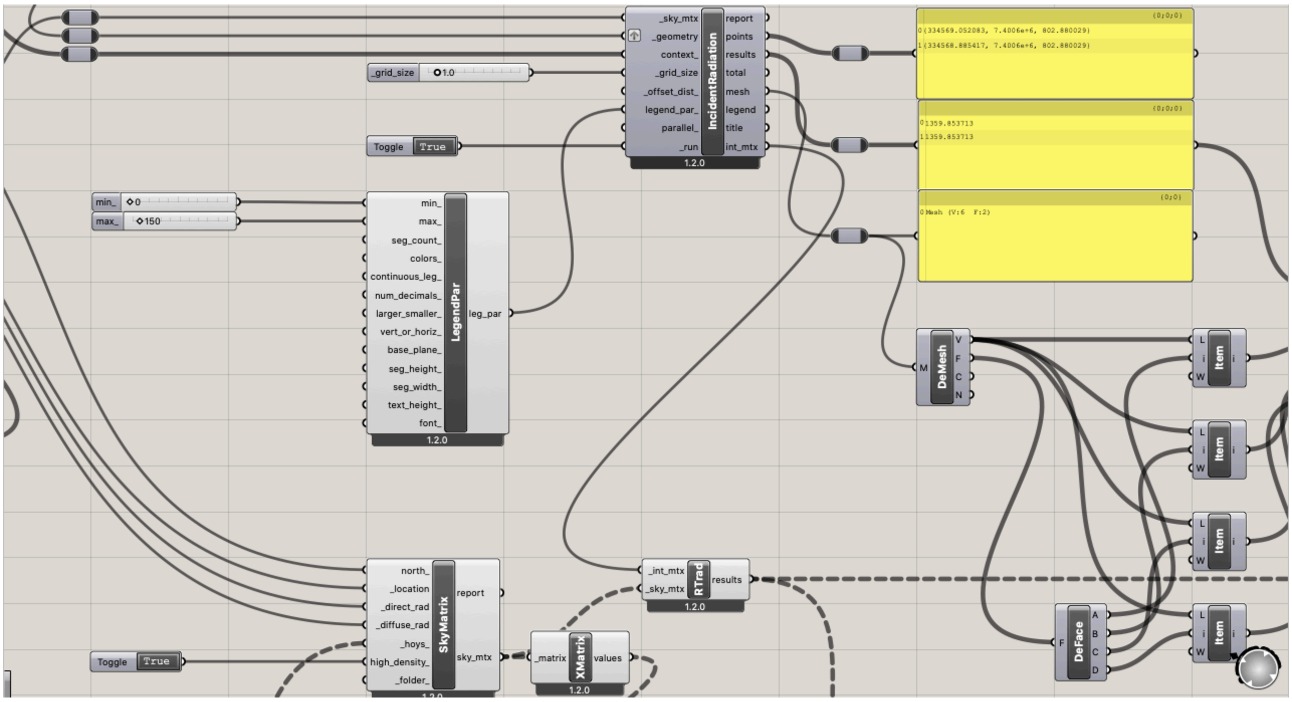


Figure 4-55 Ladybug - Real Time Solar Irradiance component.

The _int_mtx input comes from LB Incident Radiation, which stores the single sky matrix for the whole year; the _sky_mtx input, in turn, receives multiple (dashed line) sub sky matrices, each one corresponding to one time interval under analysis, i.e., at every hour. The output is a data tree structure that stores solar radiation (kWh) values per building surface.

With the geometrical and weather data inputs prepared, the Incident Radiation component of ladybug can be finally run. The grid size of the solar module is defined as 1.0m – the same spatial resolution for the raster products in GRASS and ArcGIS. By the end of this process, the algorithm computes one value of solar radiation (kWh) per triangle in the building thematic roof. These radiation values are summated per thematic roof and the result is divided by the area of the thematic roof, thus resulting in a single value of solar irradiation (kWh/m²) per roof.

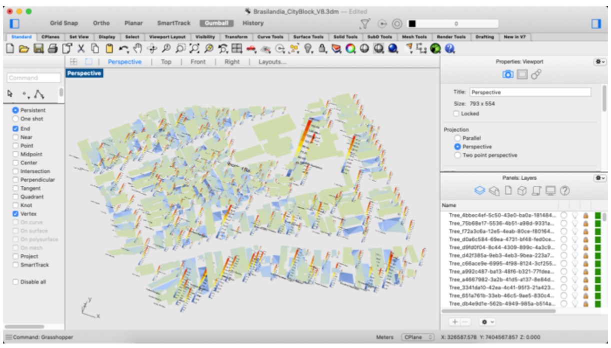


Figure 4-56 Ladybug – Solar radiation values – JUN.

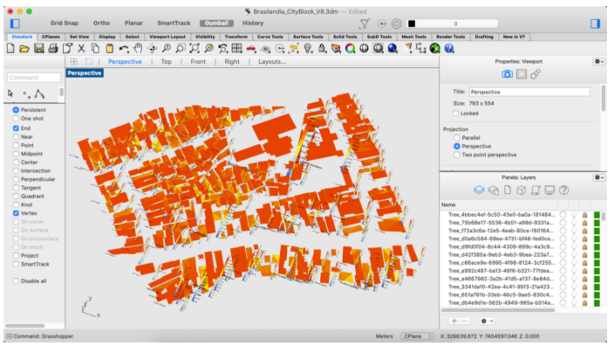


Figure 4-57 Ladybug – Solar radiation values – DEC.

With the results available, it is possible to respond the sub-research questions regarding Ladybug.

- **What are the minimum data requirements to run the solar irradiation module?**
  The Solar Incident Radiation component of ladybug demands Rhino boundary representations or Rhino meshes, which can be imported from CityGML via a DWG file. It also demands a Ladybug SkyMatrix component, which is derived from an EPW weather file.

- **How automatized is the solar irradiation pipeline? From a user perspective, how complex is setting up the simulation?**
  Among all software analysed, Ladybug is the least straightforward one since it requires a step-by-step implementation from the user side. This leads to an extra challenge of keeping track of the data tree in the workbench.

- **Is the solar irradiation module self-contained, or is it part of a greater package of solutions?**
  The ladybug Incident Radiation module is self-contained, thus delivering a total value of radiation (kWh) for all surfaces combined, or multiple irradiation (kWh/m²) values according to the number of faces in the resulting mesh.

- **What is the computational running time of the module if this metric is retrievable?**
  In the final simulation, each monthly computation took approximately five minutes for the 301 roofs in the scenario of Brasilândia, resulting in HH:MM:SS = 1:04:00.

- **What kind of urban features – among buildings, vegetation, and relief – does the module support, and how do these features participate in the simulation?**
  Strictly speaking, all sorts of urban features can participate in the simulation, as long as they are firstly converted into Rhino B-reps or Meshes. These urban features can participate either as context or as geometries for which the irradiation values will be calculated.

- **What is the degree of flexibility offered by the solar module in terms of selecting which features of the city model will be assessed and which will take part only as shading elements?**

  Ladybug is completely flexible in terms of selecting which features will be part of the context and which features will be assessed. The workbench environment of Grasshopper allows this sort of flexibility.

- **How is the output data delivered by the module in terms of time granularity and data model?**

  The Incident Radiation component of Ladybug has three main outputs: Results, which is “a list of numbers that aligns with the points. Each number indicates the cumulative incident radiation received by each of the points from the sky matrix in kWh/m2”; Total, which is “a number for the total incident solar energy falling on all input geometry in kWh”; and Mesh, which is “a coloured mesh of the geometry representing the cumulative incident radiation received by the input geometry”. In any case, the values of radiation and irradiation are related to the time interval provided in the list of HOYs, so the Lady Bug Real Time Incident Radiation Component must be used to produce multiple values of solar irradiation, one for each list of HOYs.

4.6 Virtual City Systems

Virtual City Systems is a Germany-based and private-held company that provides its clients with solutions on 3D City Models, 3D Geodata Infrastructure, Digital Urban Planning, Oblique Aerial Images and Urban Simulations. Among these Urban Simulations is the so-called *3D Solar Potential Analysis*[29] which, according to the company, “can produce monthly measurements of direct, diffuse, and global solar irradiation on the roofs and exterior walls of buildings, (…) even in the shade caused by vegetation, other buildings, and other objects on the premises in question.”

[29] The name ‘3D Solar Potential Analysis’ is derived from the company’s website.

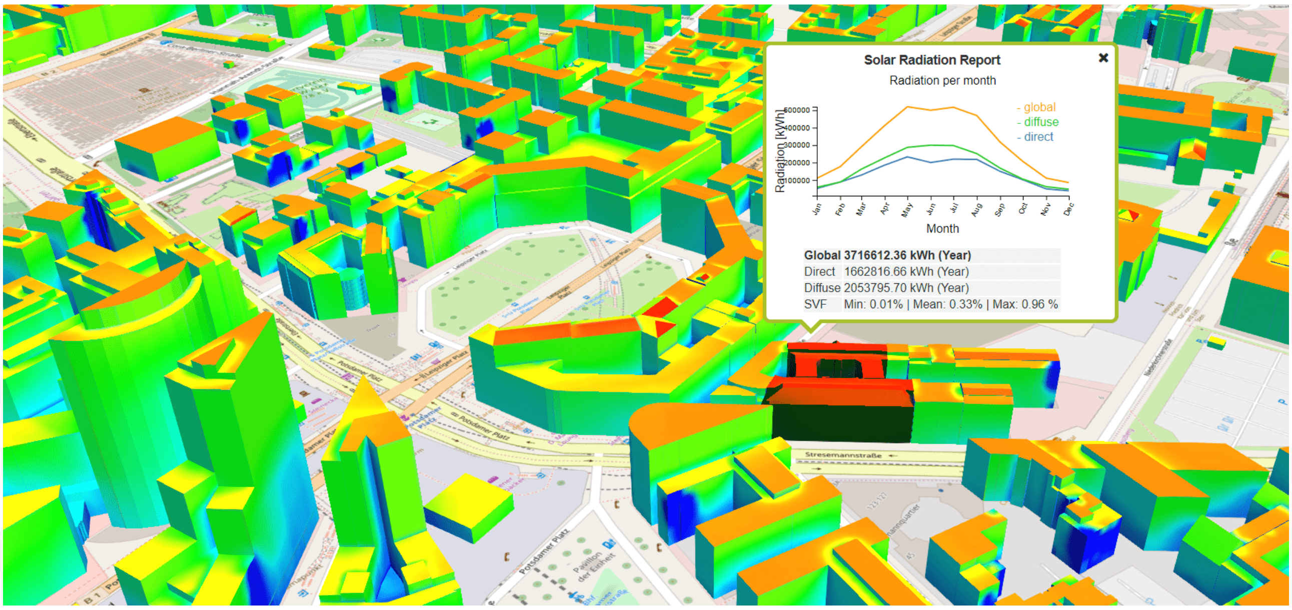


Figure 4-58 3D Virtual City Systems – 3D Solar Potential Analysis – example of data visualisation.

“Annual courses of all radiation types for selected object surfaces.”

Image and caption source: [Virtual City Systems](#).

Considering the commercial aspect of the simulation solutions, the experiments conducted with the 3D Solar Potential Analysis simulation tool were kindly carried out by the Virtual City Systems staff, who agreed to participate in the result comparison. The author was in direct contact with the team and provided them with two CityGML models, i.e., one for the scenario of Brasilândia and another one for Santana.

These two models, used in the simulations on Virtual City Systems’ tool, attend the same designing principles as the ones continuously used on all the other software, except for their extensions: while the version V8 of Santana and Brasilândia are reduced to their lowest common denominator, i.e., a model with an extension that can be handled by all software, the version V6 of these models, sent to Virtual City Systems, is the one with the extension of a whole tile of the original point cloud dataset of São Paulo. This reduction is necessary to maintain the best methodological consistency possible between the experiments that are conducted in the present research, but unfortunately the decision was not taken before the experiments on the Virtual City Systems’ tool had been already conducted. The following figures represent Brasilândia and Santana in their version V6.

Figure 4-59 Virtual City Systems – Brasilândia model V6.

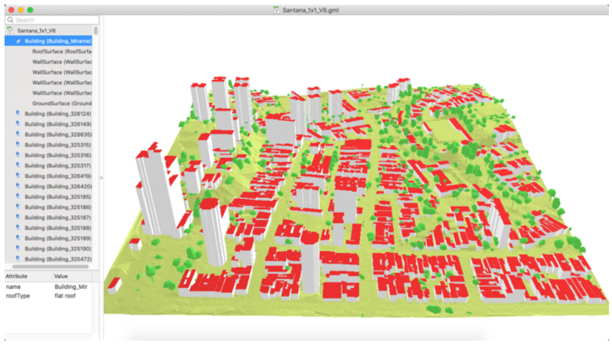


Figure 4-60 Virtual City Systems – Santana model V6.

The information hereby presented regarding the solar module and its operability are based on the literature, especially CHATURVEDI et al. (2017). In addition, the present research also analyses the output that is sent back from Virtual City Systems.

When it comes to data requirements, "(…) the only input data required is a 3D city model according to the CityGML standard in Level of Detail 2 (LoD2), having roof and wall surfaces represented as thematic surfaces." (CHATURVEDI et al., 2017, p.27) This is the case of the CityGML models that are sent to Virtual City Systems, in which not only LoD2 buildings but also trees and relief are modelled.

Regarding the weather data, in turn, the solar module "(…) is calibrated using freely available data from the NASA Atmospheric Science Data Centre. In an iterative process the model is adjusted to 22-year mean radiation parameters from the NASA surface meteorology and Solar Energy (SSE) mission (NASA, 2017), which can be queried by LAT/LON coordinates." (CHATURVEDI et al., 2017, p.27)

Once the geometrical and weather data are input, the simulation tool can be run to estimate direct, diffuse, and global radiations on building façades and roofs, whereas the "reflected radiation is neglected." (CHATURVEDI et al., 2017, p.27). The authors state that the direct solar radiation is computed adopting the 'transition model' of FU et. RICH, 1999 and by means of an algorithm that computes the sun position according to GRENA, 2012. When it comes to the diffuse radiation, an approximation of the sky dome is performed with points. At each point, the radiation is computed using the 'standard overcast model', as in Fu and Rich, 1999.

Sequentially, “a calculation basis on facades and roofs of the buildings of the city model is required. Therefore, a point grid is created on roof and wall surfaces, where each point represents the same fraction of the area of the surface it is attached to. These points are used as reference points for the irradiated solar energy.” (CHATURVEDI et al., 2017, p.27)

In addition, the authors also mention the urban features that take part in the simulation. Albeit the solar radiation results are only available for building façades and roofs, “the model considers the shadowing effects of buildings, vegetation objects, and optionally a digital terrain model (DTM) or DSM.” (CHATURVEDI et al., 2017, p.27) This is possible by means of a ray tracing approach. According to the authors, “for each point of the (…) point grid on the buildings, rays to all sun and hemisphere points are created and tested for intersection with surrounding buildings and the DTM/DSM using the ray / triangle intersection test according to (Möller and Trumbore, 2005).” (CHATURVEDI et al., 2017, p.27)

The resulting solar analysis is presented in a single, valid, and enriched CityGML file. For comparison, the input file Santana V6 is 76.2Mb large, and Brasilândia V6 is 76.9Mb large. Their corresponding outputs are 168.6Mb and 383.4 Mb large, respectively. When these outputs are loaded into a CityGML visualiser, such as FME Inspector, it is possible to observe that the relief and vegetation features are indeed missing, which already confirms what is described in the literature: LoD2 buildings take part in the simulation as both shading devices and as assessed features, whereas the vegetation objects and relief features act only as shading elements.

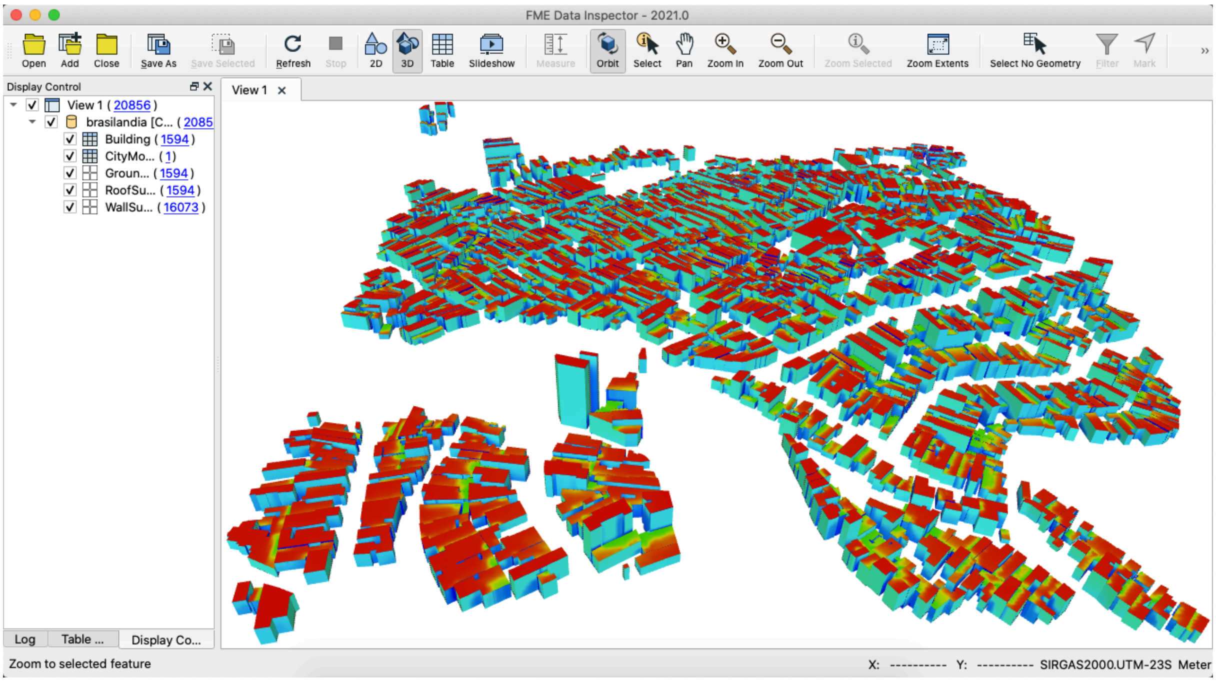


Figure 4-61 Virtual City Systems – Visualisation of some results for the scenario of Brasilândia.

Another interesting aspect to observe is the presence of different radiation values within the same building surfaces, both building walls and building roofs. This is possible since the 3D Solar Potential Analysis module creates the grid of points on the building's surfaces and roofs. It is worth mentioning that this spatial pattern is unique among all analysed software.

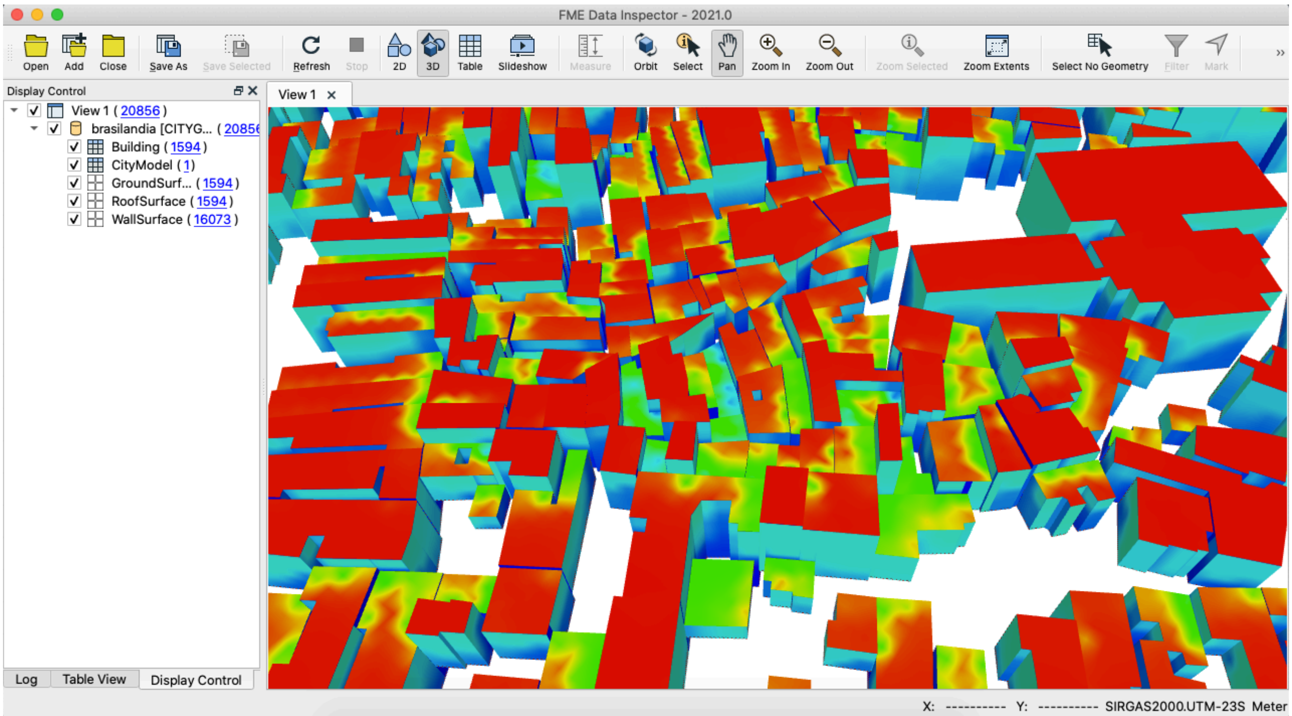


Figure 4-62 Virtual City Systems - Spatial pattern of solar analysis results, with interpolated values at building surfaces.

The zoomed area in the figure above is the one that, within the city block of Brasilândia, has the typical urban morphological characteristics of a favela, i.e., "dense and multi-layered structures built up around labyrinthine street networks." (SALAZAR MIRANDA et al., 2021, p.2). It is interesting to notice that, for this specific area, the spatial pattern of solar radiation results in high gradient values, whereas for the other portions of the city block this is not the case. Therefore, it is possible to affirm that the 3D Solar Potential Analysis simulation of Virtual City Systems offers the advantage of better understanding the spatial distribution of irradiation values in buildings of favelas, in which the shadowing effect from one building into another produces significatively different values within the same building roof.

Regarding the data structure of the results, the enriched CityGML file presents generic attributes of Sky View Factor (SVF), as well as Direct, Diffuse, and Global Solar Radiation (RAD) values both at building level and thematic surface level, in the case of walls and roofs. The SVF attributes describe the portion of the hemisphere that can be seen from a surface, ranging therefore from 0 to 1. The RAD attributes, instead, are radiation (energy) values in kWh, and thus neither irradiation in kWh/m^2 nor irradiance values in kW/m^2.

Therefore, the 3D Solar Potential Analysis tool has also the benefit of exporting diffuse, beam (direct) and global horizon radiance values, but only for monthly and yearly time intervals. These monthly radiation values are then divided by the area of each surface, so that they can be compared to the values coming from the other software.

With the results available, it is then possible to respond the sub-research questions regarding the 3D Solar Potential Analysis tool of Virtual City Systems:

- **What are the minimum data requirements to run the solar irradiation module?** Given the commercial perspective of the 3D Solar Potential Analysis tool, only the CityGML models are necessary for the experiments conducted in the present research. The weather data are handled directly by the company and, according to the literature, come from the NASA Atmospheric Science Data Centre.

- **How automatized is the solar irradiation pipeline? From a user perspective, how complex is setting up the simulation?**

As the simulations with Virtual City Systems are carried out directly by the company, there is no information available regarding this sub-research question.

- **Is the solar irradiation module self-contained, or is it part of a greater package of solutions?**
  As the simulations with Virtual City Systems are carried out directly by the company, there is no information available regarding this sub-research question.

- **What is the computational running time of the module if this metric is retrievable?**
  As the simulations with Virtual City Systems are carried out directly by the company, there is no information available regarding this sub-research question.

- **What kind of urban features – among buildings, vegetation, and relief – does the module support, and how do these features participate in the simulation?**
  Buildings are assessed and they also participate as shading surfaces for other buildings. According to the literature, vegetation objects and DTM/DSM features, if present, also participate but only as shading surfaces, and are therefore not assessed.

- **What is the degree of flexibility offered by the solar module in terms of selecting which features of the city model will be assessed and which will take part only as shading elements?**
  From what it is observed, there is no degree of flexibility, and only buildings are assessed. It is possible to imagine, however, that the solar radiation analysis based on the grid of points, which is currently implemented for building roofs and façades, could be also reproduced for other urban features, such as streets, for instance.

- **How is the output data delivered by the module in terms of time granularity and data model?**
  The output data is delivered in a valid and enriched CityGML file, with monthly records of diffuse, beam (direct) and global solar radiation (kWh).

4.7 Roof Mapping

The building footprints resulting from the 3D reconstruction algorithm are represented in the figure below (light pink fill) and confronted with the official building footprints from the cadastre (black edges and transparent fill).

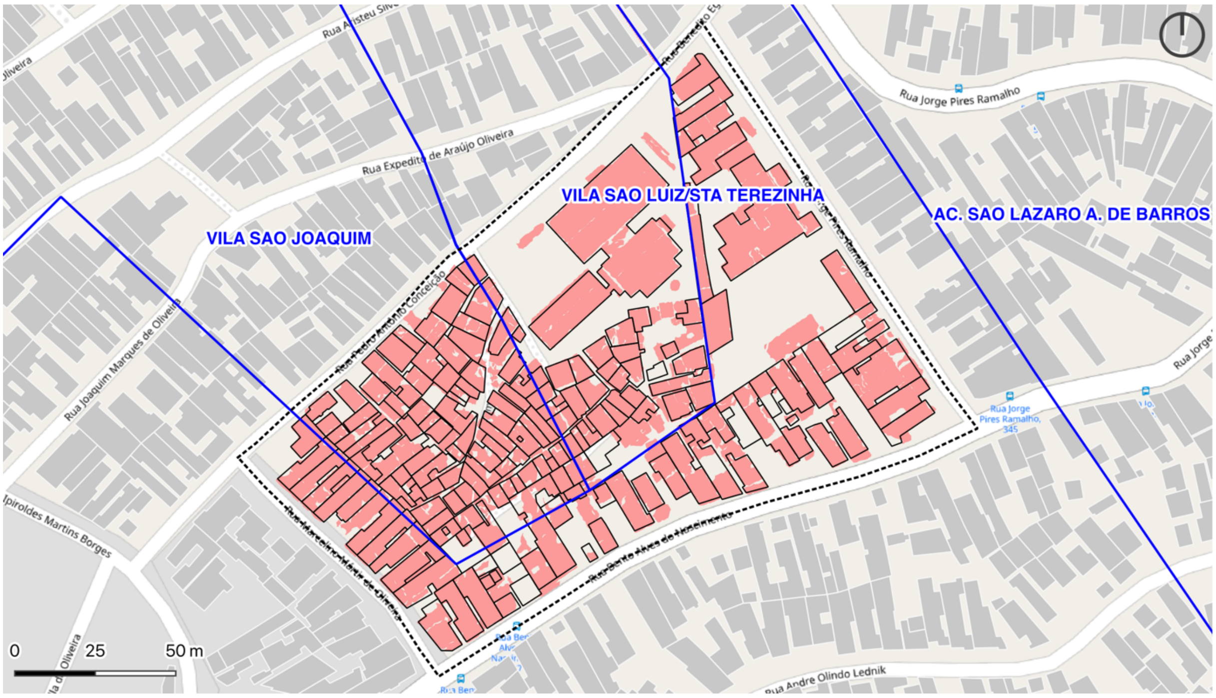


Figure 4-63 Roof Mapping - Reconstructed 2D footprints (light pink fill) vs. official cadastre footprints (black edges). The blue perimeters correspond to the irregular land parcelling informal settlements.

When it comes to the city block in its totality, it is possible to notice that the two datasets have much in common with respect to their covering area, i.e., the overlapping between these two. Considering only the footprints that lie within the city block, the official dataset covers a total built area of 11.629,41 m², whereas the reconstructed dataset covers 12.305,06 m², i.e., 5,81% more than the one from the cadastre. Such increase is expected since the cadastre dataset was released prior to the point cloud dataset. Therefore, recent buildings are still not mapped by the cadastre authority but are detected by the reconstruction algorithm. The overall overlapping area among these two datasets is 10.148,58 m², which corresponds to 87,27% of the built area from the official dataset and 82,5% from the reconstructed dataset.

However, when the number of footprints in each dataset is taken into consideration, a major difference is observed. The official dataset has 168 footprints within the city block, with a median area value of 49,9 m², an average area value of 69,2 m² and a standard deviation value of 111,5 m² - this last value due to the very large school building in the city block. The reconstructed dataset, in turn, has 525 features (3.13 times more than its counterpart), with a median area value of 9,6 m², an average area value of 23,4 m² and a standard deviation value of 33,5 m². The statistics are summarized in the table below:

	Official dataset	Reconstructed dataset
Total Built area	11.629,41 m²	12.305,06 m²
% Official	100 %	105,81 %
Overlapping area	10.148,58 m²	
% Overlapping area	87,27%	82,47%
Footprint polygons	168	525
Footprint area minimum	7,53 m²	0,55 m²
Footprint area median	49,85 m²	9,63 m²
Footprint area average	69,22 m²	23,44 m²
Footprint area maximum	1.307,90m²	312,78 m²
Area standard deviation	111,48 m²	33,51 m²

In practical terms, this difference means that the 3D building reconstruction algorithm marks as individual features:

- portions of these building roofs, which are reconstructed in true LoD2, and therefore slanted roofs become 'multi-part' buildings when projected into the 2D plane. This is mostly the case of houses with ceramic tile roofs outside the informal settlement perimeter.
- many small features that are not buildings per se, but part of these buildings, such as water tanks on the roofs. This is the typical case of dwellings lying in the informal settlement perimeter, in which the planar roofs are simple but contain extra elements.

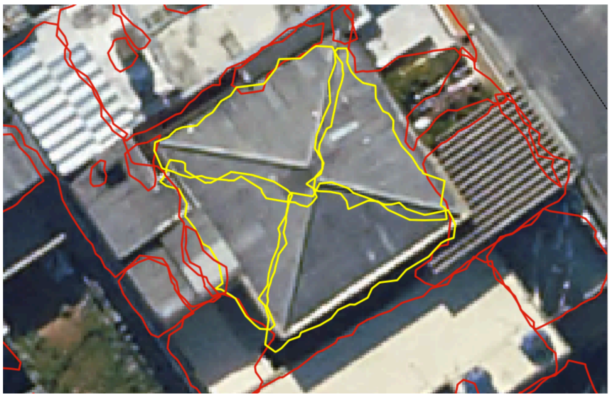


Figure 4-64 Roof Mapping – A slanted-roof building reconstructed with four features.

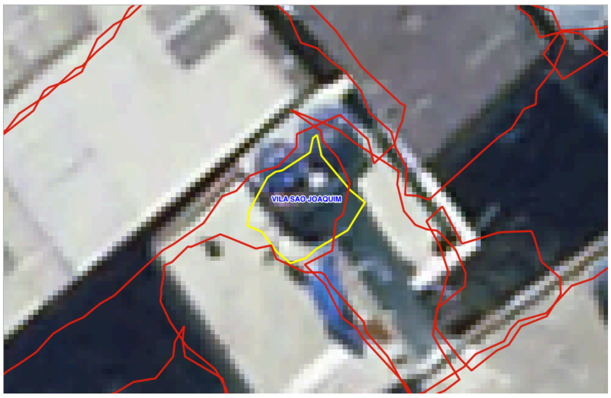


Figure 4-65 Roof Mapping – A small water tank reconstructed as an individual building.

Since such differences cannot be noted at city block scale, the research focuses the analysis on the buildings that are within the official perimeters of informal settlements, i.e., the area that has a typical urban morphology of a favela. The following figure highlights in pink the reconstructed footprints, whereas the official footprints in the surroundings are represented in light yellow.

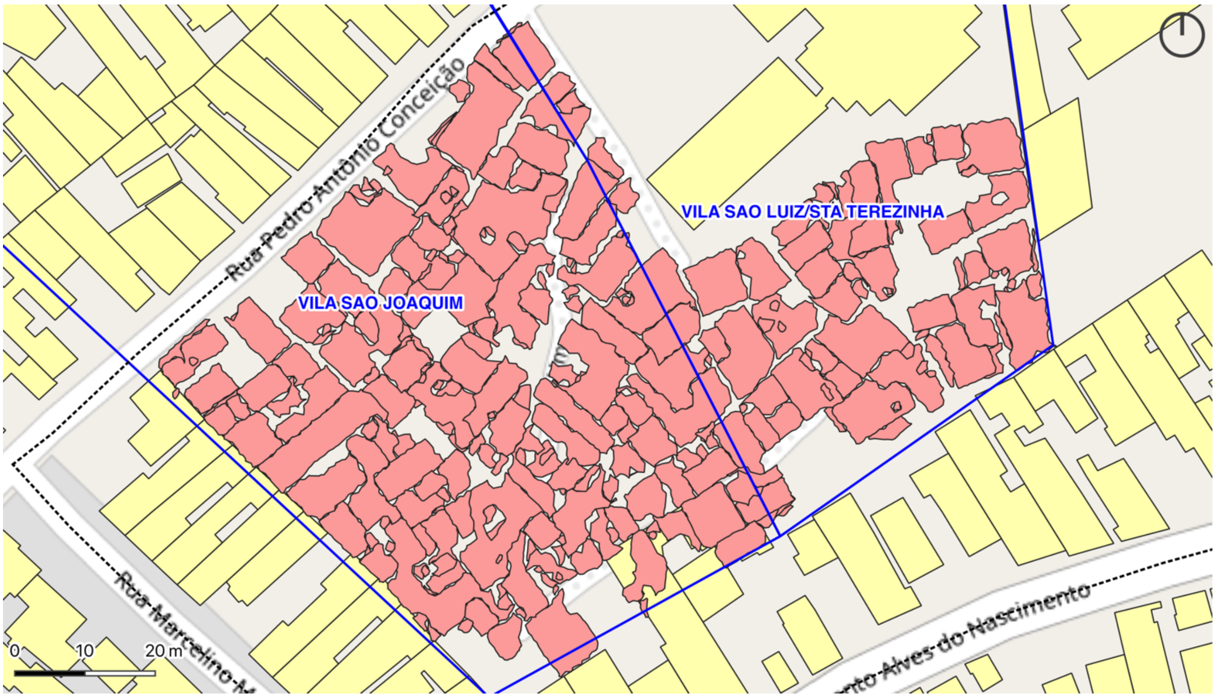
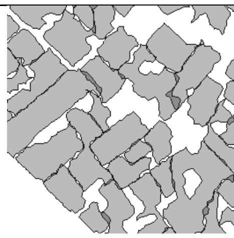
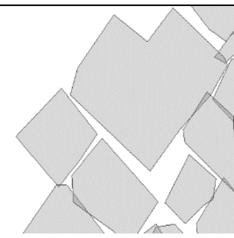
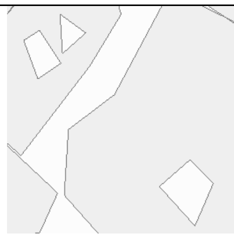
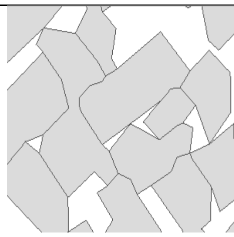
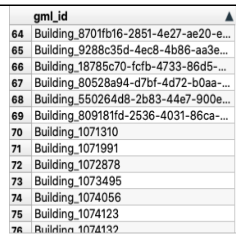


Figure 4-66 Roof Mapping - Reconstructed footprints within the informal settlement perimeter.

These reconstructed footprints have, in general, too many edges. They would result, therefore, in excessively complex LoD2 buildings in CityGML, which in turn have a negative impact in the computing efficiency of the solar irradiation modules. Moreover, as mentioned, many of these small features in the reconstructed dataset do not correspond to actual buildings, but only some structures on the roofs. Therefore, some experiments are conducted on an FME workbench to mitigate these undesired characteristics. The general steps are described below.

The reconstructed 3D features are spatially filtered according to the perimeters of the informal settlement, resulting in only 282 features. These are then clipped by the official footprints in the surroundings, so that they do not overlap.	
The resulting features are filtered by their area, and only features larger than 10 m² are preserved. This threshold is considered so that it eliminates water tanks but preserves actual buildings that may have a small footprint area. Only 122 features remain.	
A Douglas Peucker algorithm is implemented with a tolerance of 0.5m to generalise these buildings into simpler shapes. A higher tolerance results in triangular buildings, whereas a lower one still preserves many unnecessary edges.	
A Donut Hole Extractor removes the inner holes of these features. This simplification is important since many of these holes only exist due to the water tanks that are originally reconstructed.	
An Area Gap and Overlap Cleaner, with a tolerance of 0.5m, removes all intersections between features and reshapes these so to remove small gaps among them.	
As the previous operations modify the shapes of these features, another Clipping and further Area Gap and Overlap Cleaner is performed to re-adequate the reconstructed footprints to the boundaries of the official footprints in the surroundings.	

The features are once more filtered by area to eliminate small fragments that result from the geometrical operations. Unique UUID GML ids are given to the resulting features, and these are joint with the ones from the official dataset.



	gml_id
64	Building_8701fb16-2851-4e27-ae20-e...
65	Building_9288c35d-4ec8-4b86-aa3e...
66	Building_18785c70-fcfb-4733-86d5-...
67	Building_80528a94-d7bf-4d72-b0aa-...
68	Building_550264d8-2b83-44e7-900e...
69	Building_809181fd-2536-4031-86ca-...
70	Building_1071310
71	Building_1071991
72	Building_1072878
73	Building_1073495
74	Building_1074056
75	Building_1074123
76	Building_1074132

From the resulting post-processed footprint dataset, the workbench models the scenario in a vector-based data model just as it is the case for the previously presented scenarios of Brasilândia and Santana. The results are presented in Chapter 5.

RESULTS
& ANALYSIS

5 RESULTS & ANALYSIS

The present chapter aims to assess the overall results of the experiments conducted with the six solar analysis tools and to expose the reconstructed building footprints of the city block in Brasilândia. The analysis is hereby divided into three topics: firstly, a comparison among the solar modules is performed qualitatively, indicating their potentials and limitations, especially for favelas; sequentially, the resulting irradiation values are quantitatively confronted with the ground truth of Santana and correction factors are applied to the values of Brasilândia; finally, the roof mapping experiment is evaluated.

5.1 Solar Analysis – Qualitative comparison

In the previous chapter, the sub-search questions regarding each solar irradiation module are responded extensively. For comparison purposes, these are summarised in the table below and further detailed. It is important to emphasise that these qualitative results emerge solely from the user experience that the author has had during their MSc thesis.

	GRASS	ArcGIS	CitySim	SimStadt	Ladybug	VCS Solar
Minimum data requirements	Raster DSM	Raster DSM	CityGML + CLI + (Opt. HOR)	CityGML (Opt. TMY3)	B-rep/ Mesh + EPW	CityGML
Automatization / Complexity	Multiple steps	One-click basis	One-click basis	One-click basis	Multiple steps	Info. not available
Self-contained or package	Self-contained	Self-contained	Package of Solutions	Self-contained	Self-contained	Info. not available
Running time	HH:MM:SS 00:19:01	HH:MM:SS 04:43:41	BRA ±3 hr SAN ±7.5 hr	HH:MM:SS 00:08:06	HH:MM:SS 01:04:00	Info. not available
Assessable feature types	All features accepted	All features accepted	Buildings + Trees + Relief	Only buildings	All features accepted	Only buildings
Simulation flexibility	No flexibility	Minimal flexibility	No Flexibility	No Flexibility	Full flexibility	No Flexibility
Time granularity	Daily	Daily	Hourly	Hourly	Hourly	Monthly
Output data	Raster file	Raster file	TSV file	OUT file	Data tree	Enriched CityGML

With ‘Minimum data requirements’, the research aims to answer: ‘**What are the minimum data requirements to run the solar irradiation module?**’

> In the case of GRASS GIS, ArcGIS, SimStadt and Virtual City Systems, only the geometrical model is mandatory, i.e., a raster-based DSM for the first two and a CityGML file for the last two.
>
> For the context of a favela in which cadastral data is unavailable and the building reconstruction algorithm outputs unrealistic results, the point-cloud-derived DSM is already sufficient to offer at least an overview of solar irradiation within the territory of this favela. Therefore, visual interpretation of the results could indicate the optimal location of PV panels.
>
> These four abovementioned modules derive the weather/climate data from third-party databases, which result in a more straightforward workflow, despite the limited control of the input irradiation values. In this sense, SimStadt offers the optimal solution, allowing the user to either opt for the PVGIS or INSEL databases, or to input their own TMY3 fie. CitySim and Ladybug, in turn, demand respectively a climate file and an EPW file, which result in a weather data pre-processing step.

With ‘Automatization / Complexity, the research aims to answer: ‘**How automatized is the solar irradiation pipeline? From a user perspective, how complex is setting up the simulation?**’

> ArcGIS, CitySim and SimStadt run the solar simulations on a one-click basis, which means that, once all input datasets are loaded into the software, the solar irradiation analysis is performed automatically for the whole year. GRASS GIS and Ladybug, in turn, rely on a multiple-step approach. The former is optimized if aspect, slope, and horizon maps are computed by the user beforehand, whereas the latter demands a great number of intermediate steps in the workbench, which in turn demand solid knowledge about data structure. As the simulations with Virtual City Systems are carried out directly by the company, there is no information available regarding this sub-research question.

With ‘Self-contained or package’, the research aims to respond: ‘**Is the solar irradiation module self-contained, or is it part of a greater package of solutions?**’

Except for CitySim, all solar irradiation modules are self-contained, which means that they can be triggered individually to return only solar radiation / irradiation / irradiance values for the selected time interval. In CitySim, this is not possible since the software will run the whole collection of algorithms, even if the user is not interested in other variables, such as building heat demand. As the simulations with Virtual City Systems are carried out directly by the company, there is no information available regarding this sub-research question.

With ‘Running time’, the research aims to respond: ‘**What is the computational running time of the module if this metric is retrievable?**’

GRASS GIS, SimStadt and Ladybug return solar radiation / irradiation / irradiance values for all 301 buildings in the scenario of Brasilândia in less than or almost an hour.

A running time of approximately one hour is a good metric, considering the complexity of the built environment of a favela, in which gathered and compact dwellings result in a great number of building surfaces. Moreover, in slopy informal settlements – which is often the case – the terrain model cannot be highly generalised, thus resulting in complex relief components for the vector-based approach.

ArcGIS has a longer running time since it is not possible to provide the software with aspect / slope / horizon maps beforehand, and the loop statement recomputes these for each day of the year. It would be possible to run all days of the year in a single command – without a loop – but the computer of the author crashed for uncleared reasons. Further research could evaluate this possibility.

As the simulations with Virtual City Systems are carried out directly by the company, there is no information available regarding this sub-research question.

With ‘Assessable feature types’, the research aims to respond: ‘**What kind of urban features – among buildings, vegetation, and relief – does the module support, and how do these features participate in the simulation?**’

As the raster-based modules (GRASS GIS and ArcGIS) require a Digital Surface Model as their geometrical input, all urban features are assessed, if they are present in the DSM in the first place.

CitySim computes solar irradiance values for all building, tree and relief surfaces present in the model.

SimStadt has the disadvantage of only inputting building surfaces, and the impact of other urban features is neglected. This is particularly negative for the solar analysis in building of favelas, which are often built in slopy areas of the city, like the ones in the hills Rio de Janeiro but also as many in the case of São Paulo. Therefore, neglecting the terrain results in inaccurate values for favelas.

With Ladybug, every single geometrical B-rep or Mesh that is input in the ‘geometry’ port will be assessed, which makes this module the most powerful one under this criterium.

In the case of Virtual City Systems, results are delivered regarding buildings exclusively, but the simulation also considers the vegetation and relief when these are present.

With ‘Simulation flexibility’, the research aims to respond: ‘**What is the degree of flexibility offered by the solar module in terms of selecting which features of the city model will be assessed and which will take part only as shading elements?**’

Under this criterium, all but Ladybug have a negative remark.

In GRASS GIS, all pixels of the input DSM are assessed in the simulations that the author performs. An experiment could be conducted to verify if a masking layer would handle such degree of flexibility, but this procedure is out of the scope of the present research. In ArcGIS, the user could assess a list of points instead of the whole extent of the DSM, but this might result in a long list of points corresponding to features that are subject of investigation. In CitySim and SimStadt, all input geometries will be assessed, and such inflexibility results in unnecessary running time due to features which only matter as shading elements.

Ladybug, in turn, offers full flexibility, since not only it is possible to distinguish which feature types will be assessed and which will act just as shading surfaces, but also which single geometries will be assessed, and which will not. Given this flexibility, for the context of a favela, the user can already establish pre-defined thresholds, such as the minimum roof area that allows the installation of PV panels.

Regarding Virtual City Systems, from what is possible to observe in the output CityGML file, only and all buildings are assessed, and the remaining features are excluded from the results.

With 'Time granularity' and 'Output data', the research aims to respond: '**How is the output data delivered by the module in terms of time granularity and data model?**'

With GRASS GIS, ArcGIS, CitySim, SimStadt and Ladybug, solar irradiation values can be derived in at least a daily basis: from a raster file, in the case of the first two; from a spreadsheet file, in the case of the following two; and from a data tree, in the case of the last one, which needs to be further processed. Daily values are ideal to observe a trend within a year interval.

The results from Virtual City Systems are output only on a monthly basis.

5.2 Solar Analysis – Quantitative comparison

In parallel with the qualitative comparison, this MSc thesis also aims to respond, for each solar module under analysis, the following sub-research question: 'With respect to the meteorological station of São Paulo, which offers ground truth data, what is the level of accuracy of the module in terms of solar irradiation?'

The quantitative comparison is therefore structured in two sub-parts. The first one regards the scenario of Santana, in which the results coming from each simulation software are confronted to the values present in the ground truth dataset of the meteorological station. The second one regards the scenario of Brasilândia, here understood as a *terra incognita* also from a weather data perspective. In this case, a comparison is not possible by means of ground truth data, but only among the results coming from each software, which are corrected based on the first comparison made with the data for Santana. The following pages detail these analyses.

Santana

From the two raster-based output datasets regarding the scenario of Santana – namely the ones coming from GRASS and ArcGIS –, daily solar irradiation values are extracted based on the coordinates of the meteorological station of Mirante de Santana, and further aggregated into monthly and yearly values.

In the case of the four vector-based modules, only the values regarding the roof of the 'artificial' building Mirante de Santana are extracted, and these are further aggregated into monthly and yearly values.

The results from each simulation are plot in the following daily, monthly, and yearly charts, which are also confronted with respect to the ground truth data coming from the INMET station.

Regarding the yearly chart, the input values present in the CLI file for CitySim, in the TMY3 file for SimStadt and in the EPW file for Ladybug are also represented, so that the comparison between input and output values for each one of these software can also be made. The daily curve representing the 3D Solar Potential Analysis module of Virtual City Systems is dashed since these are always constant within a month.

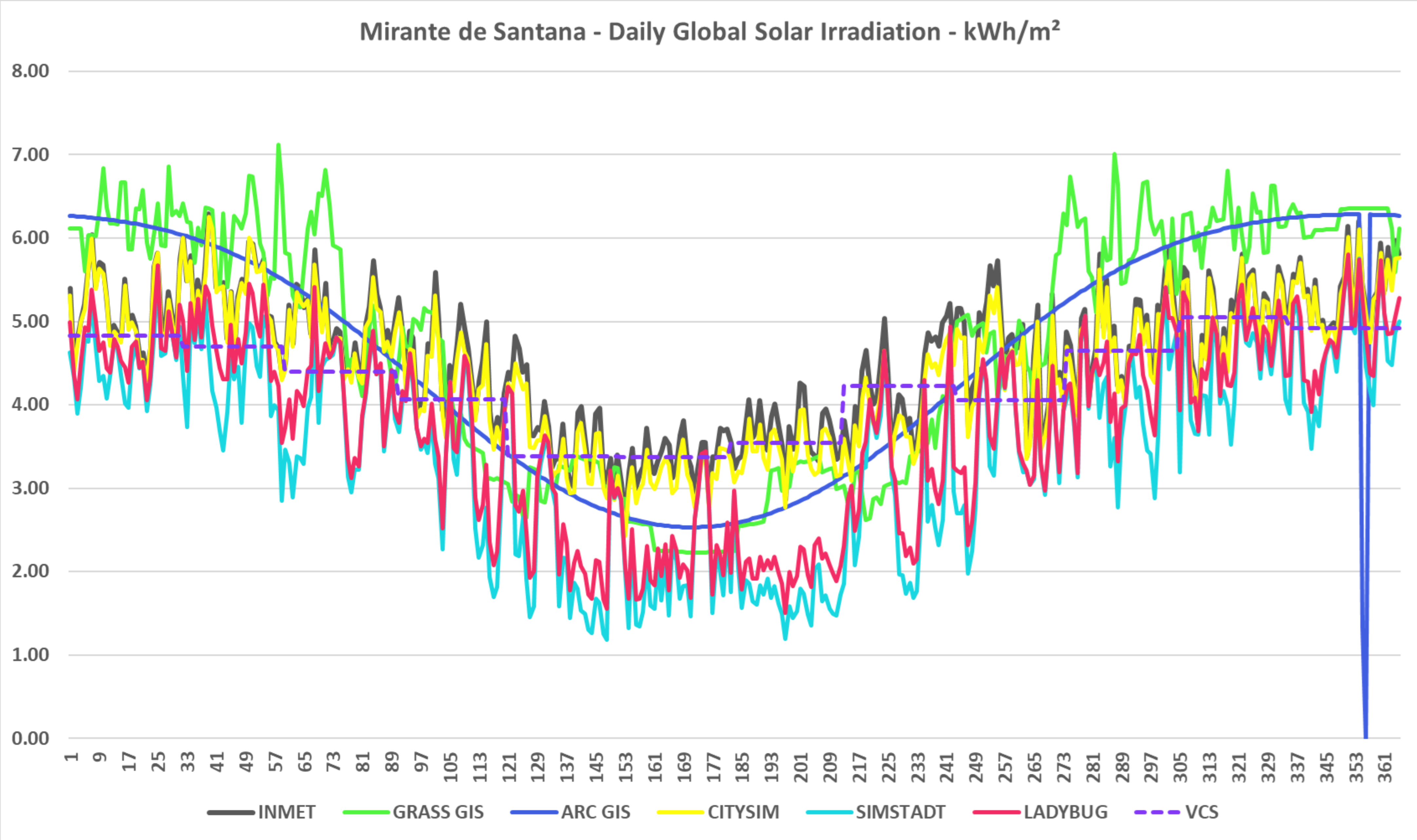


Figure 5-1 Mirante de Santana – Daily Global Solar Irradiation – kWh/m²

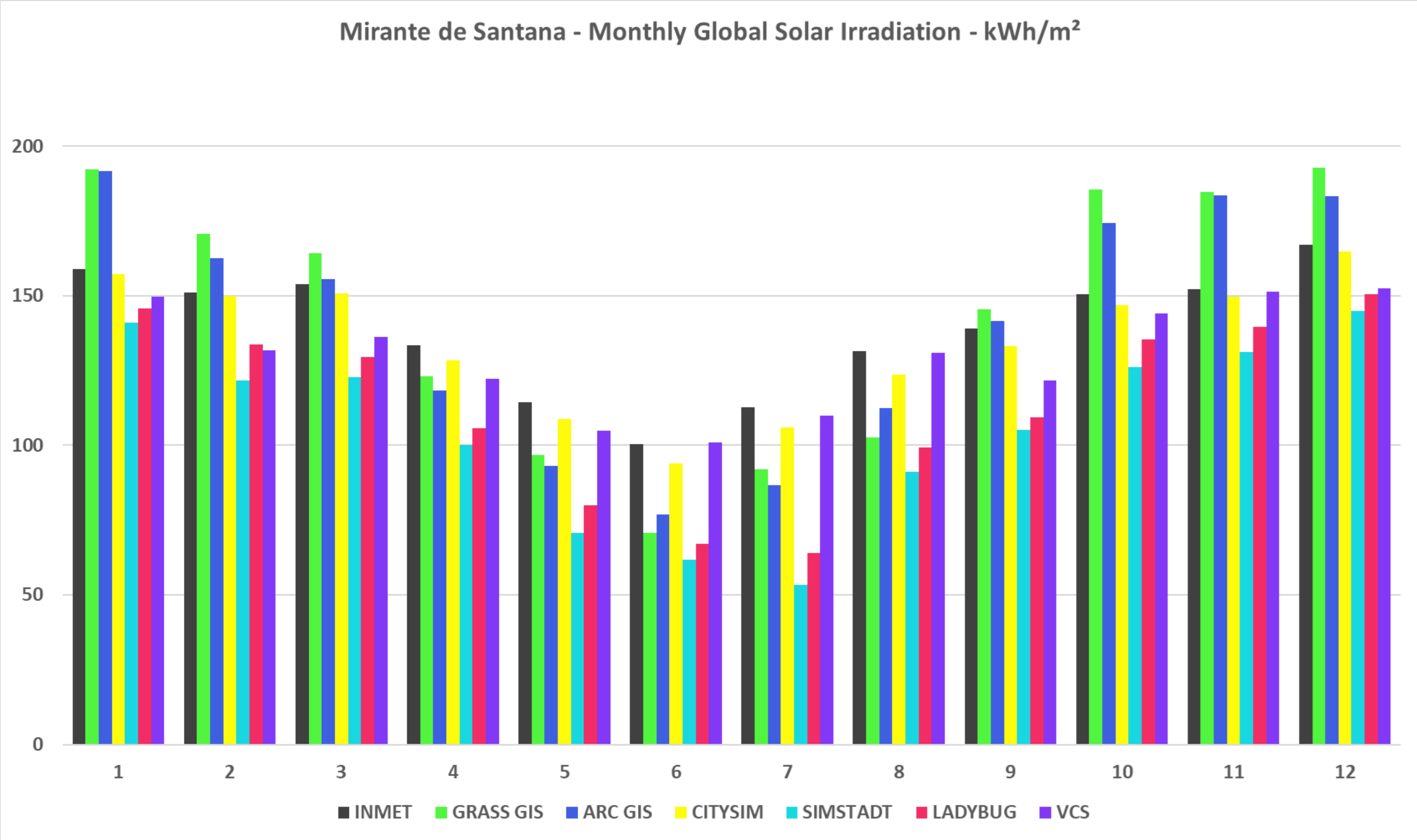


Figure 5-2 Mirante de Santana – Monthly Global Solar Irradiation – kWh/m²

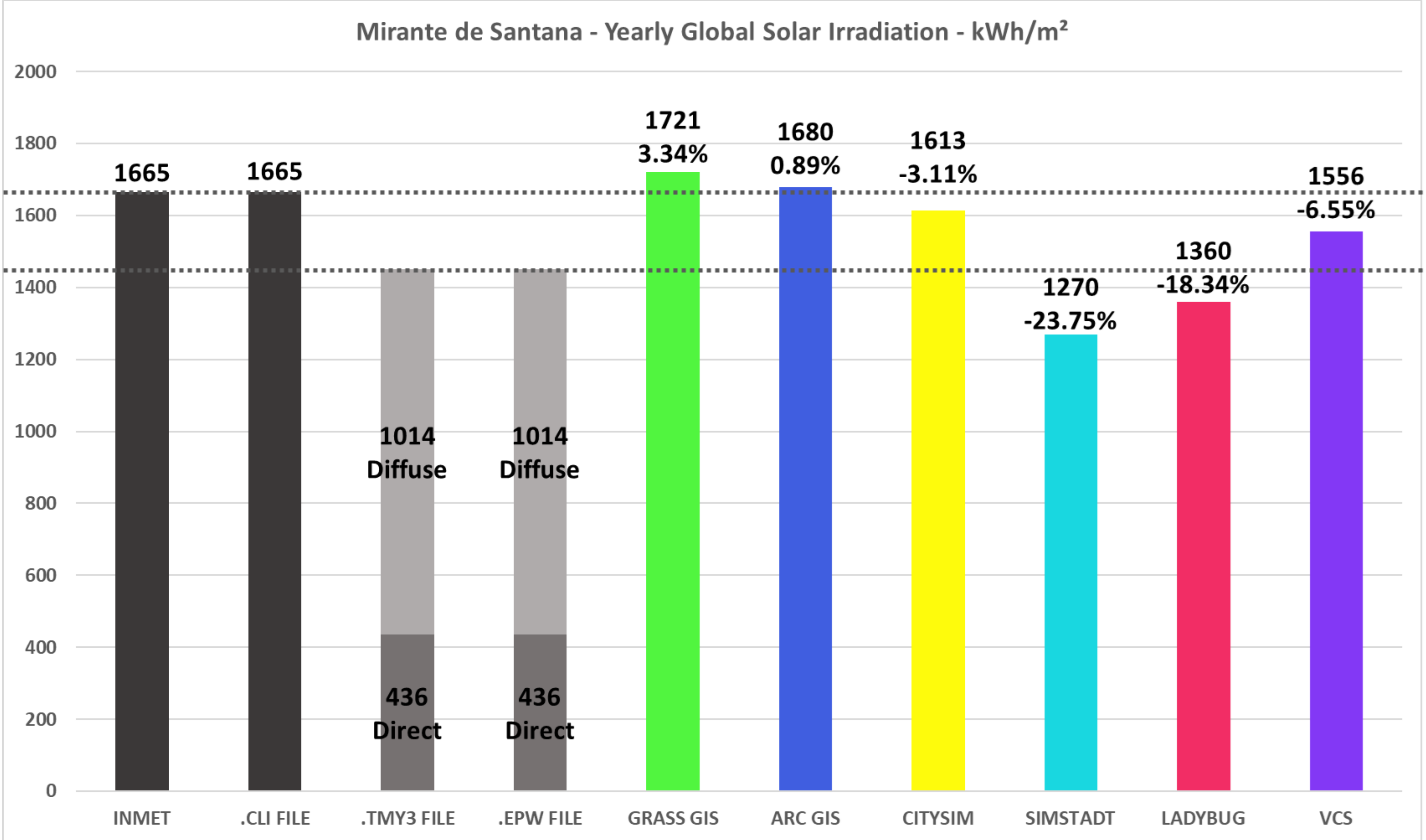


Figure 5-3 Mirante de Santana – Yearly Global Solar Irradiation – kWh/m²

The percentual values correspond to the differences of annual summation of global irradiation between each solar module and the ground truth dataset.

Complementary, for the purpose of assessing not only the amount of solar irradiation per time interval and per software, but also the level of accuracy of each software with respect to ground truth, root-square mean error (RMSE) values are calculated on monthly and yearly basis for each software, according to the formula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n-1}}$$

where:

n – number of records, e.g., 365 for the annual analysis.

i – variable solar irradiation (in kWh/m²)

xi – software value (GRASS, ArcGIS, CitySim, SimStadt, Ladybug, VCS)

$\hat{x}$i – ground truth value (INMET)

The resulting RMSE values (in kWh/m²) are presented in the table below. Values closer to zero indicate a greater coincidence between the software data series and the ground truth one, whereas larger values indicate a greater discrepancy between these two datasets. Exclusively for visualisation purposes, a continuous colour ramp represents lower to higher values from green to red colour hues.

	GRASS GIS	ARC GIS	CITYSIM	SIMSTADT	LADYBUG	VCS
01 – JAN	1.19	1.15	0.07	0.72	0.52	0.54
02 – FEB	0.92	0.60	0.05	1.20	0.70	0.85
03 – MAR	0.83	0.46	0.11	1.18	0.90	0.70
04 – APR	0.85	0.71	0.18	1.35	1.11	0.65
05 – MAY	0.87	0.80	0.20	1.68	1.32	0.60
06 – JUN	1.08	0.85	0.22	1.46	1.26	0.29
07 – JUL	0.80	0.92	0.23	2.00	1.63	0.32
08 – AUG	1.08	0.76	0.26	1.53	1.21	0.55
09 – SEP	0.86	0.82	0.23	1.39	1.19	0.88
10 – OCT	1.36	0.91	0.13	0.98	0.60	0.55
11 – NOV	1.24	1.17	0.09	0.84	0.50	0.48
12 – DEC	0.94	1.54	0.08	0.86	0.63	0.62

	GRASS GIS	ARC GIS	CITYSIM	SIMSTADT	LADYBUG	VCS
YEAR	1.00	0.92	0.17	1.30	1.02	0.60

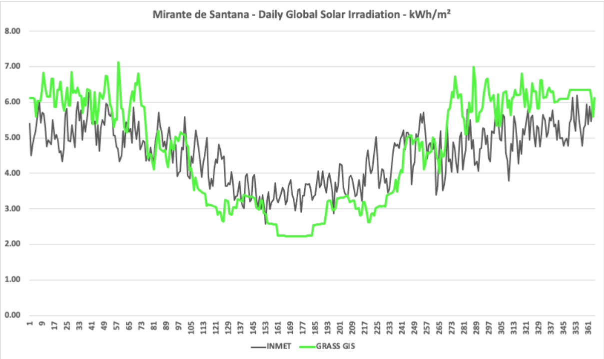


Figure 5-4 GRASS GIS – computed daily global irradiation at Mirante de Santana.

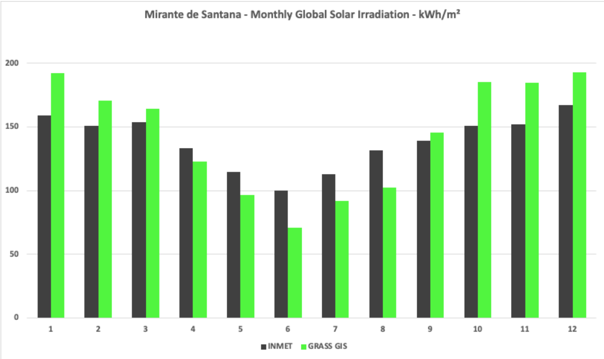


Figure 5-5 GRASS GIS – computed monthly global irradiation at Mirante de Santana.

GRASS GIS (green curve) presents a spiky daily curve. In a monthly analysis, it is possible to observe that it overestimates irradiation values for the spring/summer period (September to March), whereas during autumn/wintertime it underestimates these values. On a yearly basis, the global solar irradiation value is **3.34% higher than ground truth**. The **monthly RMSE values range in between 0.80 and 1.36 kWh/m²**, whereas the **annual RMSE results in 1.0 kWh/m²**.

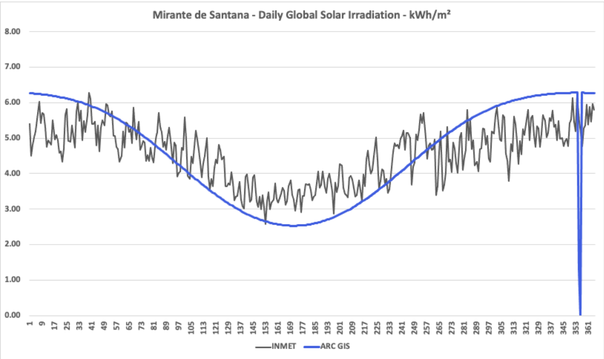


Figure 5-6 ArcGIS – computed daily global irradiation at Mirante de Santana.

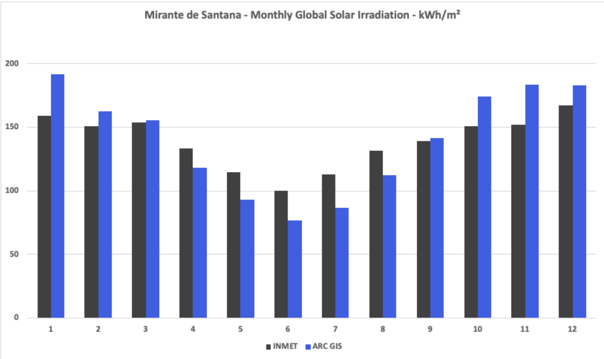


Figure 5-7 ArcGIS – computed monthly global irradiation at Mirante de Santana.

ArcGIS (blue curve), in turn, presents a more regular daily bell curve, which indicates that the Area Solar Radiation module reads irradiance values from a mathematical function according to the latitude of the site. The only noticeable outliers are the values of December 21st and 22nd, which are set as 1.33 and 0.00 kWh/m² in all simulations run throughout this research. This might be a bug in the software for this specific location. In a monthly analysis, ArcGIS also overestimates irradiation values for the spring/summer period (September to March) and underestimates them for the autumn and winter seasons. On a yearly basis, the global solar irradiation value is **only 0.89% higher than ground truth**, making ArcGIS the software that delivers the closest amount of energy with respect to ground truth. The **monthly RMSE values range between 0.46 and 1.54 kWh/m²**, with a greater amplitude than the ones from GRAS GIS. The **annual RMSE results in 0.92 kWh/m²**.

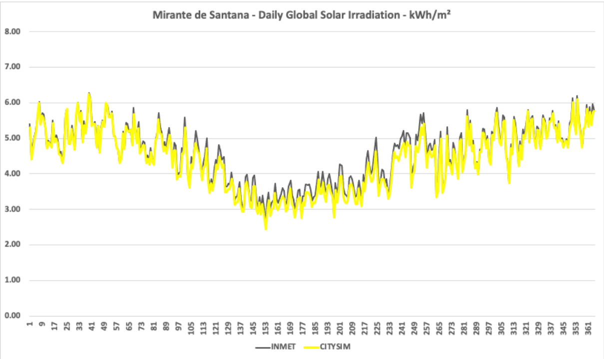


Figure 5-8 CitySim – computed daily global irradiation at Mirante de Santana.

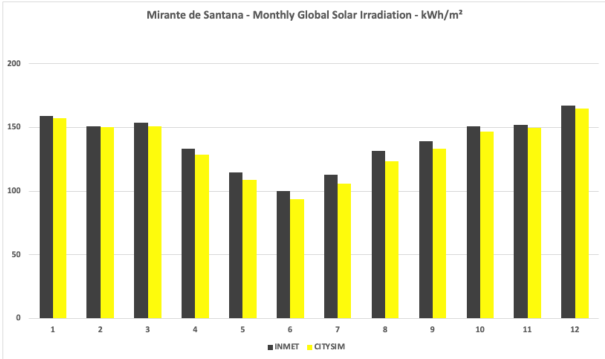


Figure 5-9 CitySim – computed monthly global irradiation at Mirante de Santana.

CitySim (yellow curve) presents a spiky curve that is very-well aligned with the one coming from the meteorological station. The climate file, mandatory for CitySim, is designed with the same dataset as the one containing ground truth values. Moreover, since the climate file stores only global irradiance values instead of the extra two derived direct and diffuse irradiance attributes, the error propagation resulting from the previous workarounds are avoided. Nevertheless, slight underestimations of global solar irradiation values are observed throughout the year, more expressively from April to September. On a yearly basis, the global solar irradiation value is **3.11 % lower than ground truth**, therefore a larger discrepancy than the one resulting from the ArcGIS values. However, CitySim results the lowest **RMSE values, ranging from 0.05 to 0.26 kWh/m² on a monthly basis, and 0.17 kWh/m² on an annual basis.** This makes the CitySim the software containing the solar module with most accurate results.

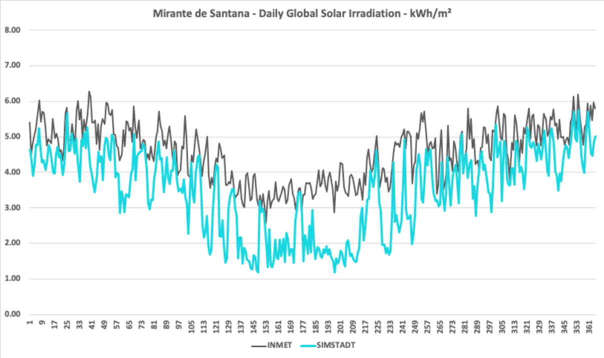


Figure 5-10 SimStadt – computed daily global irradiation at Mirante de Santana.

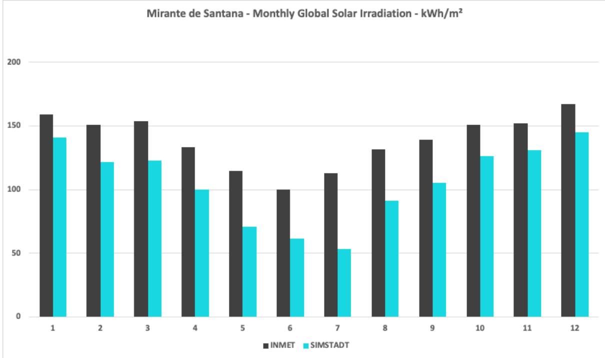


Figure 5-11 SimStadt –– computed monthly global irradiation at Mirante de Santana.

SimStadt (cyan curve) – in which the simulation is based on the global, direct, and diffuse irradiation values in the designed hourly TMY3 file – presents an intensely spiky daily curve. This curve mostly follows the ground truth trend in an offset – albeit with much lower values –, but at some points it also deviates from ground truth data significantly. From the GUI and documentation of SimStadt, it is possible to observe that the global irradiance curve is input, but it is not possible to understand how it influences the output curve. Nevertheless, it becomes clear that, in a monthly analysis, the underestimation of global solar irradiation values is constant throughout the year. This might indicate that the direct and diffuse irradiance curves play a much more significant role in the RSA algorithm, but further investigation would be needed for conclusions. On a yearly basis, the global solar irradiation value is **23.75% lower than ground truth**, but the discrepancy is lower when compared to the summation of direct and diffuse irradiation in the TMY3 file, only 12.44% lower. Still, taking into consideration the ground truth dataset, **the monthly RMSE values range from 0.72 to 2.00 kWh/m²**, and therefore this is the dataset that results in the largest RMSE values. **The annual RMSE results in 1.30 kWh/m²**, also the largest one with respect to the results coming from other software.

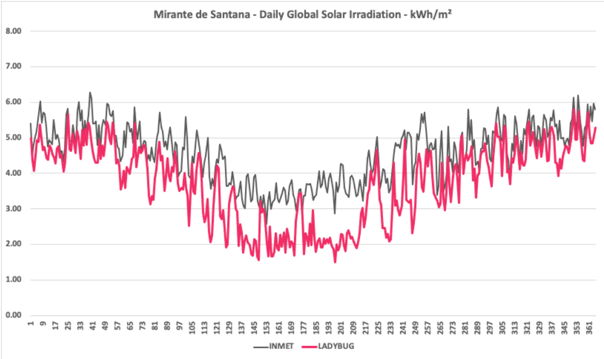


Figure 5-12 Ladybug -– computed daily global irradiation at Mirante de Santana.

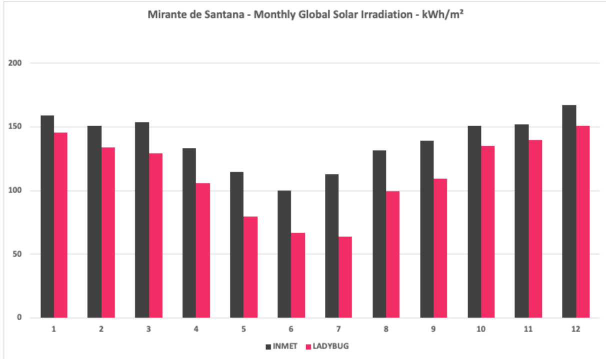


Figure 5-13 Ladybug – computed monthly global irradiation at Mirante de Santana.

Ladybug (pink curve), – in which the simulation is based exclusively on the direct and diffuse irradiation values in the designed hourly EPW file – presents an intensely spiky daily curve, parallel to the one coming from SimStadt but with slightly greater values. In a monthly analysis, the underestimation of global solar irradiation values is also constant throughout the year. This is coherent since, according to the assumptions carried out when designing the EPW file, the summation of direct and diffuse irradiation does not reach the summation of global irradiation. Unlike SimStadt, it is clear from the beginning that the results coming from Ladybug will not be based on the amount of global irradiation present in the ground truth dataset, but rather partial irradiation values artificially derived from the INMET dataset. Still, on a yearly basis, the global solar irradiation value is **18.34% lower than ground truth**, which represents a much lower discrepancy as the one resulting from SimStadt. Additionally, the discrepancy is even lower when compared to the summation of direct and diffuse irradiation in the EPW file, only 6.23% lower. Considering the ground truth dataset, **the monthly RMSE values range from 0.50 to 1.63 kWh/m²**. **The annual RMSE results in 1.02 kWh/m²**, a similar value to the one coming from GRASS GIS.

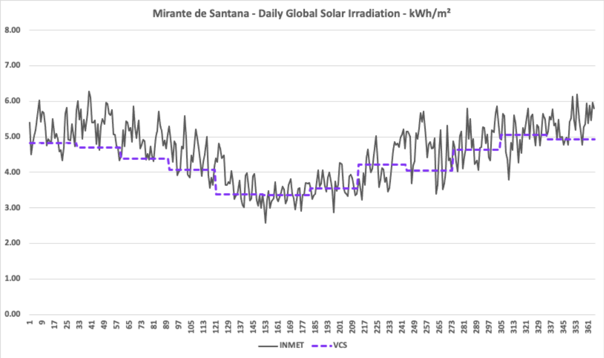


Figure 5-14 VCS – derived daily global irradiation at Mirante de Santana.

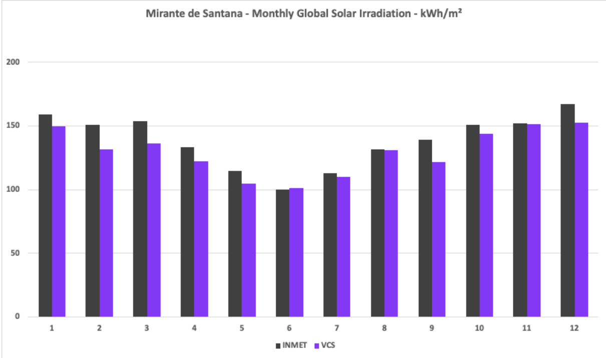


Figure 5-15 VCS – computed monthly global irradiation at Mirante de Santana.

VCS Solar (purple curve) – in which the simulation is based on the irradiation values that the company queried from third-part weather databases – does not present daily values, and therefore the VCS Solar values are dashed in the daily graph. In a monthly analysis, the values are very well aligned with those from ground truth, with larger underestimations of global solar irradiation only in the months of February and September. On a yearly basis, the global solar irradiation value is **6.55% lower than ground truth**. The **monthly RMSE values range in between 0.29 and 0.88**, whereas the **annual RMSE results in 0.60**, which is the second-lowest annual RMSE value.

Therefore, based on the comparison among the annual amount of solar energy coming from each software, it is possible to affirm that ArcGIS outputs an annual summation of irradiation values that is the closest to the one of ground truth values, the former only 0.89% greater than the latter.

Nevertheless, from an round square mean error perspective, CitySim is the software that outputs a curve that has the best correspondence to the one of ground truth. The monthly RMSE values are much lower than the RMSE values coming from all the other software, and so is the case with the annual RMSE value of CitySim. As already indicated, this relates to the fact that CitySim is the only software that is based exclusively on global irradiance values coming from the INMET weather data.

GRASS GIS presents an annual summation of solar irradiation that is, in modulus, almost as distant as the one coming from CitySim, respectively 3.34% above and 3.11% below ground truth. Nevertheless, the monthly and yearly RMSE values coming from CitySim are still lower compared to the ones of GRASS GIS, and therefore the first is still preferred.

The 3D Solar Potential Analysis tool from Virtual City Systems, in turn, offers relatively low monthly and yearly RMSE values – but not as low as the ones coming from CitySim. Nevertheless, the yearly summation of global solar irradiation is 6.55% underestimated, and therefore further away from ground truth with respect to GRASS GIS, ArcGIS and CitySim.

Lastly, SimStadt and Ladybug present much lower summations of global irradiation values, and this has to do with the fact that the summation of direct and diffuse values in the TMY3 and EPW files, respectively, do not reach the global irradiation values present in the INMET dataset. Their monthly and annual RMSE values are also the highest among all software.

Brasilândia

When it comes to Brasilândia, there is no ground truth dataset for comparison, and so this site becomes a *terra incognita* also from a weather data perspective. Therefore, an adaption must be performed such that the datasets coming from the software can be compared to each other.

In this sense, based on the discrepancies between monthly ground truth values and monthly global solar irradiation values that each software computes for the meteorological station of Mirante de Santana, a sequence of monthly correction factors are applied to the values extracted for the buildings in Brasilândia.

$$F_{s,i} = \frac{\hat{x}_i}{x_{s,i}}$$

F – monthly correction factor

s – software (GRASS GIS, ArcGIS, CitySim, …)

i – month of the year [1,12]

$\hat{x}$ – measured irradiation value as ground truth

x – computed irradiation value

With this procedure, the research mitigates the intrinsic weather data heterogeneity among results coming from different solar modules, and therefore it only evaluates the differences that result from the implementation of the algorithms.

The following chart presents the collection of monthly corrections factors that are applied to the global solar irradiation values encountered for the buildings in Brasilândia. If a correction factor is greater than one, the computed value for the meteorological station of Santana is considered underestimated with respect to ground truth, and therefore the 301 values for the building roofs in Brasilândia are increased; likewise, if a correction factor is lower than one, the computed value for the meteorological station of Santana is considered overestimated with respect to ground truth, and therefore the 301 values for the building roofs in Brasilândia are decreased.

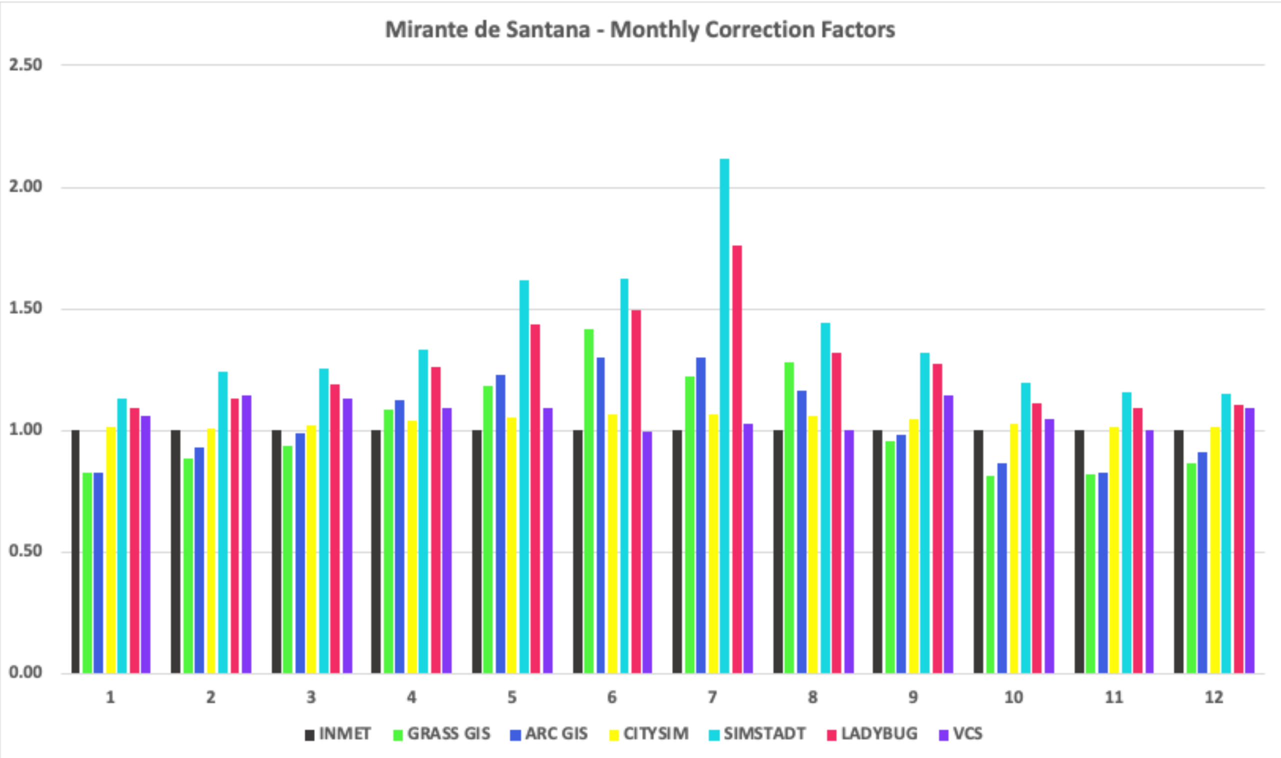


Figure 5-16 Mirante de Santana – Monthly Correction Factors.

It is important to observe that SimStadt and Ladybug present factors greater than one throughout the year, and these factors increase significantly during winter months, especially in the case of July. This means that the underestimation of solar irradiation for the meteorological station is exacerbated in wintertime with Ladybug and SimStadt, which are the ones based on direct and diffuse irradiance values. The other four software, in turn, present values that fluctuate around the threshold of 1.00.

Once the global solar irradiation values for Brasilândia are adjusted, annual and monthly analyses are carried out. Such analyses are performed by:

- visual interpretation of the results by means of an atlas, which represents the spatial and temporal distribution of this phenomenon across the buildings in the scenario. In this sense, thirteen boards are generated – an annual one- and twelve-monthly ones –, each of these containing eight maps (13x8): GRASS GIS and ArcGIS – both in their original raster and vector models –, CitySim, SimStadt, Ladybug and the VCS Solar. Analogously, the monthly maps are also presented in another eight boards, each one confronting solar irradiation values coming from the same software for the twelve months of the year (8x12).
- an individual comparison among values for some of these 301 buildings roofs. In total, eight buildings roofs are chosen for such comparison:
  - The one resulting in the **lowest** corrected annual **global irradiation** value (399 kWh/m²) among all software, i.e., the building with GML id “Building_1078833”.
  - The one resulting in the **highest** corrected annual **global irradiation** value (2609 kWh/m²) among all software, i.e., the building with GML id “Building_1078195”.
  - The one resulting in the **smallest absolute difference** of corrected global irradiation value (129.7 kWh/m²) among all software, i.e., the building with GML id “Building_1083095”.
  - The one resulting in the **largest absolute difference** of corrected global irradiation value (1593.44 kWh/m²) among all software, i.e., the building with GML id “Building_1085317”.
  - The one with the **smallest footprint area** (7.52 m²), i.e., the building with GML id “Building_1080661”.
  - The one with the **largest footprint area** (1308.1 m²), i.e., the building with GML id “Building_1088141”.
  - The **lowest building roof** (839.22 m above sea level), i.e., the building with GML id “Building_1077433”.
  - The **highest building roof** (867.41 m above sea level), i.e., the building with GML id “Building_1071008”.

The visual and numeric results are firstly assessed on an annual basis, and sequentially on a monthly basis.

Brasilândia City Block – Corrected Global Solar Irradiation – **ANNUAL** – kWh/m²

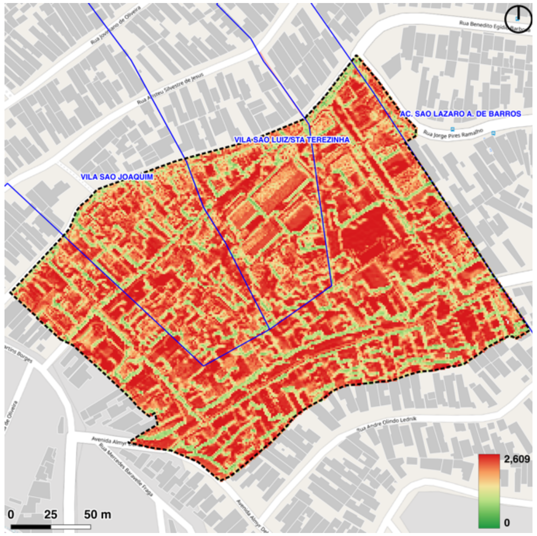


Figure 5-17 Brasilândia GRASS raster ANNUAL.

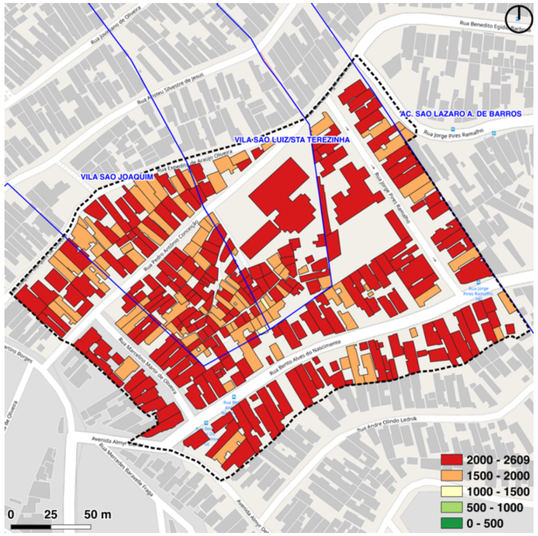


Figure 5-18 Brasilândia GRASS vector ANNUAL.

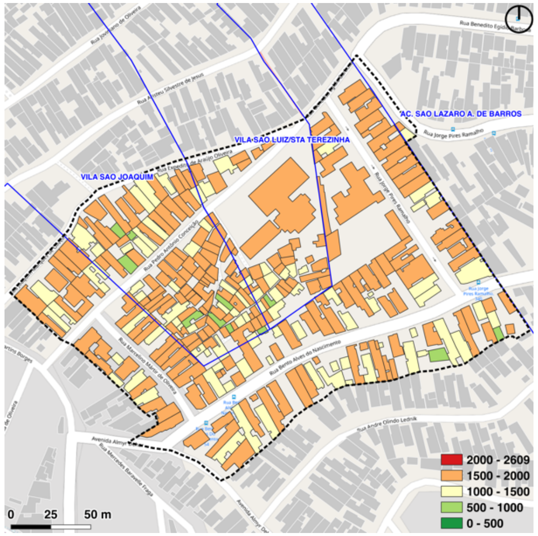


Figure 5-19 Brasilândia CitySim ANNUAL.

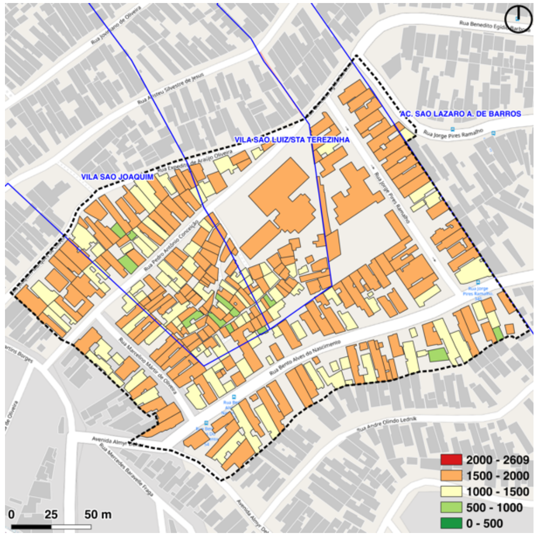


Figure 5-20 Brasilândia Ladybug ANNUAL.

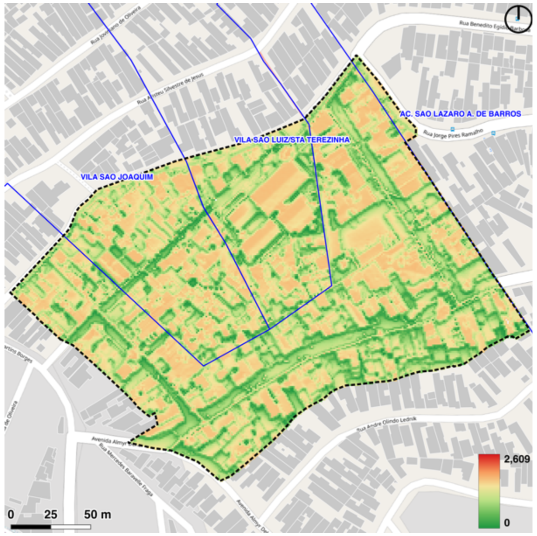


Figure 5-21 Brasilândia ArcGIS raster ANNUAL

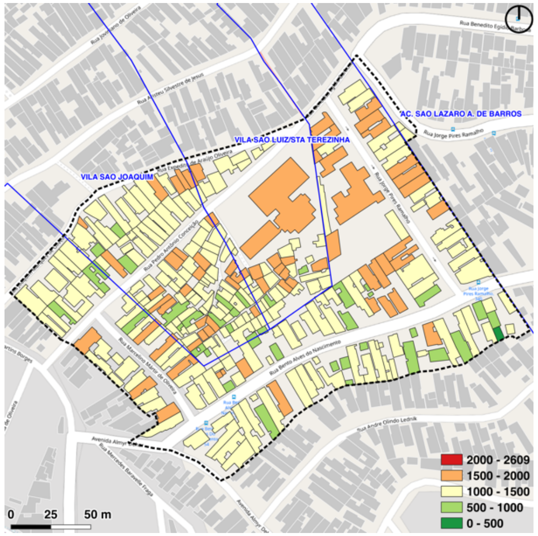


Figure 5-22 Brasilândia ArcGIS vector ANNUAL.

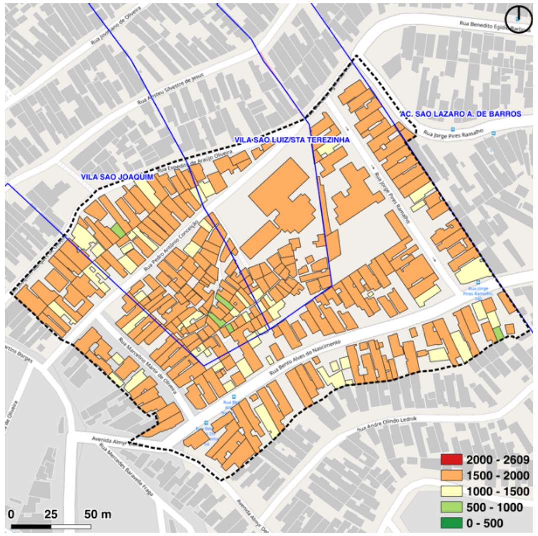


Figure 5-23 Brasilândia SimStadt ANNUAL.

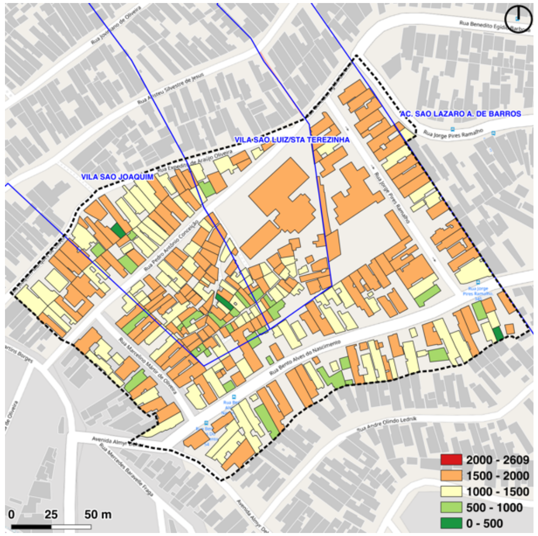


Figure 5-24 Brasilândia VCS ANNUAL.

In an overview of the annual corrected global solar irradiation values throughout the buildings in Brasilândia, it is possible to observe substantial differences among the solar modules. Even after applying the correction factors, GRASS GIS still results in elevated annual values with respect the other software, with most buildings passing the mark of 2.000 kWh/m²/year. In the case of ArcGIS, these values are, in general, less intense than the ones coming from all other models, with most buildings below the mark of 1.500 kWh/m²/year. CitySim, SimStadt, Ladybug and Virtual City Systems, in turn, result in most irradiation values on building roofs between 1.500 and 2.000 kWh/m²/year. In the case of this last one, a major contrast is noticeable, with some building roofs in the VCS Solar dataset presenting values that are even below the mark of 500 kWh/m²/year.

Complementary, for each solar module and for each building, multiplying the resulting individual irradiation values (kWh/m²) per their corresponding building roof area (m²) results in an annual amount of solar radiation energy (kWh). Sequentially, by summating these building roof values of radiation and then dividing the achieved result by the global area of roof surfaces, it is possible to retrieve six weighted averages of solar global irradiation for the scenario of Brasilândia as a whole, one average for each solar module.

ArcGIS and VCS Solar result in lower and similar weighted averages, respectively 1346 and 1460 kWh/m²/year. Slightly higher are the averages for CitySim, SimStadt and Ladybug, respectively 1543, 1667 and 1532 kWh/m²/year. GRASS GIS, in turn, is the solar module with the highest weighted average, 2112 kWh/m²/year.

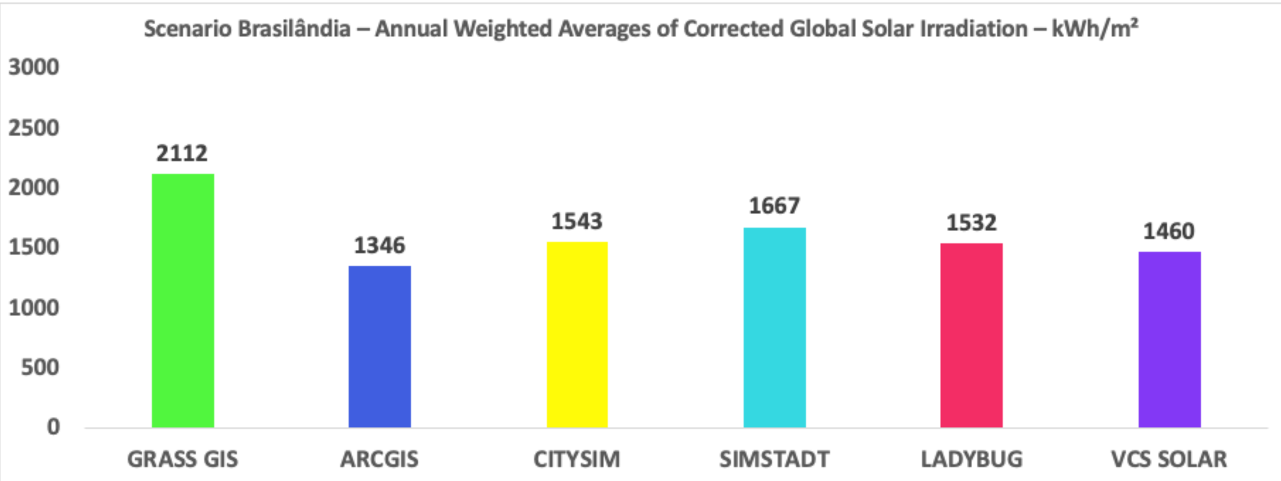


Figure 5-25 Scenario Brasilândia – Annual Weighted Averages of Corrected Global Solar Irradiation – kWh/m²

With respect to the individual building comparison, the following table presents the corrected annual global solar irradiation values for these for the eight selected buildings.

	BUILDING GML ID	ROOF ALTITUDE (m)	FOOTPRINT AREA (m²)	GRASS GIS (kWh/m²)	ARCGIS (kWh/m²)	CITYSIM (kWh/m²)	SIMSTADT (kWh/m²)	LADYBUG (kWh/m²)	VCS SOLAR (kWh/m²)	(MAX-MIN) DIFF. (kWh/m²)
LOWEST ANNUAL GLOBAL IRRADIATION	Building_1078833	844.99	39.44	1813	451	664	802	692	399	1414.22
HIGHEST ANNUAL GLOBAL IRRADIATION	Building_1078195	848.76	40.59	2609	1339	1496	1598	1416	1352	1269.85
SMALLEST ABSOLUTE IRRAD. DIFFERENCE	Building_1083095	858.82	106.47	1782	1657	1670	1786	1663	1662	129.70
LARGEST ABSOLUTE IRRAD. DIFFERENCE	Building_1085317	840.05	14.86	2129	801	1158	1186	1126	536	1593.44
SMALLEST FOOTPRINT AREA	Building_1080661	858.76	7.52	1688	939	1183	1365	1223	1021	749.46
LARGEST FOOTPRINT AREA	Building_1088141	861.68	1308.1	2136	1580	1669	1789	1659	1663	555.45
LOWEST BUILDING ROOF	Building_1077433	839.22	46.61	1744	646	820	853	787	469	1275.12
HIGHEST BUILDING ROOF	Building_1071008	867.41	28.95	1682	1511	1672	1789	1665	1665	277.57

The building with GML id **"Building_1078833"** is the one presenting the lowest corrected annual global irradiation value (399 kWh/m²/year), which comes from the Virtual City System solar module. For this building, ArcGIS also results in a similarly low irradiation value (451 kWh/m²/year). CitySim, SimStadt and Ladybug, in turn, result in similar low to moderate irradiation values, 664, 802 and 692 kWh/m²/year, respectively. GRASS GIS is the one that diverges the most, with an annual global solar irradiation of 1813 kWh/m²/year, i.e., more than three times the average of the other five values.

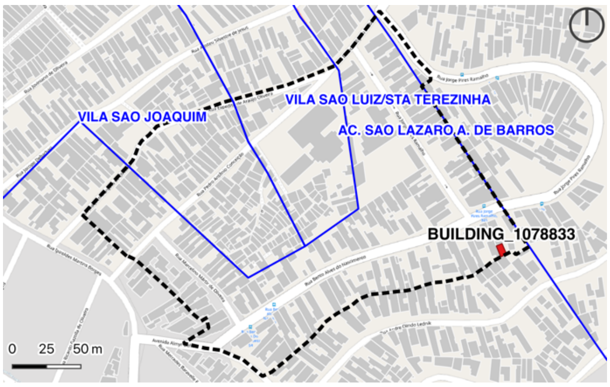


Figure 5-26 Building_1078833 in the city block of Brasilândia.

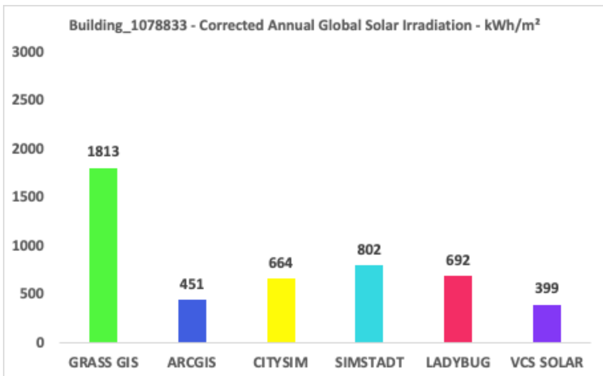


Figure 5-27 Building_1078833 – Corrected Annual Global Solar Irradiation – kWh/m².

The building with GML id **"Building_1078195"** is the one presenting the highest corrected annual global irradiation value (2609 kWh/m²/year), which comes from GRASS GIS. In the case of this building roof, it is noticeable that all other software result in relatively similar values, i.e., an average of 1440 kWh/m²/year excluding GRASS GIS, whereas the value of the latter is 1.81 times this average.

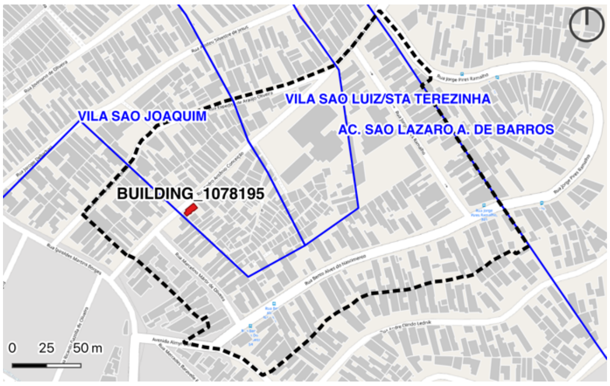


Figure 5-28 Building_1078195 in the city block of Brasilândia.

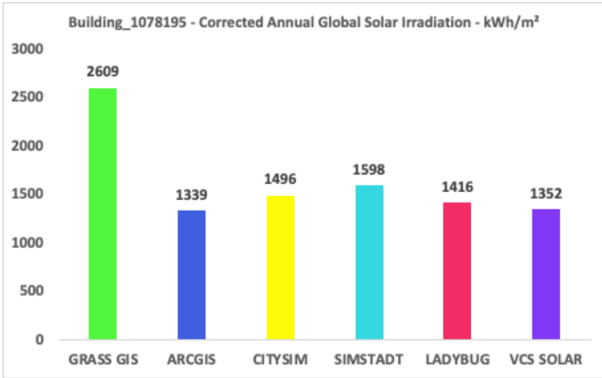


Figure 5-29 Building_1078195 – Corrected Annual Global Solar Irradiation – kWh/m².

The building with GML id **“Building_1083095”** is the one presenting the smallest annual difference in corrected global solar irradiation values: the difference from SimStadt (1786 kWh/m²/year) and VCS Solar (1662 kWh/m²/year) is only 129.7 kWh/m²/year, i.e., 7.6% of the annual average value coming from the six software.

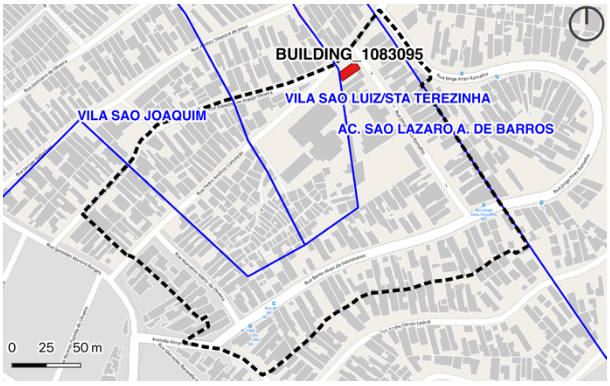


Figure 5-30 Building_1083095 in the city block of Brasilândia.

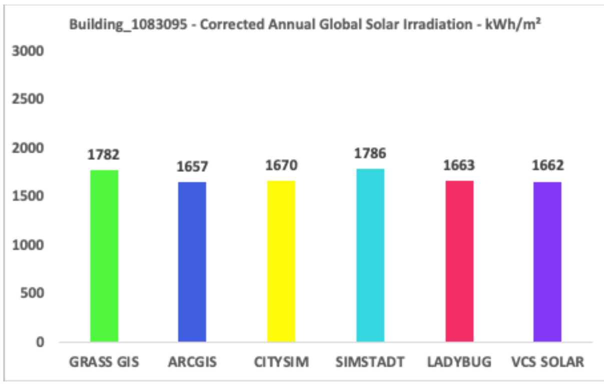


Figure 5-31 Building_1083095 – Corrected Annual Global Solar Irradiation – kWh/m².

The building with GML id **“Building_1085317”** is the one presenting the largest annual difference in corrected global solar irradiation values: the difference from GRASS GIS (2129 kWh/m²/year) and VCS Solar (536 kWh/m²/year) reaches 1593.44 kWh/m²/year, i.e., 137.9% of the annual average value coming from the six software.

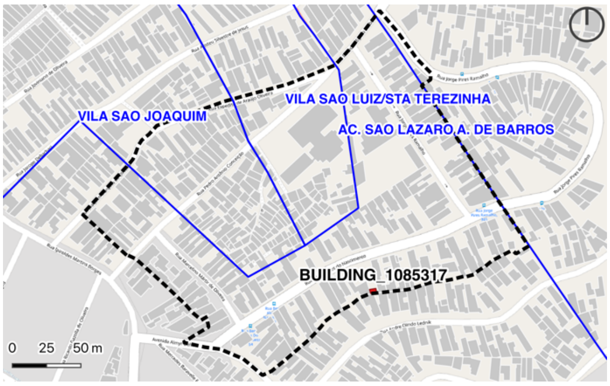


Figure 5-32 Building_1085317 in the city block of Brasilândia.

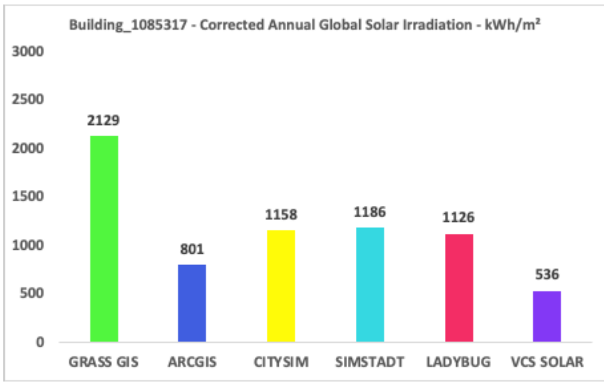


Figure 5-33 Building_1085317 – Corrected Annual Global Solar Irradiation – kWh/m².

The building with GML id **"Building_1080661"** is the one with the smallest footprint, which is only 7.52 m² large. Since irradiation is a ratio between solar radiation (kWh) and area (m²), the algorithms must perform this averaging to reach a single irradiation value from the radiation values at the building roofs. A smaller building footprint results in less computed radiation values in a grid structure of fixed spatial resolution, as the ones adopted in GRASS GIS and ArcGIS – the raster-based ones – but also in Ladybug and VCS Solar. Therefore, the irradiation values would, in theory, be subject to larger errors The corrected results coming from this building footprint indicate that GRASS GIS is still the software with the highest corrected annual value, 1688 kWh/m², whereas the other five software result in relatively low to moderate global solar irradiation values in a year.

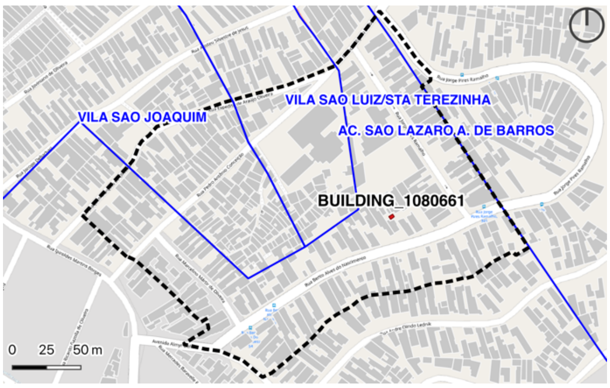


Figure 5-34 Building_1080661 in the city block of Brasilândia.

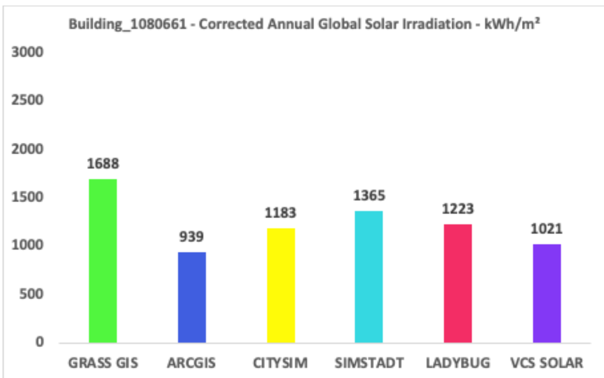


Figure 5-35 Building_1080661 – Corrected Annual Global Solar Irradiation – kWh/m².

The building with GML id **"Building_1088141"**, in turn, is the one with the largest footprint (1308.1 m²). In this case, it is possible to observe a that GRASS GIS still deviates from the other five software, with an annual value of 2136 kWh/m²/year. Nevertheless, the results coming from the other five software are much alike each other, in an average of 1672 kWh/m²/year and a standard deviation of 74.80 kWh/m²/year.

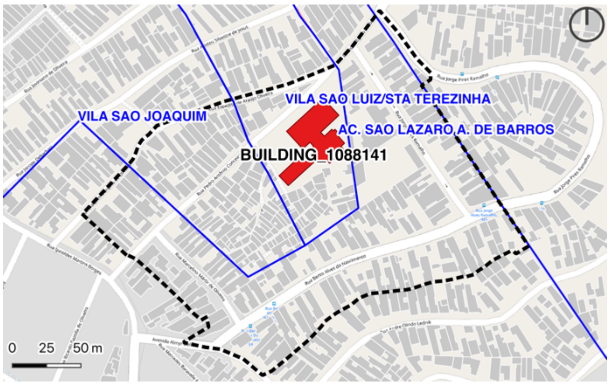


Figure 5-36 Building_1088141 in the city block of Brasilândia.

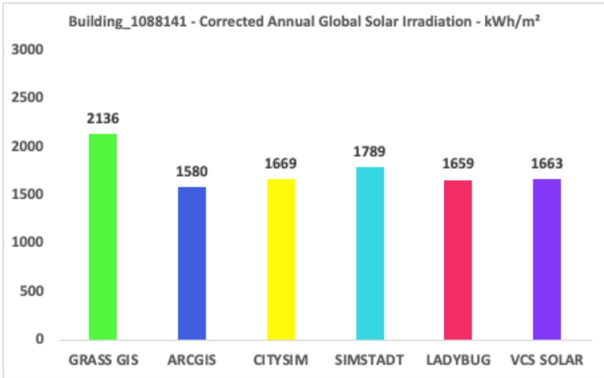


Figure 5-37 Building_1088141 – Corrected Annual Global Solar Irradiation – kWh/m².

The building with GML id **"Building_1077433"** is the one with the lowest roof among all 301 buildings in the extended scenario, i.e., 839.22 m above sea level. In the case of this building, all software but GRASS GIS result in relatively low/very low values of global solar irradiation, i.e., 469 kWh/m²/year coming from VCS Solar to 853 kWh/m²/year coming from SimStadt. GRASS GIS, in turn, is the exception, with an analogous value of 1744 kWh/m²/year, i.e., 2.44 more than the average of the other five values.

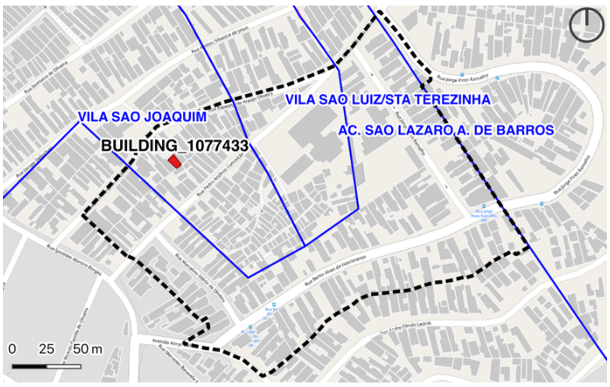


Figure 5-38 Building_1077433 in the city block of Brasilândia.

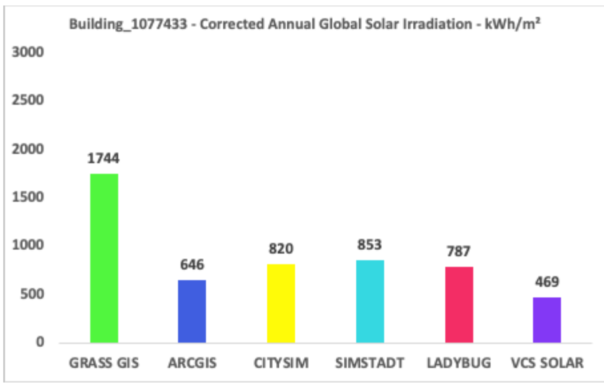


Figure 5-39 Building_1077433 – Corrected Annual Global Solar Irradiation – kWh/m².

The building with GML id **"Building_1071008"** is the one with the highest roof among all 301 buildings in the extended scenario, i.e., 867.41 m above sea level. From a geometrical perspective, the highest building roof is the one exempt of any shading effect coming from other buildings in the surroundings. Additionally, in the case of this scenario, there is neither a higher relief component / portion of DSM nor a vegetation object that is higher than this building roof. The resulting corrected annual global solar irradiation values are very much similar, ranging from 1511 kWh/m²/year (ArcGIS) to 1789 kWh/m²/year (SimStadt).

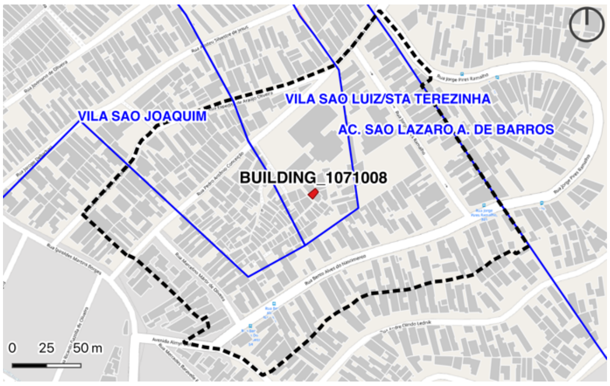


Figure 5-40 Building_1071008 in the city block of Brasilândia.

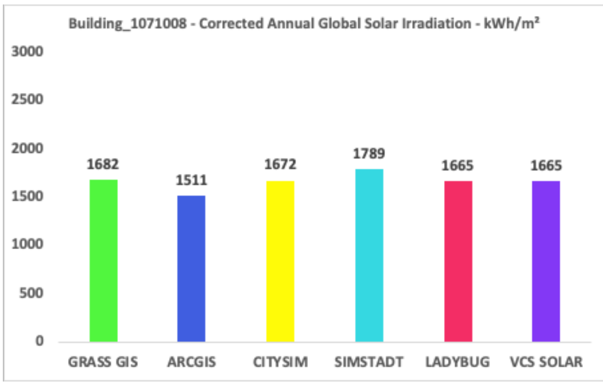


Figure 5-41 Building_1071008 – Corrected Annual Global Solar Irradiation – kWh/m².

The following pages present the atlas of Corrected Monthly Global Solar Irradiation values for the scenario of Brasilândia. Successively, general remarks are stated regarding the achieved results.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **JANUARY** – kWh/m²

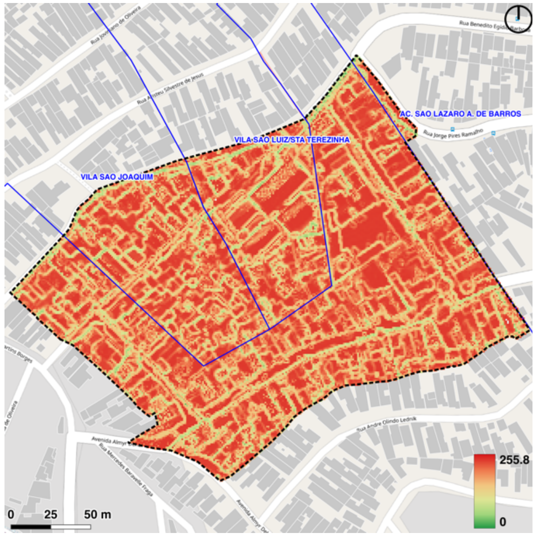


Figure 5-42 Brasilândia GRASS raster JAN.

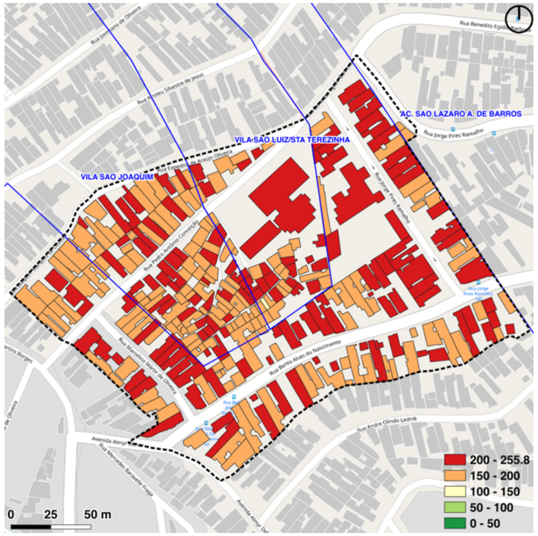


Figure 5-43 Brasilândia GRASS vector JAN.

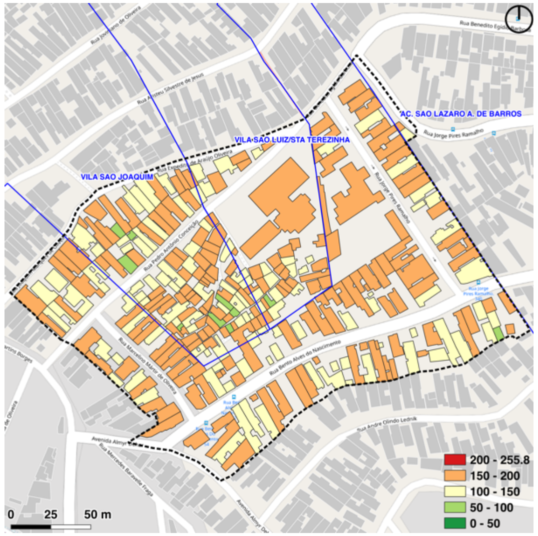


Figure 5-44 Brasilândia CitySim JAN.

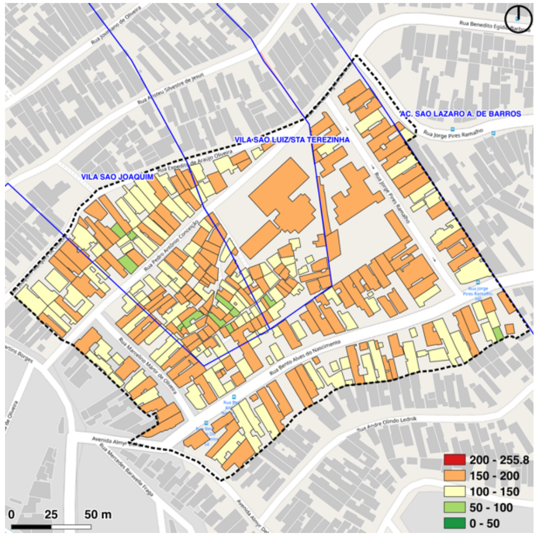


Figure 5-45 Brasilândia Ladybug JAN.

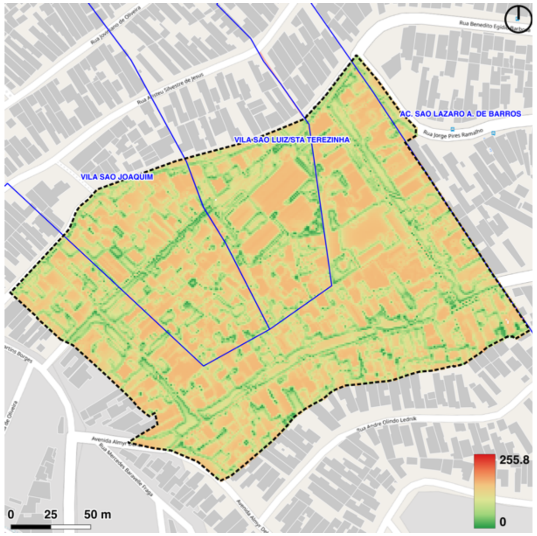


Figure 5-46 Brasilândia ArcGIS raster JAN.

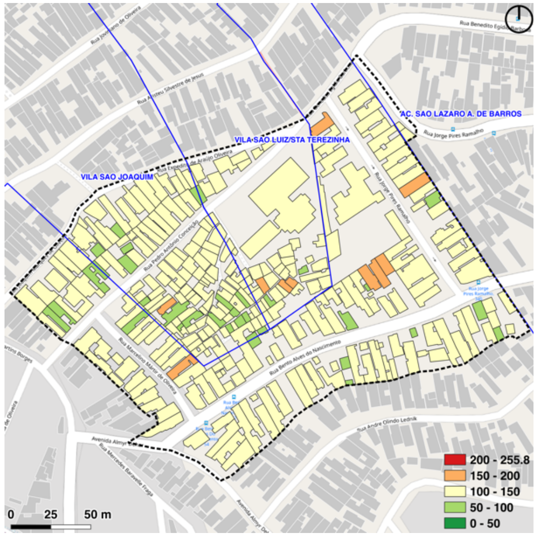


Figure 5-47 Brasilândia ArcGIS vector JAN.

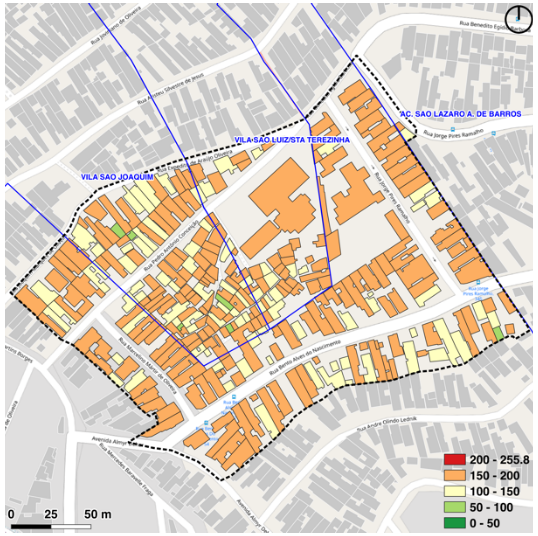


Figure 5-48 Brasilândia SimStadt JAN.

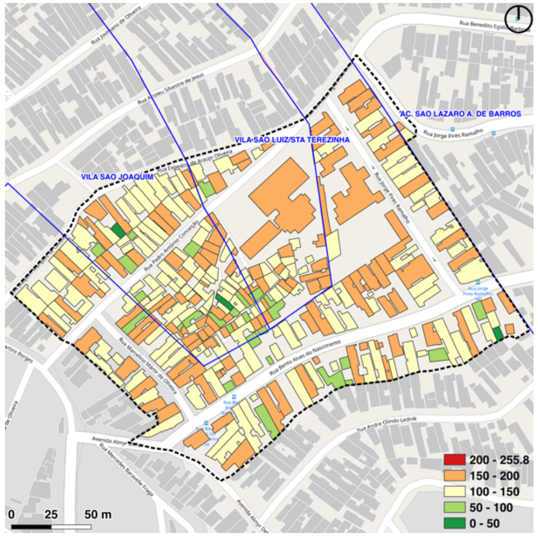


Figure 5-49 Brasilândia VCS JAN.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **FEBRUARY** – kWh/m²

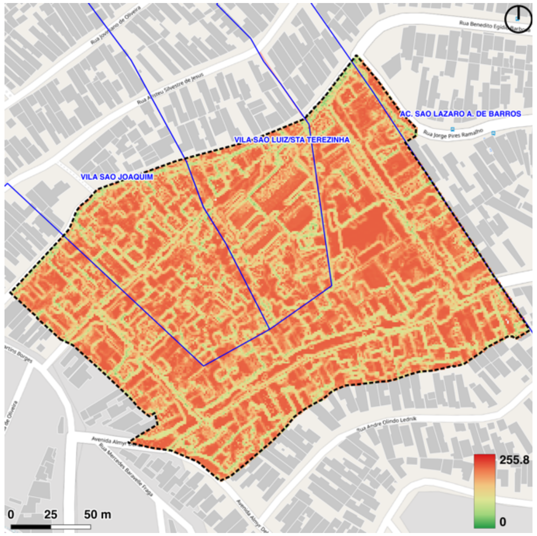


Figure 5-50 Brasilândia GRASS raster FEB.

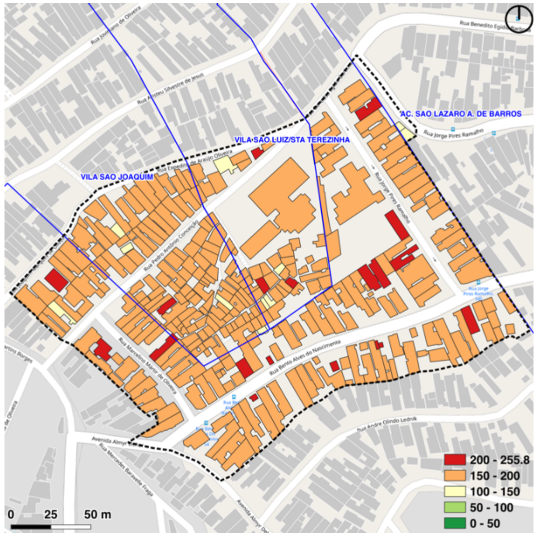


Figure 5-51 Brasilândia GRASS vector FEB.

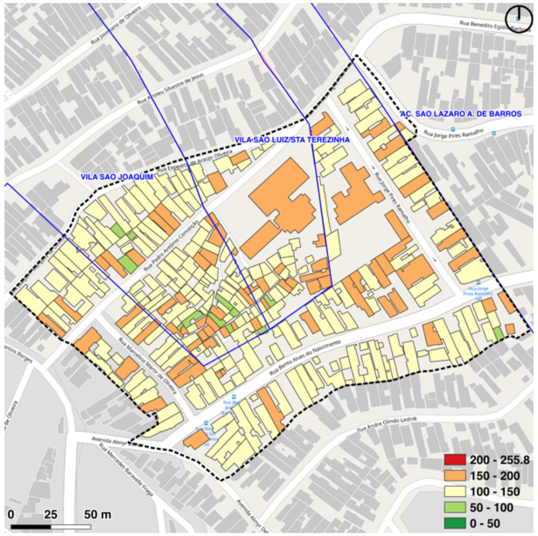


Figure 5-52 Brasilândia CitySim FEB.

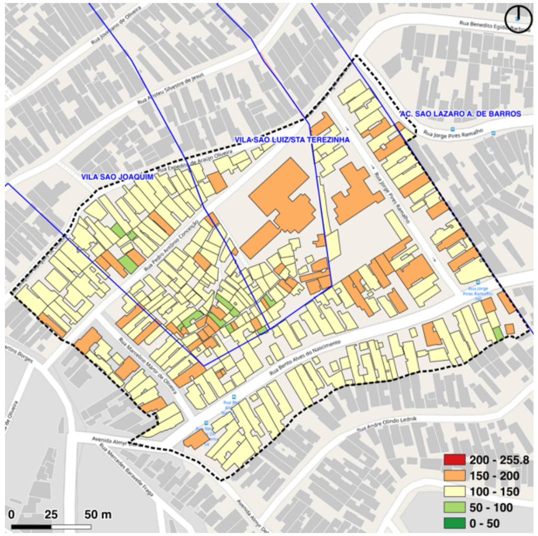


Figure 5-53 Brasilândia Ladybug FEB.

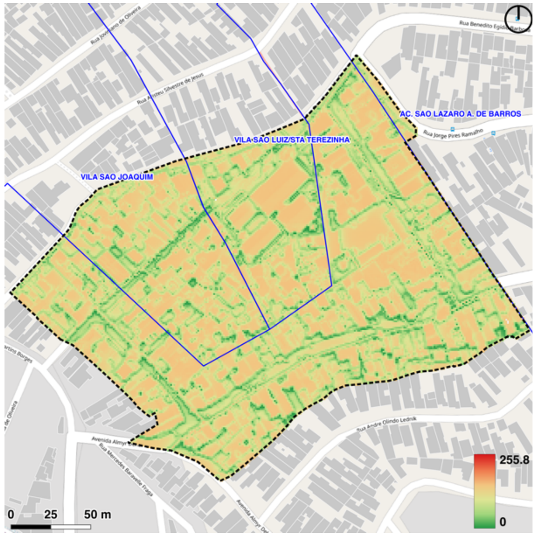


Figure 5-54 Brasilândia ArcGIS raster FEB.

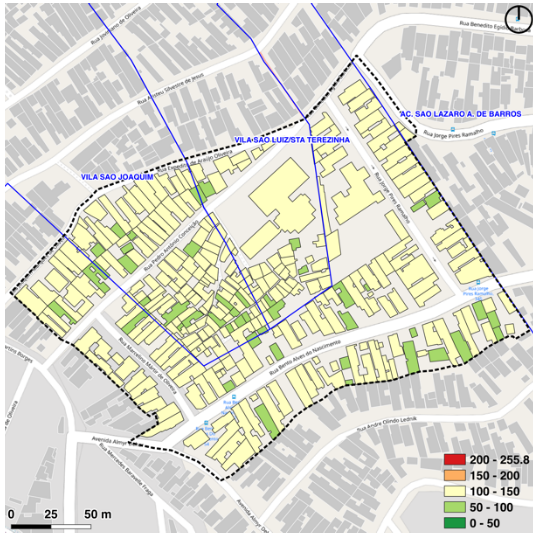


Figure 5-55 Brasilândia ArcGIS vector FEB.

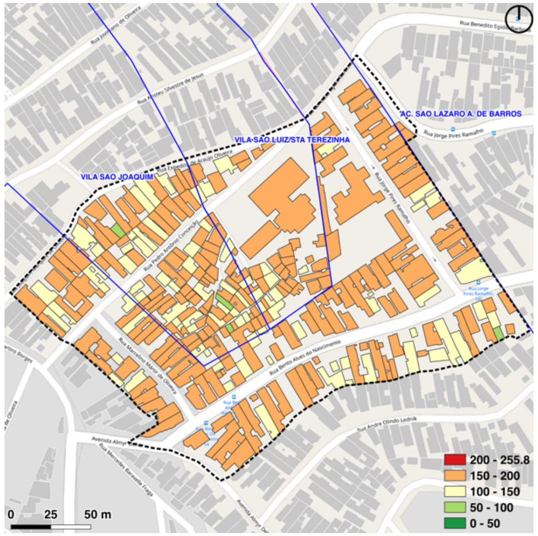


Figure 5-56 Brasilândia SimStadt FEB.

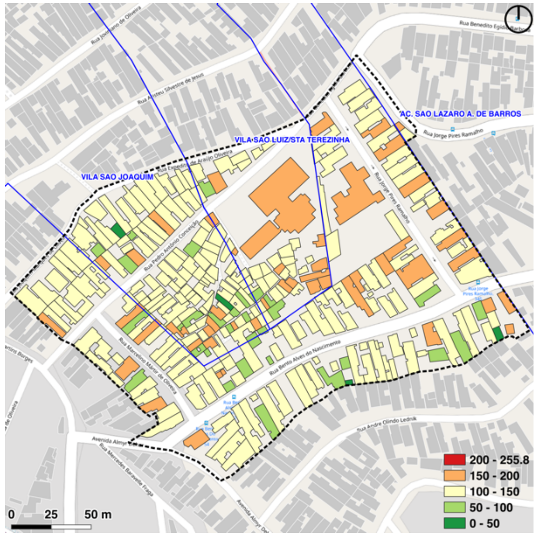


Figure 5-57 Brasilândia VCS FEB.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **MARCH** – kWh/m²

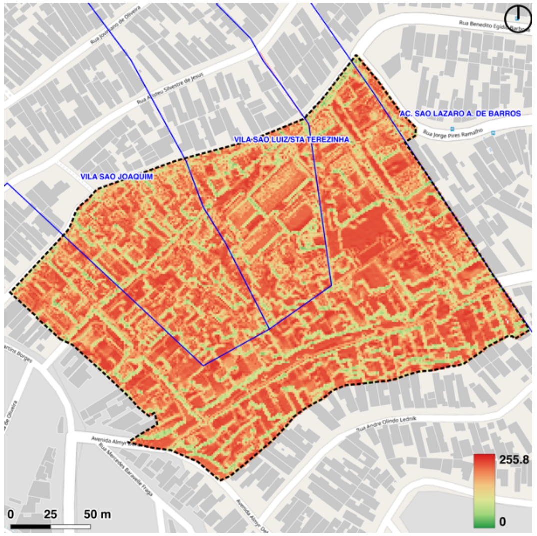


Figure 5-58 Brasilândia GRASS raster MAR.

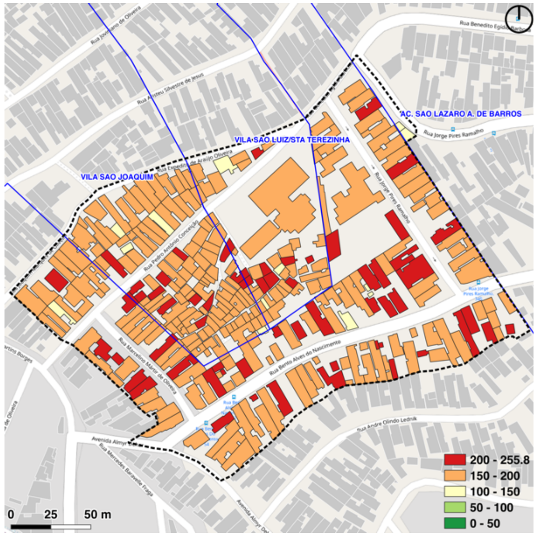


Figure 5-59 Brasilândia GRASS vector MAR

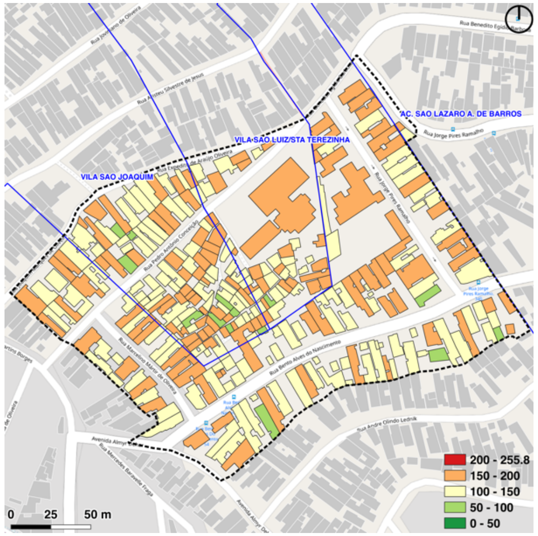


Figure 5-60 Brasilândia CitySim MAR.

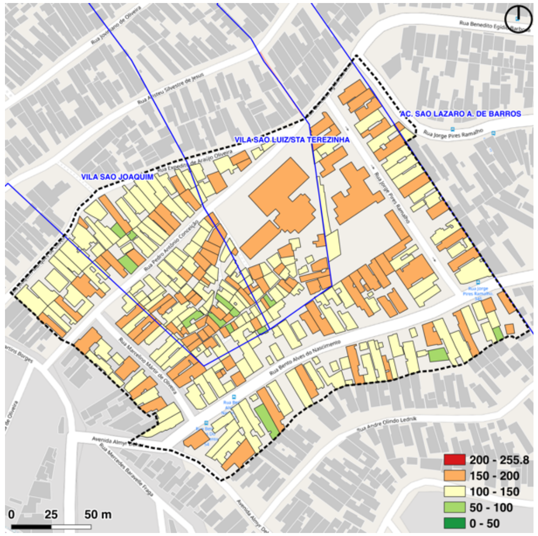


Figure 5-61 Brasilândia Ladybug MAR.

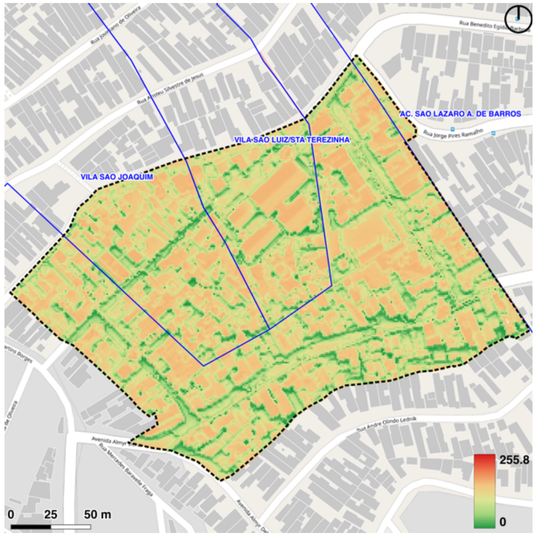


Figure 5-62 Brasilândia ArcGIS raster MAR.

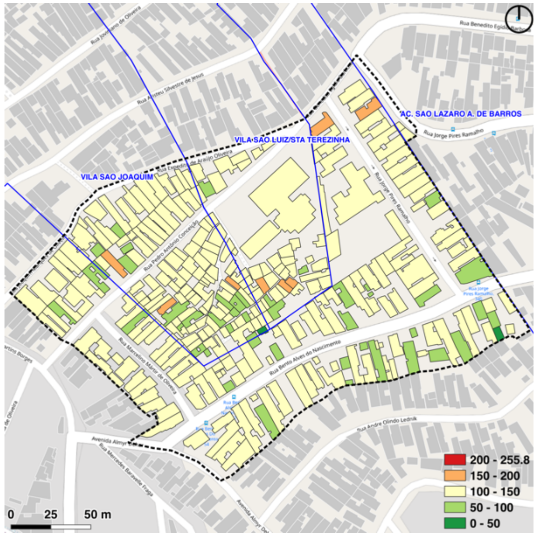


Figure 5-63 Brasilândia ArcGIS vector MAR.

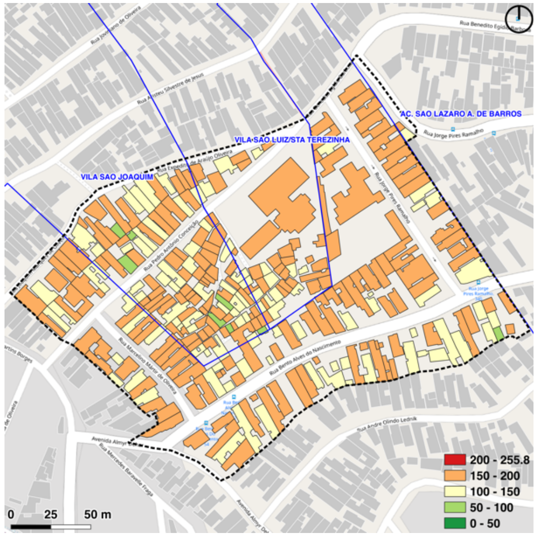


Figure 5-64 Brasilândia SimStadt MAR.

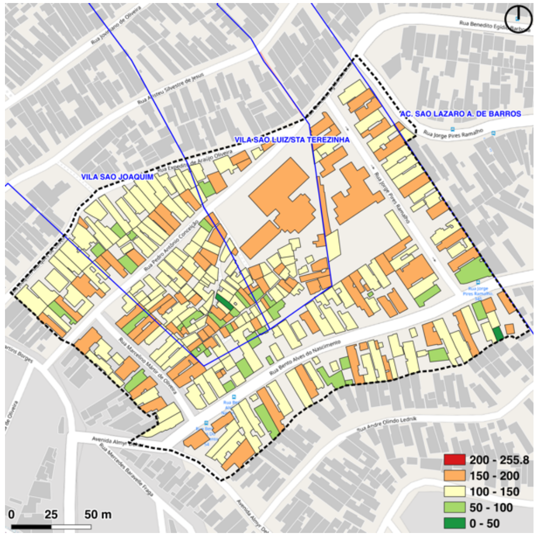


Figure 5-65 Brasilândia VCS MAR.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **APRIL** – kWh/m²

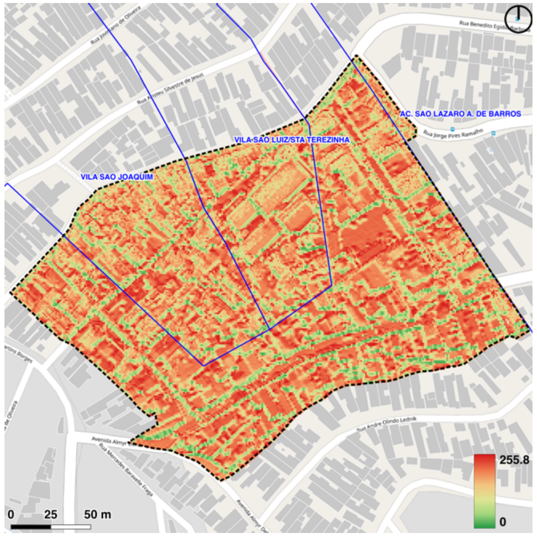


Figure 5-66 Brasilândia GRASS raster APR.

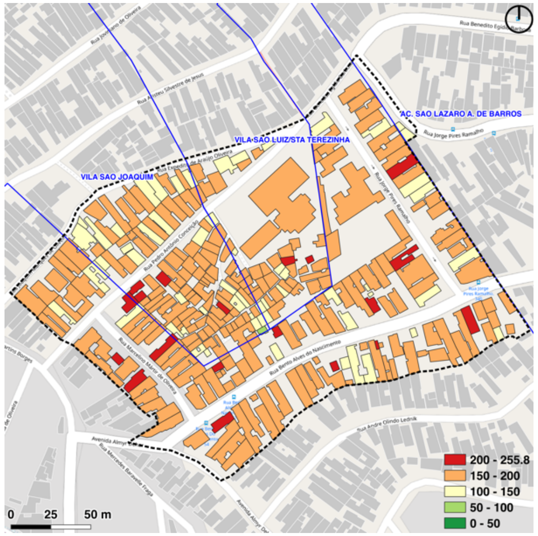


Figure 5-67 Brasilândia GRASS vector APR.

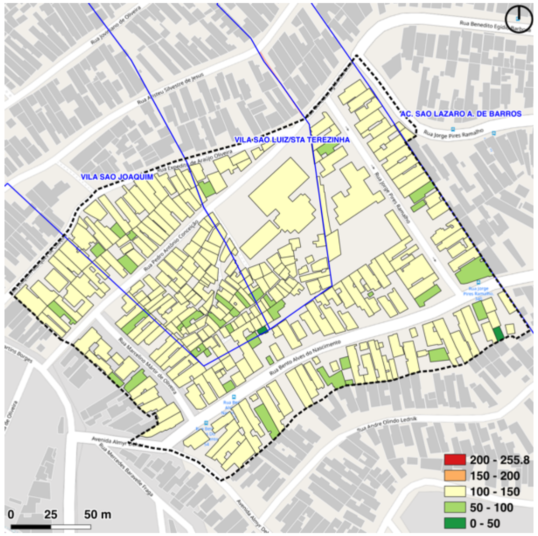


Figure 5-68 Brasilândia CitySim APR.

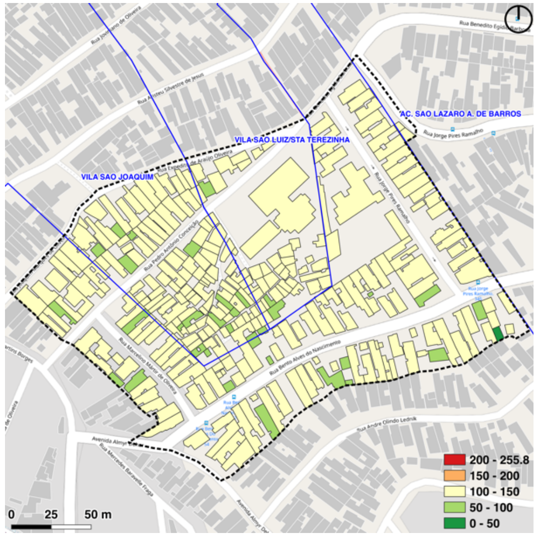


Figure 5-69 Brasilândia Ladybug APR.

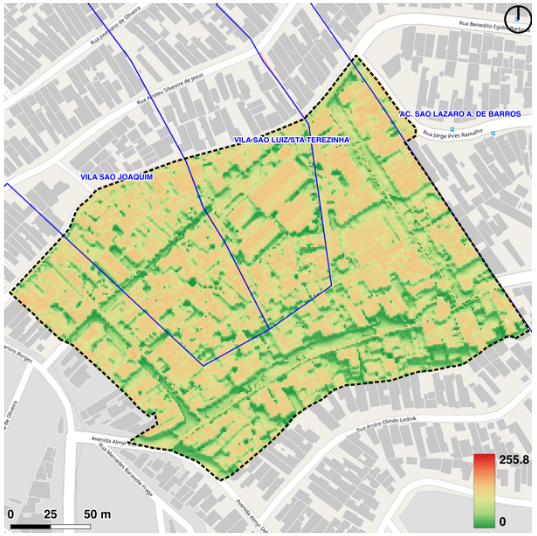


Figure 5-70 Brasilândia ArcGIS raster APR.

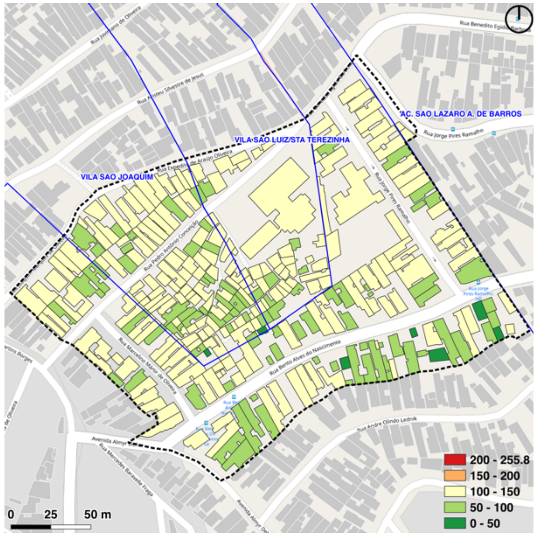


Figure 5-71 Brasilândia ArcGIS vector APR.

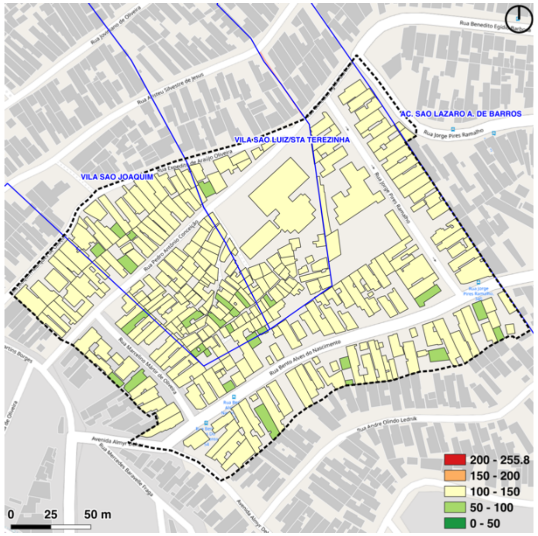


Figure 5-72 Brasilândia SimStadt APR.

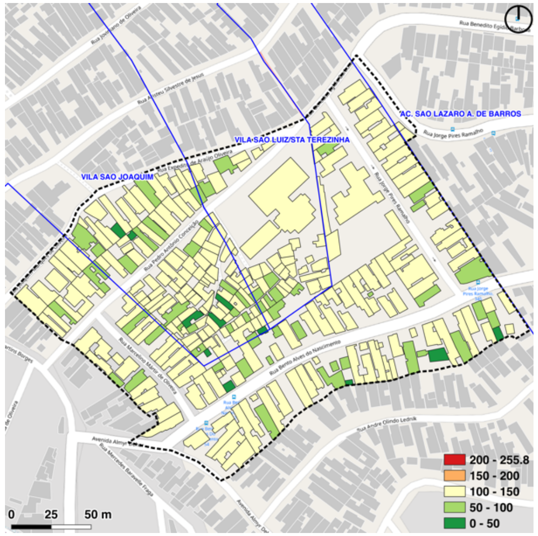


Figure 5-73 Brasilândia VCS APR.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **MAY** – kWh/m²

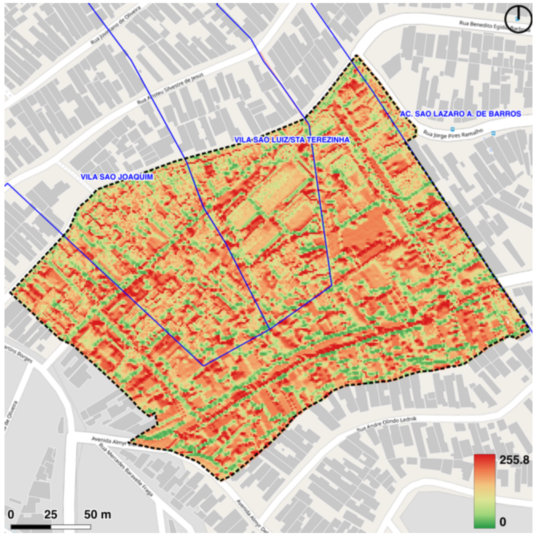


Figure 5-74 Brasilândia GRASS raster MAY.

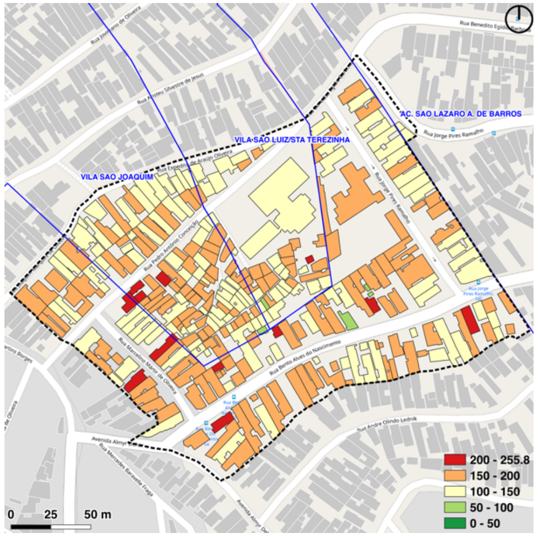


Figure 5-75 Brasilândia GRASS vector MAY.

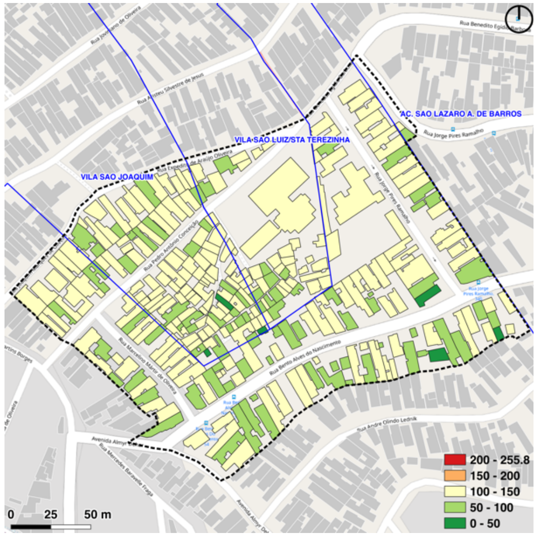


Figure 5-76 Brasilândia CitySim MAY.

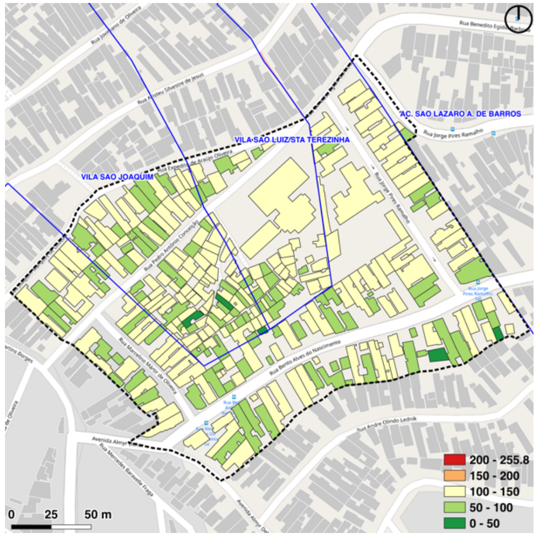


Figure 5-77 Brasilândia Ladybug MAY.

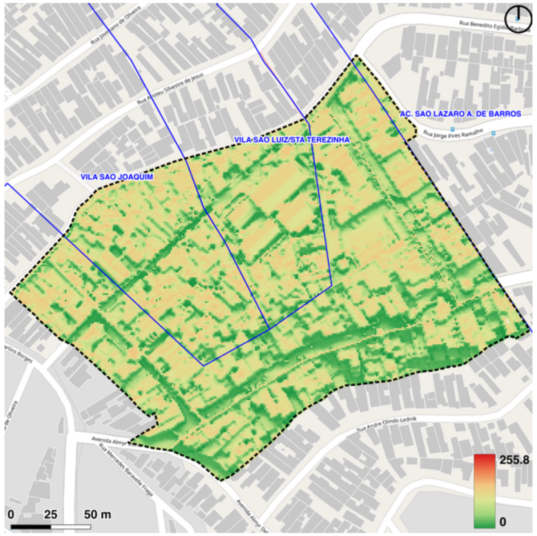


Figure 5-78 Brasilândia ArcGIS raster MAY.

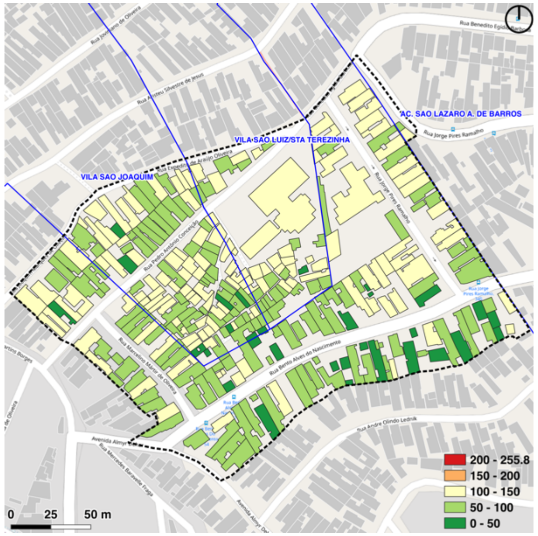


Figure 5-79 Brasilândia ArcGIS vector MAY.

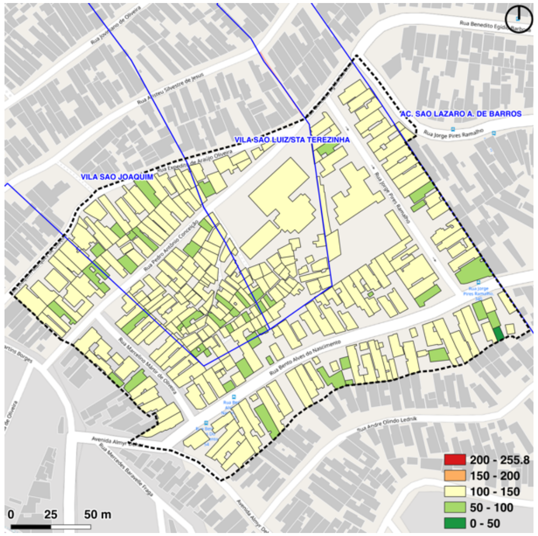


Figure 5-80 Brasilândia SimStadt MAY.

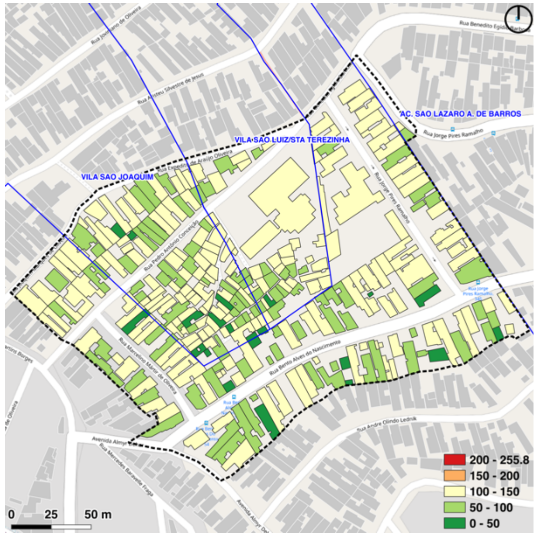


Figure 5-81 Brasilândia VCS MAY.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **JUNE** – kWh/m²

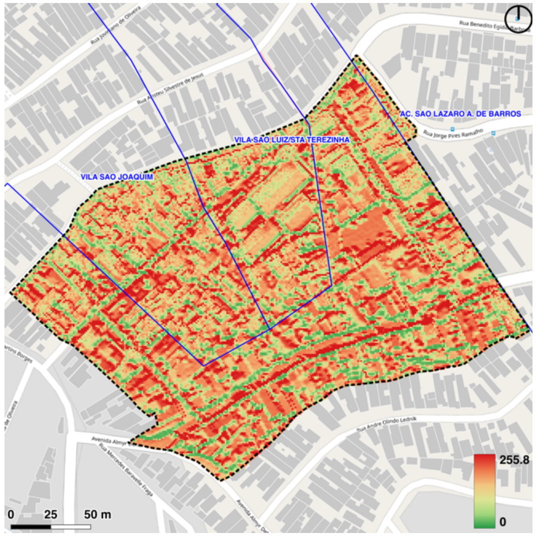


Figure 5-82 Brasilândia GRASS raster JUN.

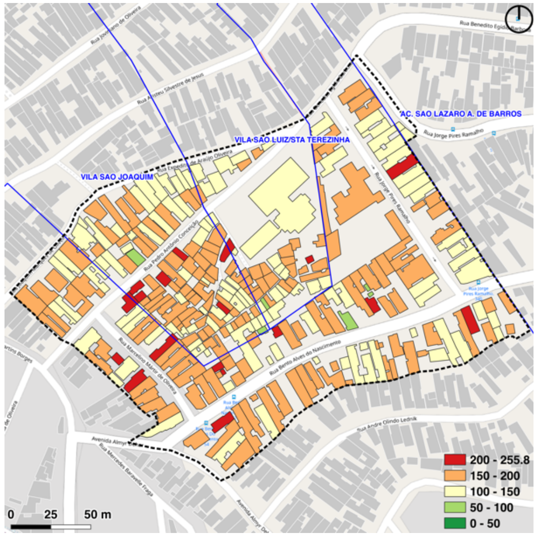


Figure 5-83 Brasilândia GRASS vector JUN.

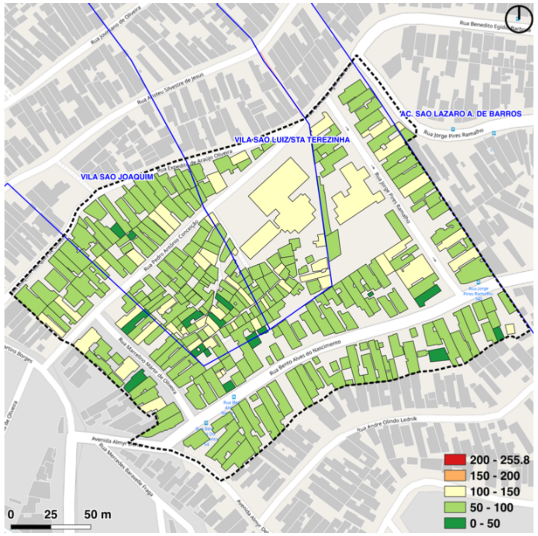


Figure 5-84 Brasilândia CitySim JUN.

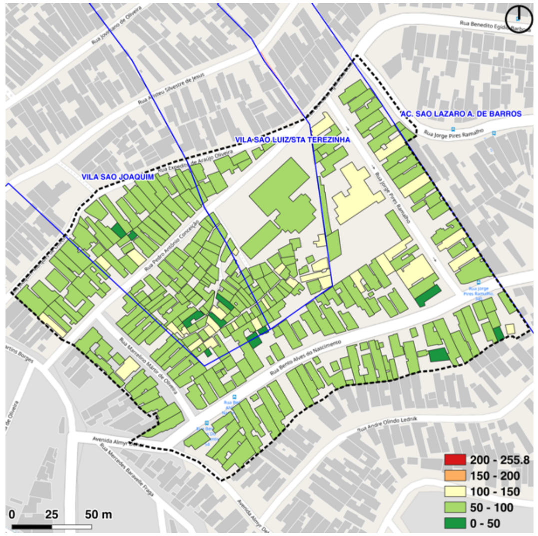


Figure 5-85 Brasilândia Ladybug JUN.

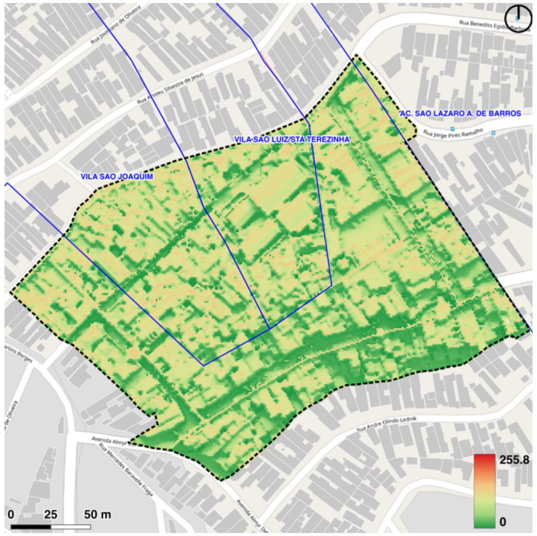


Figure 5-86 Brasilândia ArcGIS raster JUN.

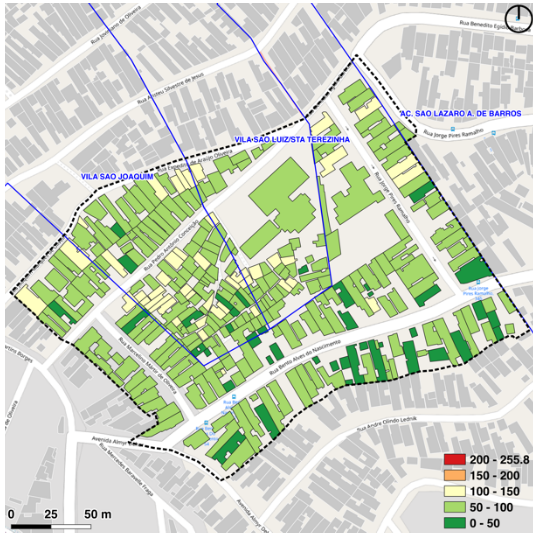


Figure 5-87 Brasilândia ArcGIS vector JUN.

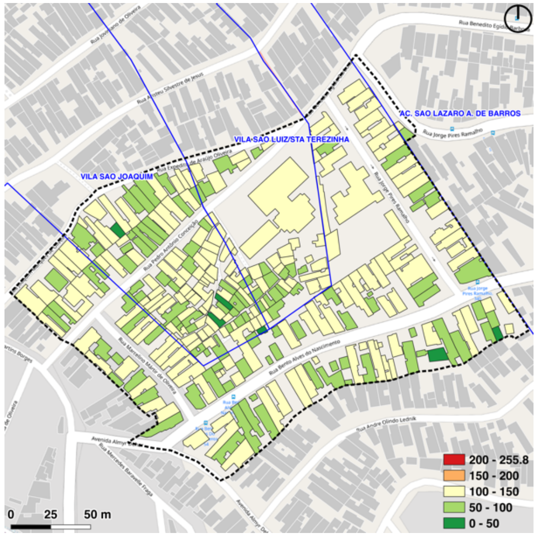


Figure 5-88 Brasilândia SimStadt JUN.

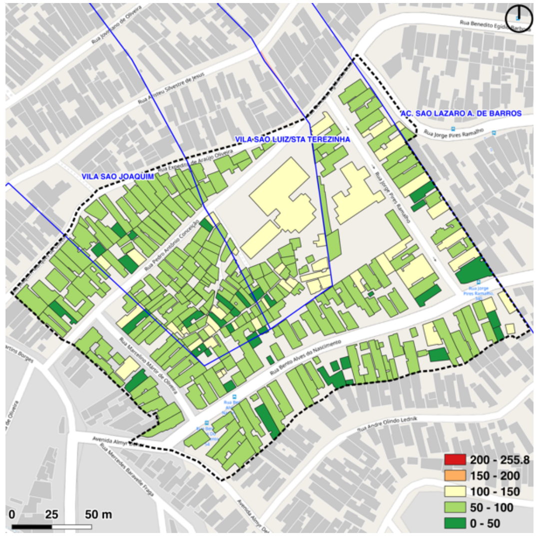


Figure 5-89 Brasilândia VCS JUN.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **JULY** – kWh/m²

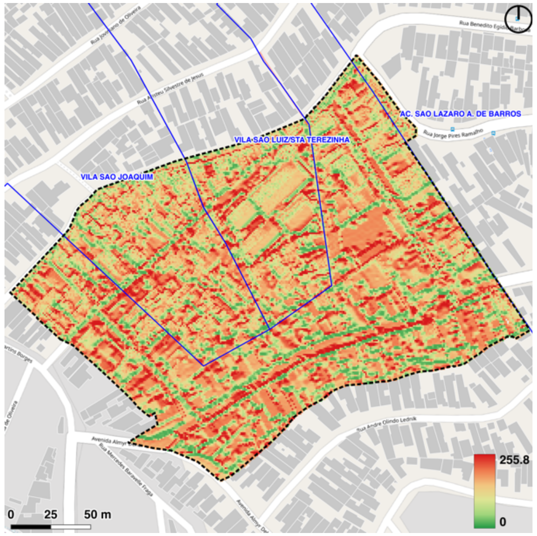


Figure 5-90 Brasilândia GRASS raster JUL.

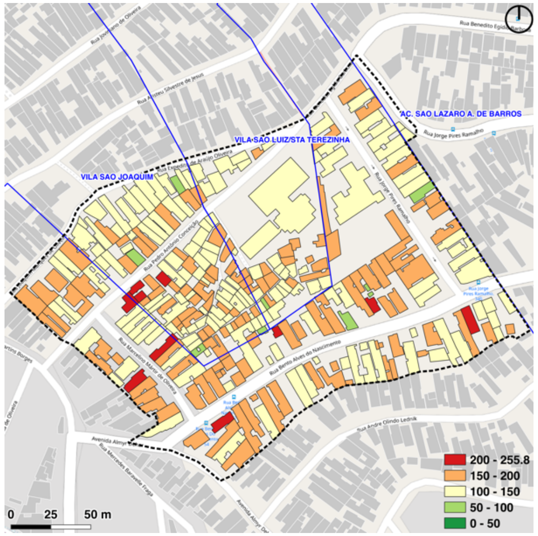


Figure 5-91 Brasilândia GRASS vector JUL.

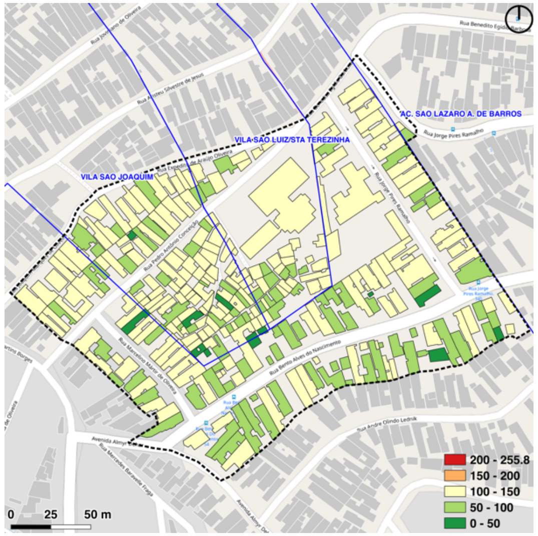


Figure 5-92 Brasilândia CitySim JUL.

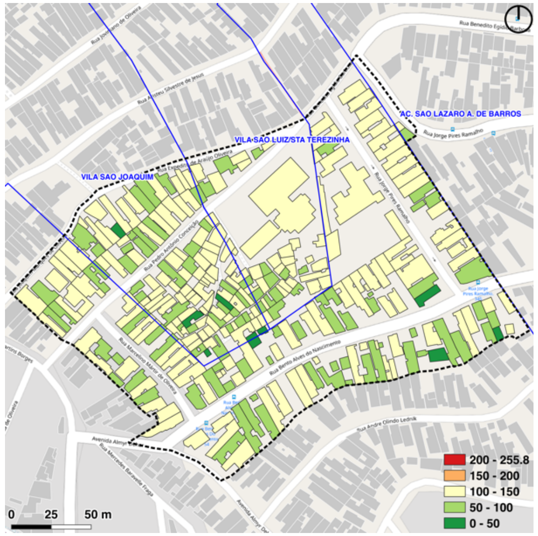


Figure 5-93 Brasilândia Ladybug JUL.

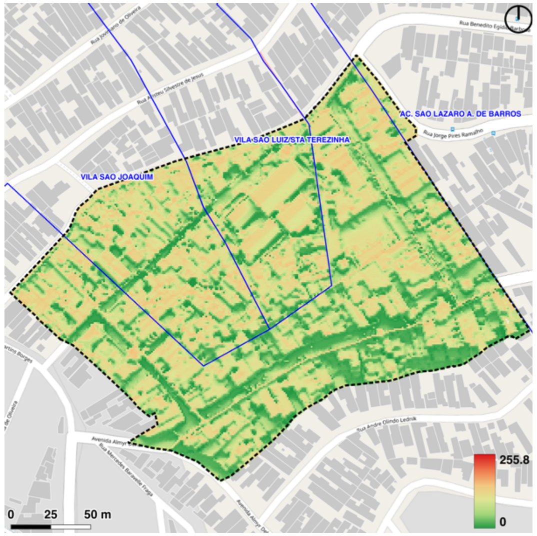


Figure 5-94 Brasilândia ArcGIS raster JUL.

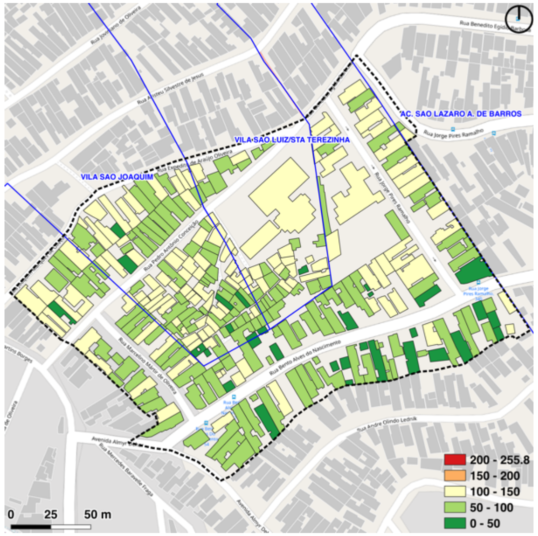


Figure 5-95 Brasilândia ArcGIS vector JUL.

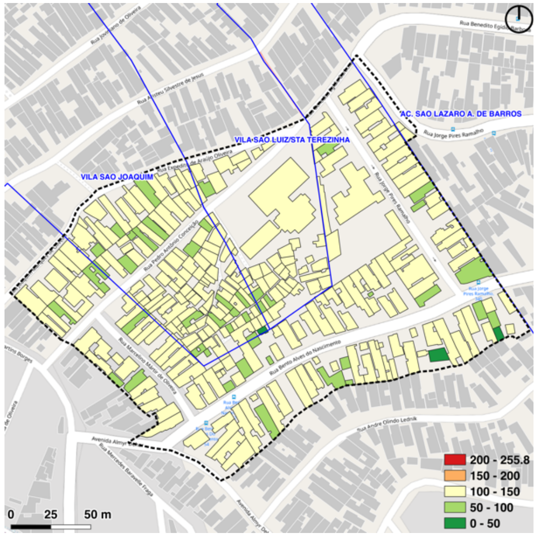


Figure 5-96 Brasilândia SimStadt JUL.

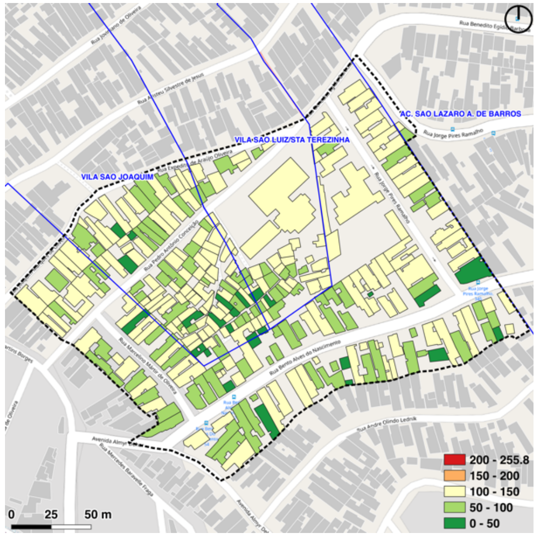


Figure 5-97 Brasilândia VCS JUL.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **AUGUST** – kWh/m²

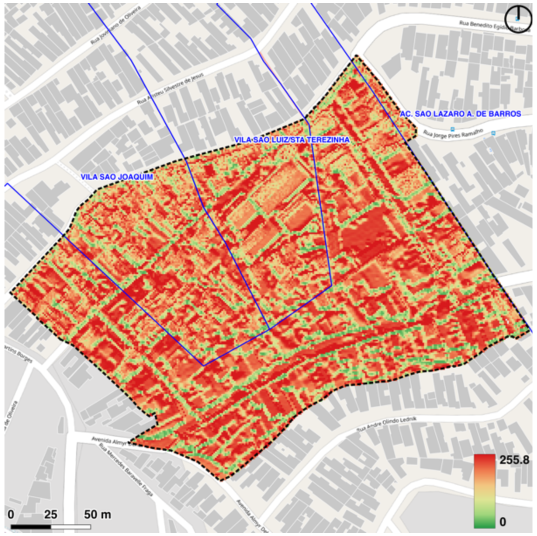


Figure 5-98 Brasilândia GRASS raster AUG.

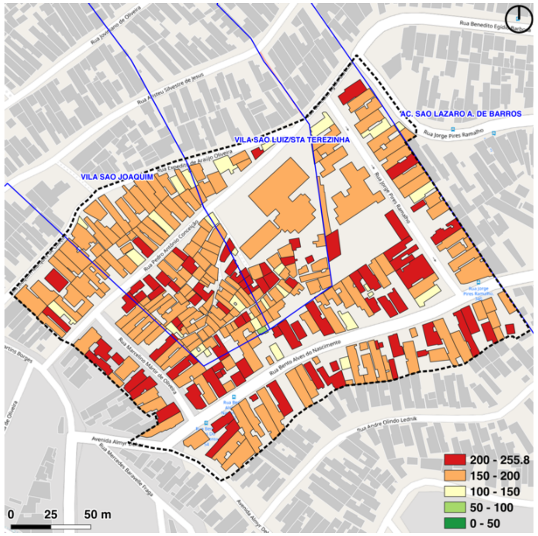


Figure 5-99 Brasilândia GRASS vector AUG.

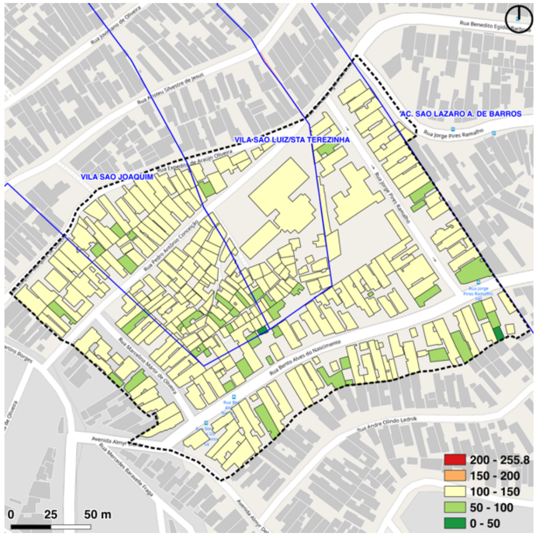


Figure 5-100 Brasilândia CitySim AUG.

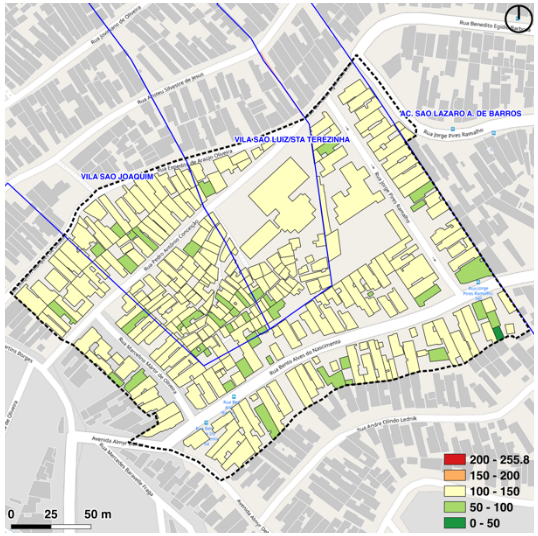


Figure 5-101 Brasilândia Ladybug AUG.

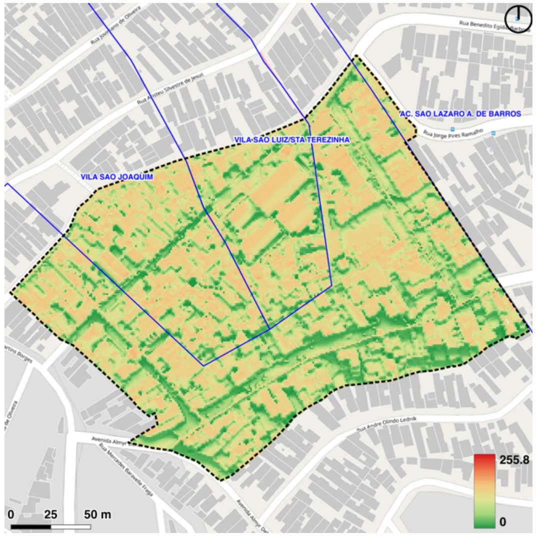


Figure 5-102 Brasilândia ArcGIS raster AUG.

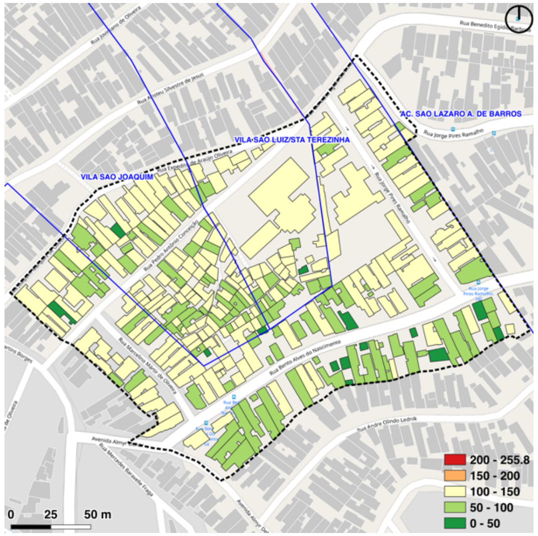


Figure 5-103 Brasilândia ArcGIS vector AUG.

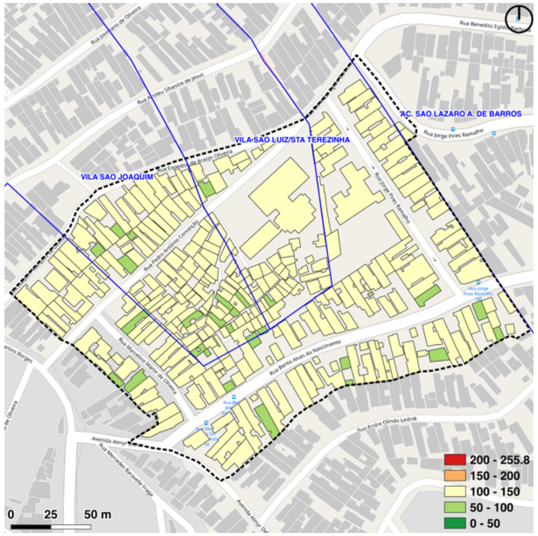


Figure 5-104 Brasilândia SimStadt AUG.

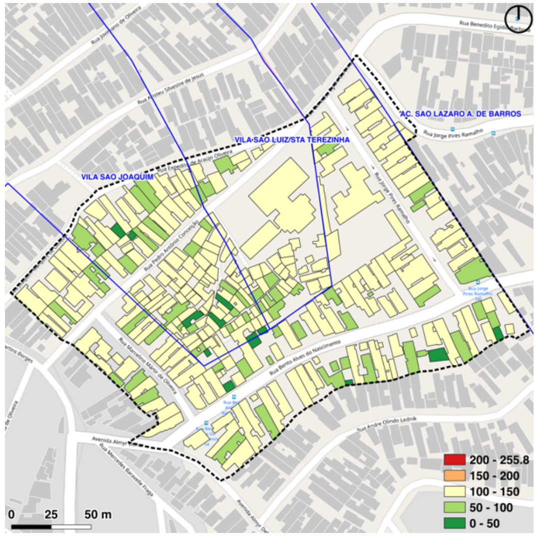


Figure 5-105 Brasilândia VCS AUG.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **SEPTEMBER** – kWh/m²

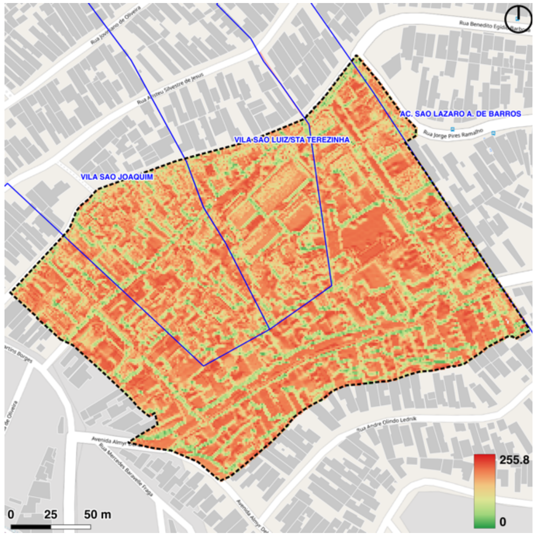


Figure 5-106 Brasilândia GRASS raster SEP.

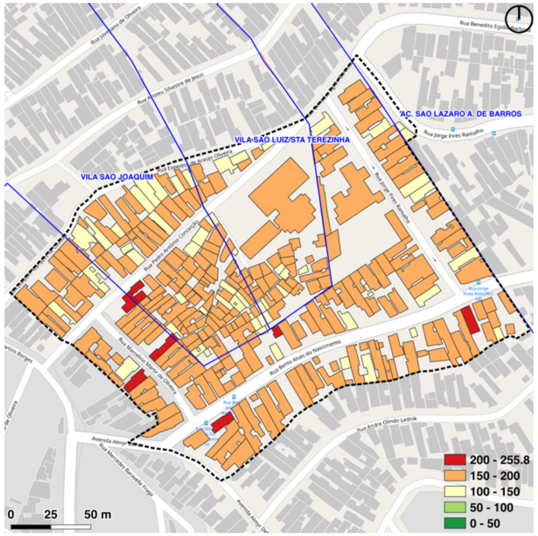


Figure 5-107 Brasilândia GRASS vector SEP.

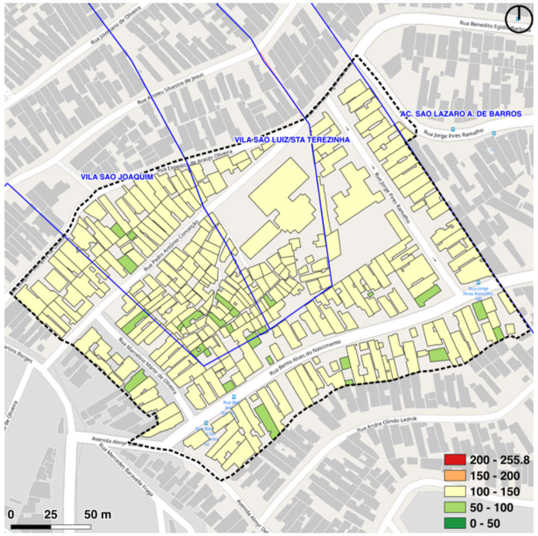


Figure 5-108 Brasilândia CitySim SEP.

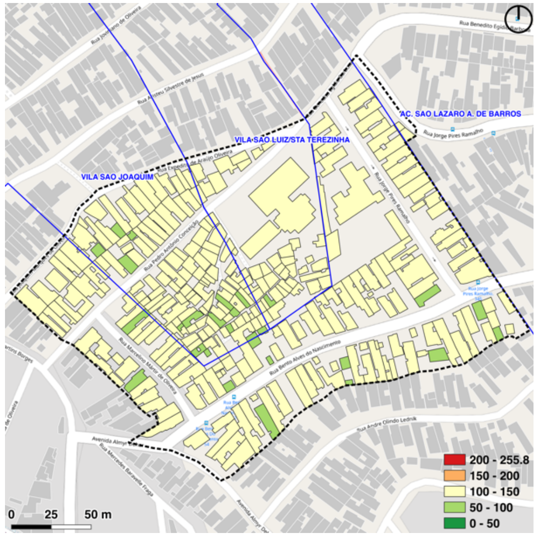


Figure 5-109 Brasilândia Ladybug SEP.

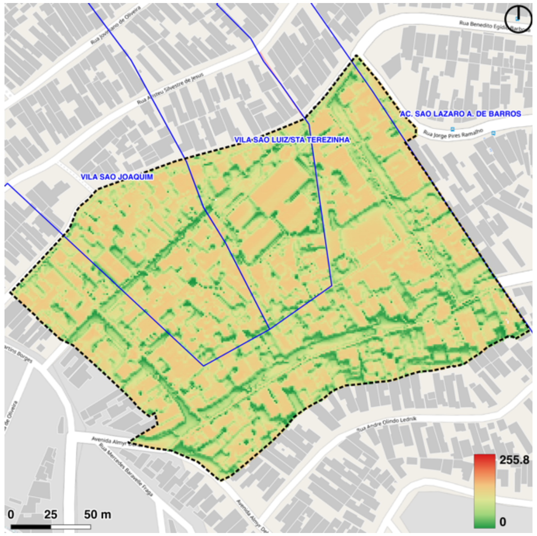


Figure 5-110 Brasilândia ArcGIS raster SEP.

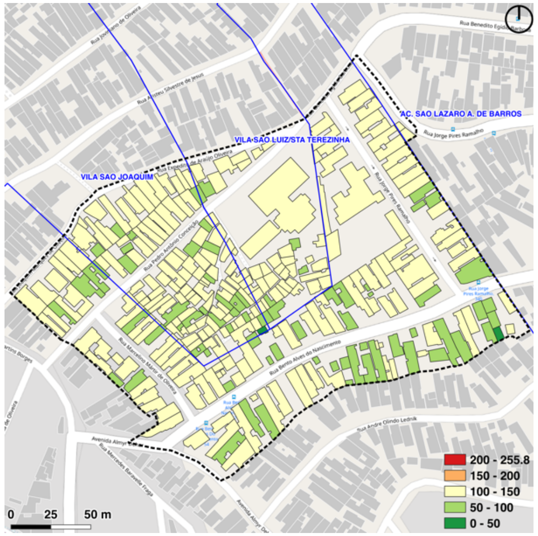


Figure 5-111 Brasilândia ArcGIS vector SEP.

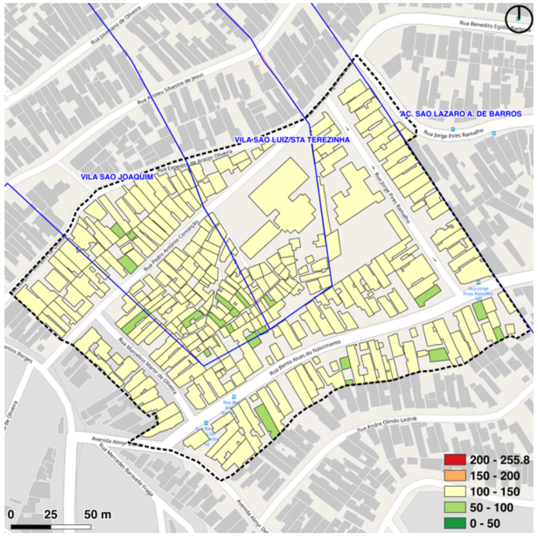


Figure 5-112 Brasilândia SimStadt SEP.

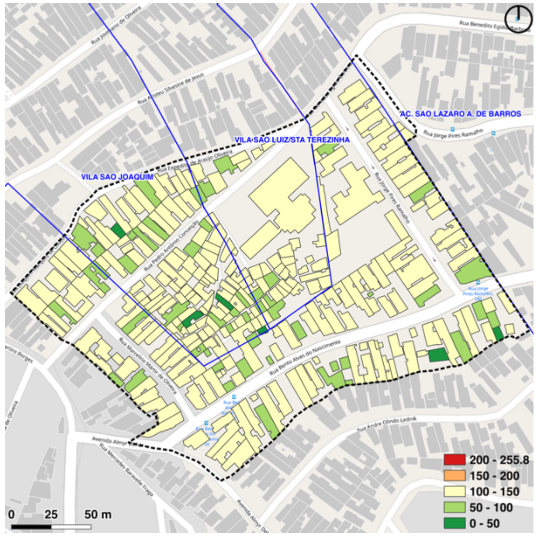


Figure 5-113 Brasilândia VCS SEP.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **OCTOBER** – kWh/m²

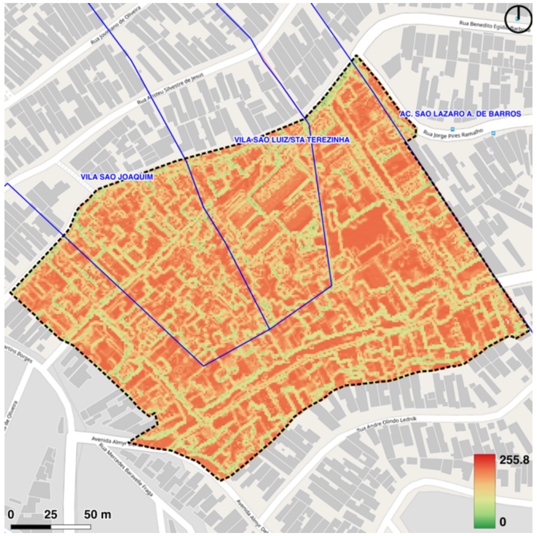


Figure 5-114 Brasilândia GRASS raster OCT.

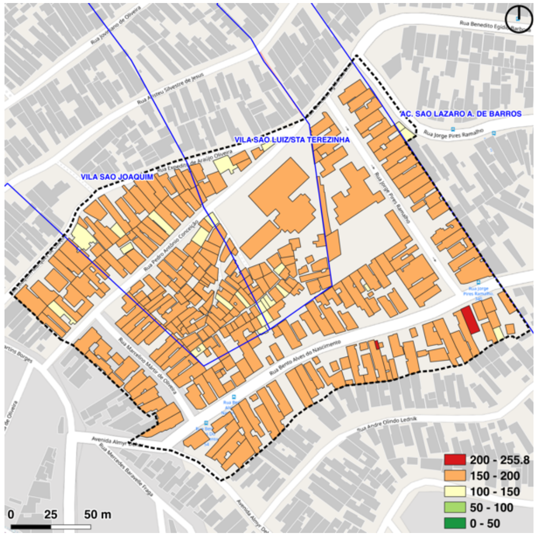


Figure 5-115 Brasilândia GRASS vector OCT.

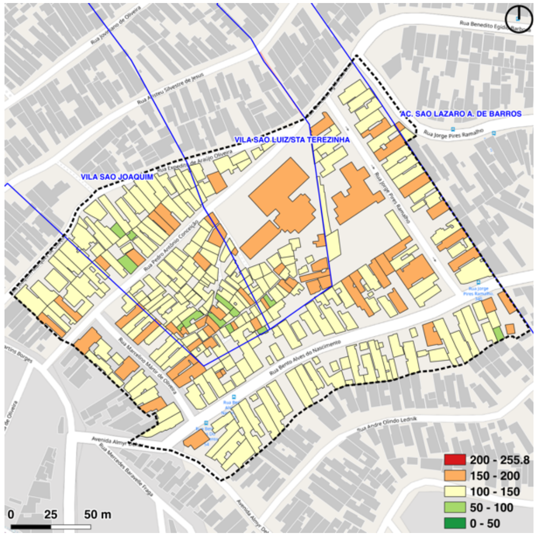


Figure 5-116 Brasilândia CitySim OCT.

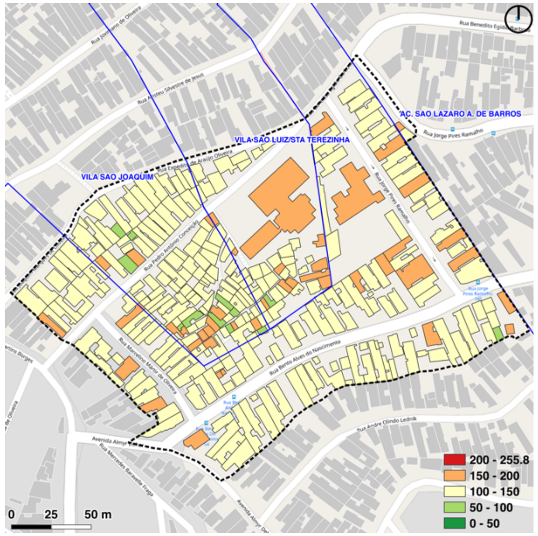


Figure 5-117 Brasilândia Ladybug OCT.

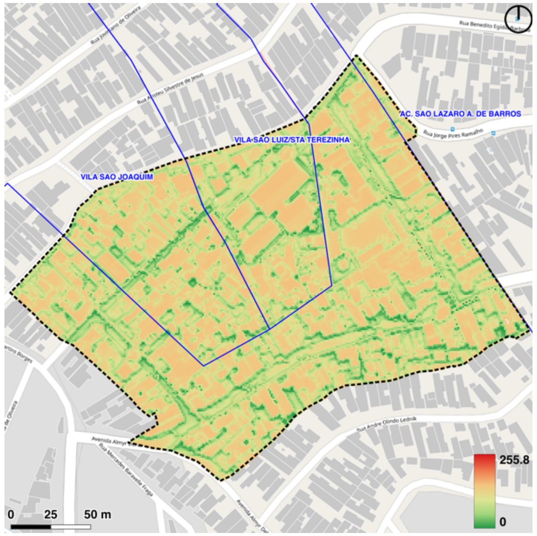


Figure 5-118 Brasilândia ArcGIS raster OCT.

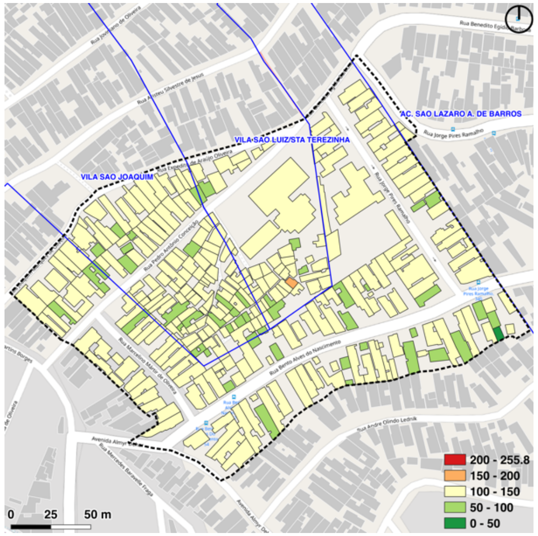


Figure 5-119 Brasilândia ArcGIS vector OCT.

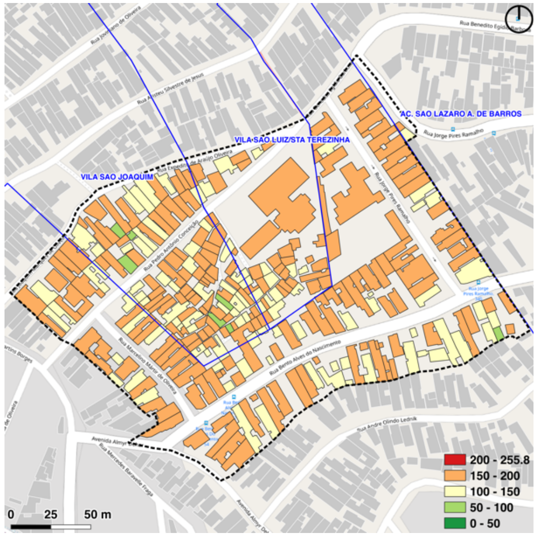


Figure 5-120 Brasilândia SimStadt OCT.

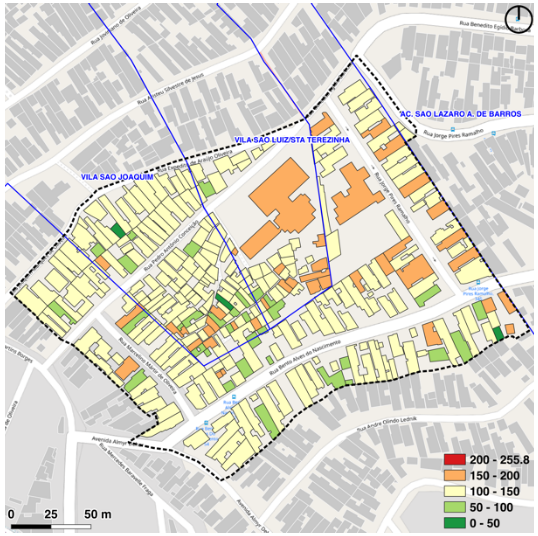


Figure 5-121 Brasilândia VCS OCT.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **NOVEMBER** – kWh/m²

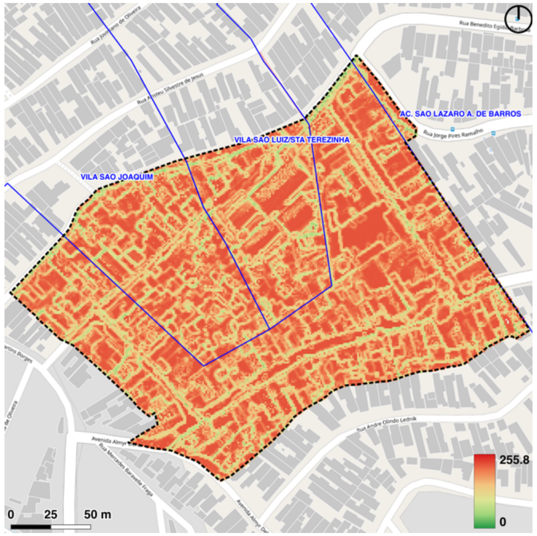


Figure 5-122 Brasilândia GRASS raster NOV.

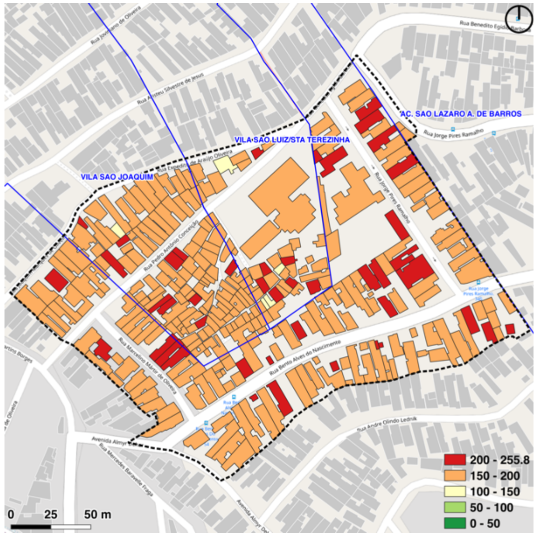


Figure 5-123 Brasilândia GRASS vector NOV.

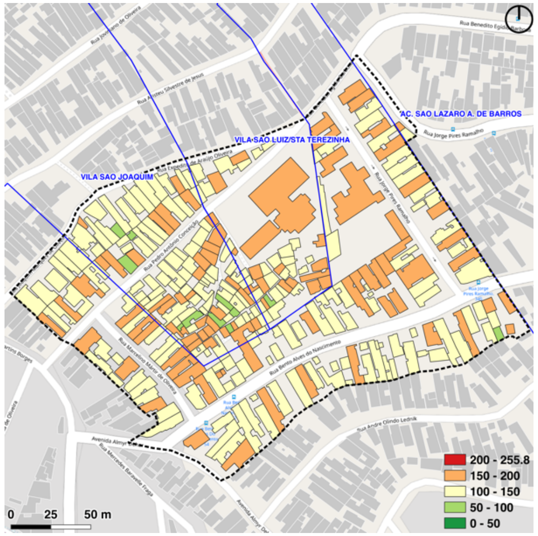


Figure 5-124 Brasilândia CitySim NOV.

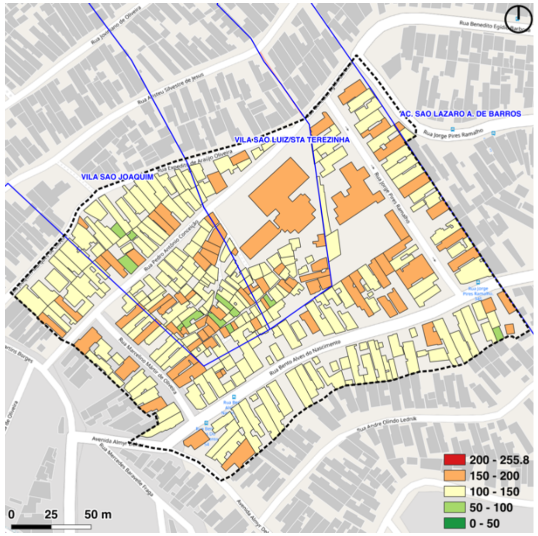


Figure 5-125 Brasilândia Ladybug NOV.

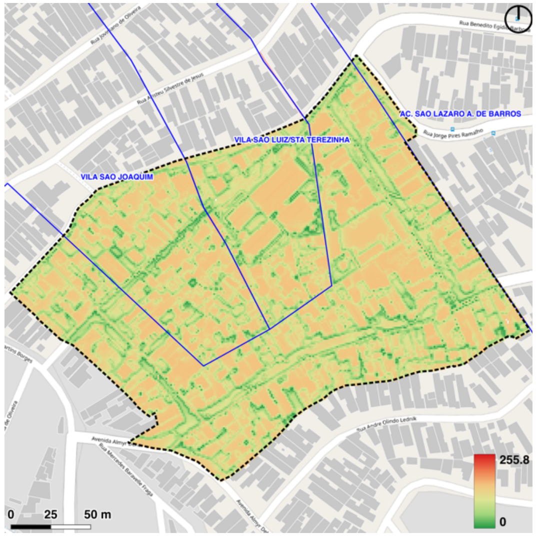


Figure 5-126 Brasilândia ArcGIS raster NOV.

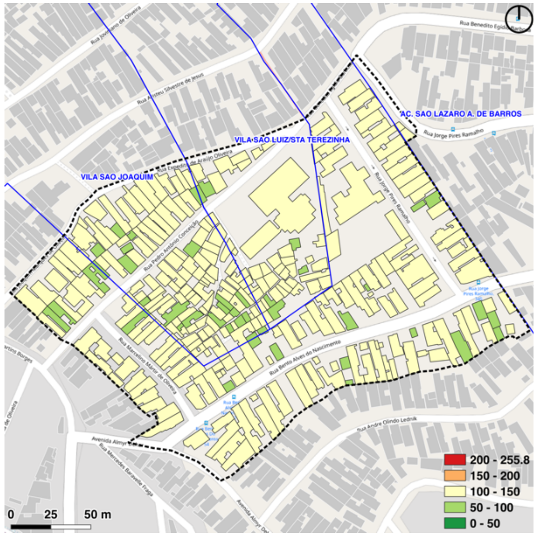


Figure 5-127 Brasilândia ArcGIS vector NOV.

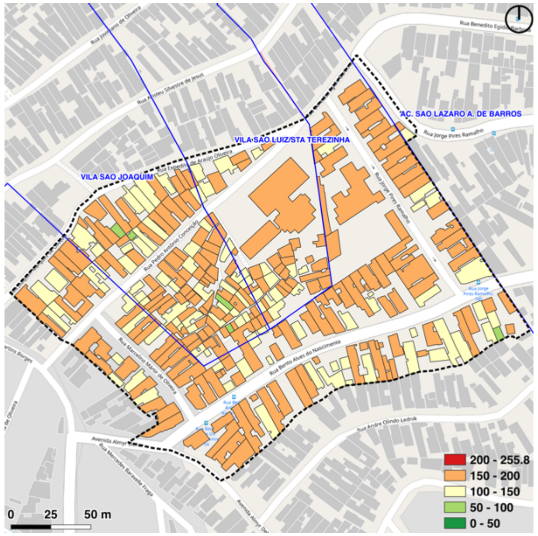


Figure 5-128 Brasilândia SimStadt NOV.

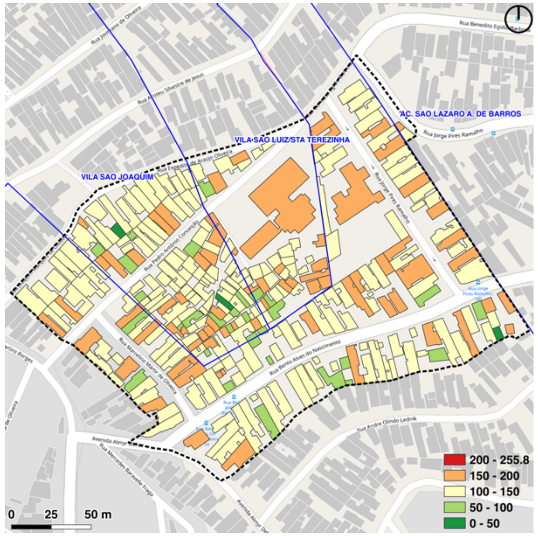


Figure 5-129 Brasilândia VCS NOV.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **DECEMBER** – kWh/m²

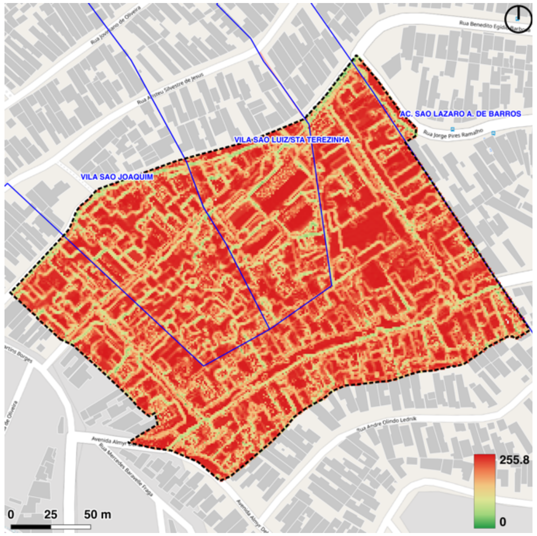


Figure 5-130 Brasilândia GRASS raster DEC.

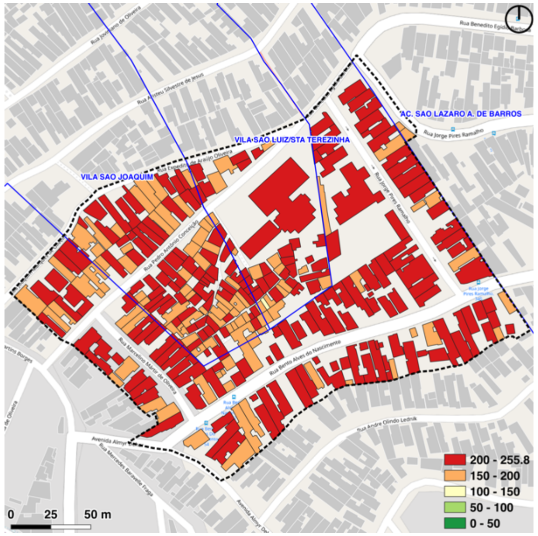


Figure 5-131 Brasilândia GRASS vector DEC.

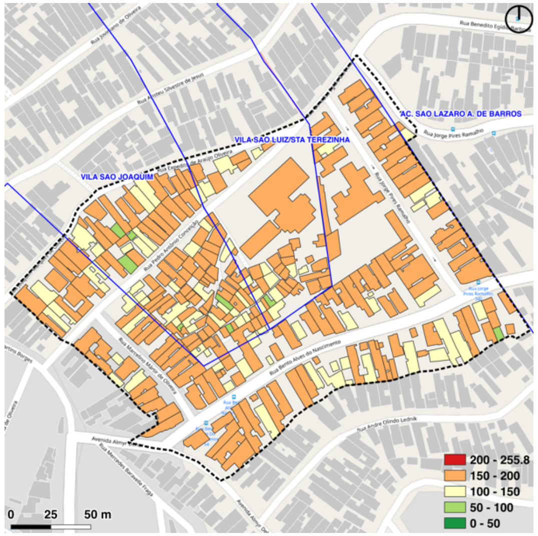


Figure 5-132 Brasilândia CitySim DEC.

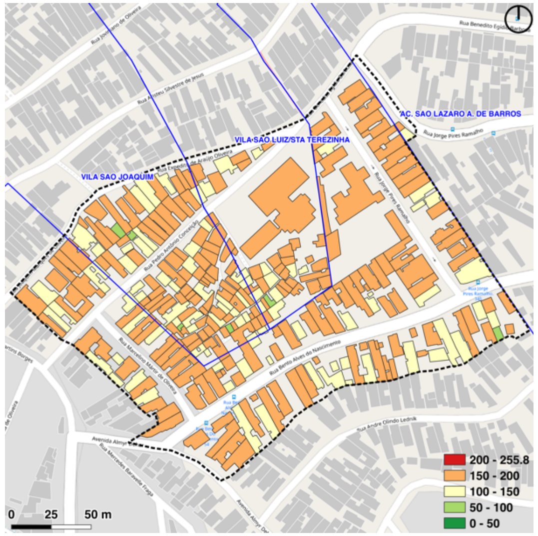


Figure 5-133 Brasilândia Ladybug DEC.

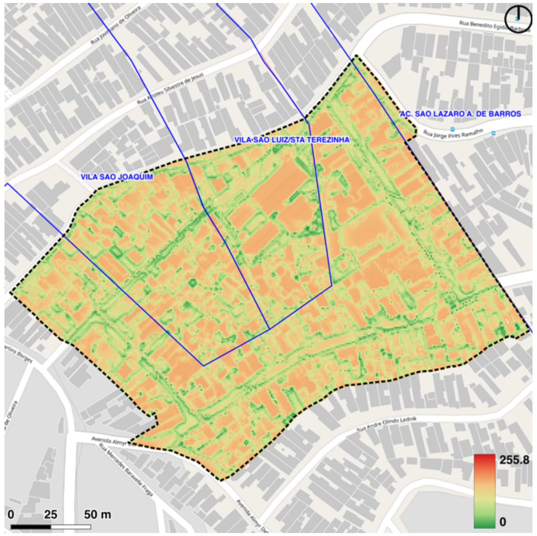


Figure 5-134 Brasilândia ArcGIS raster DEC.

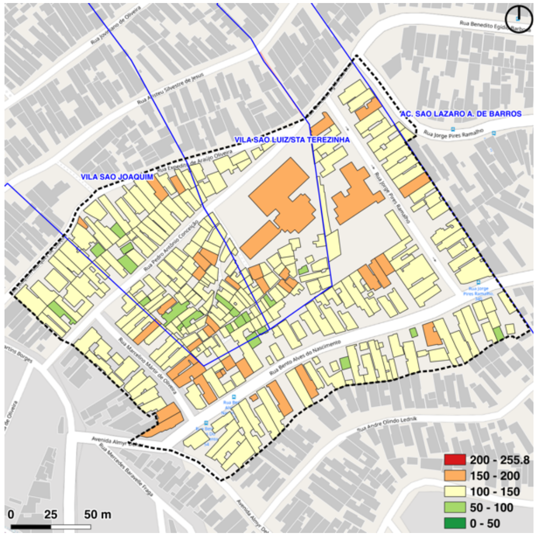


Figure 5-135 Brasilândia ArcGIS vector DEC.

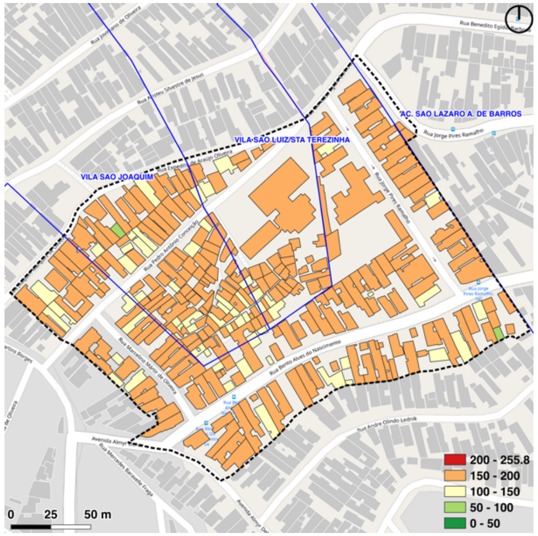


Figure 5-136 Brasilândia SimStadt DEC.

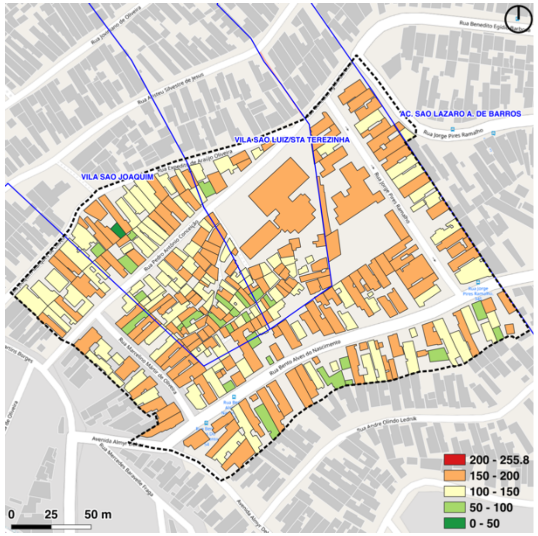


Figure 5-137 Brasilândia VCS DEC.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **GRASS raster** – kWh/m²

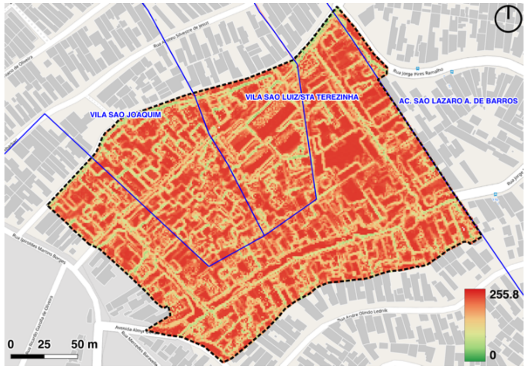


Figure 5-138 Brasilândia GRASS raster JAN.

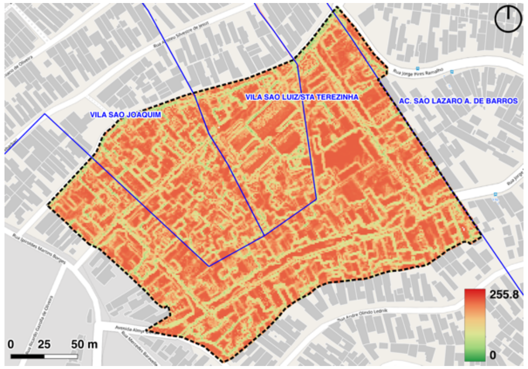


Figure 5-139 Brasilândia GRASS raster FEB.

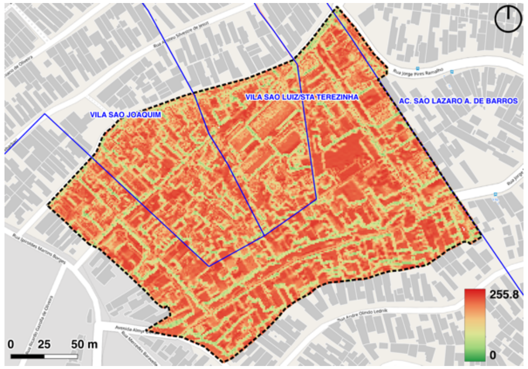


Figure 5-140 Brasilândia GRASS raster MAR.

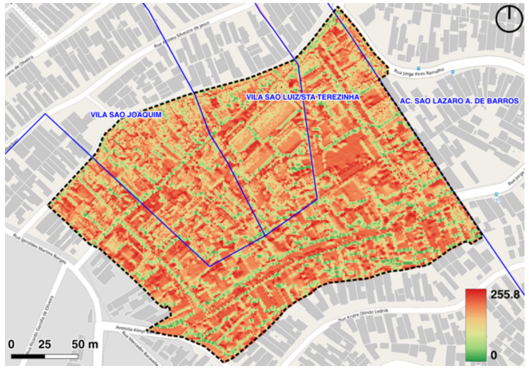


Figure 5-141 Brasilândia GRASS raster APR.

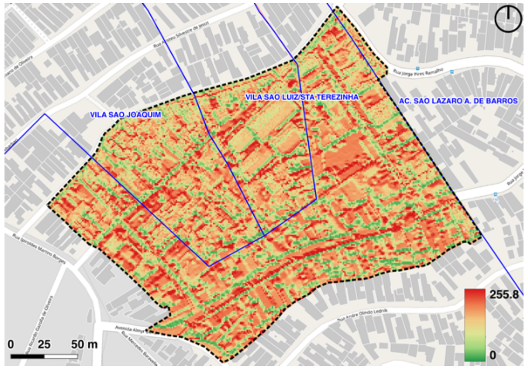


Figure 5-142 Brasilândia GRASS raster MAY.

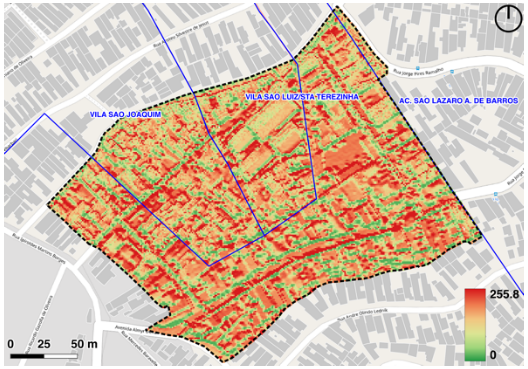


Figure 5-143 Brasilândia GRASS raster JUN.

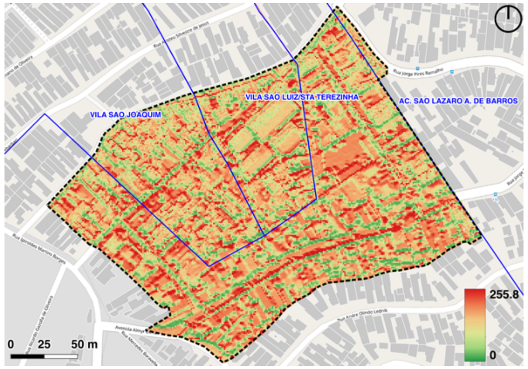


Figure 5-144 Brasilândia GRASS raster JUL.

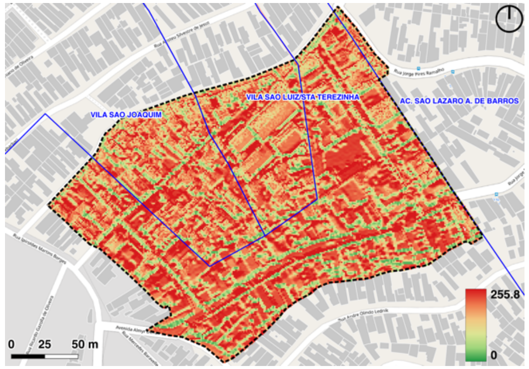


Figure 5-145 Brasilândia GRASS raster AUG.

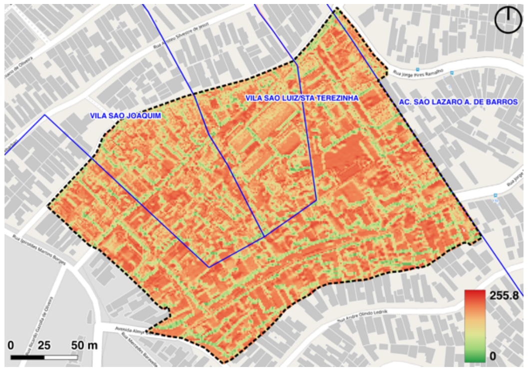


Figure 5-146 Brasilândia GRASS raster SEP.

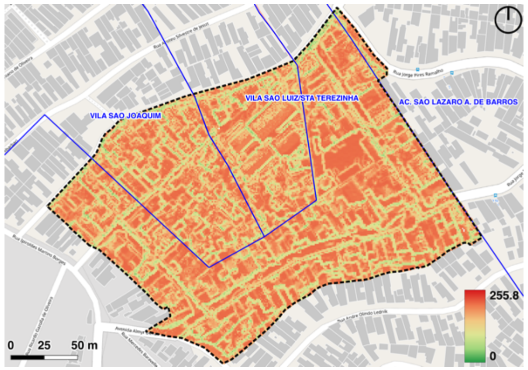


Figure 5-147 Brasilândia GRASS raster OCT.

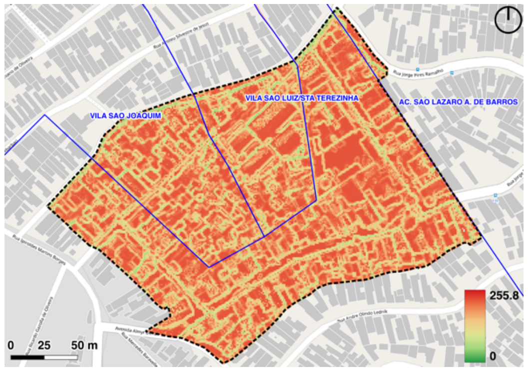


Figure 5-148 Brasilândia GRASS raster NOV.

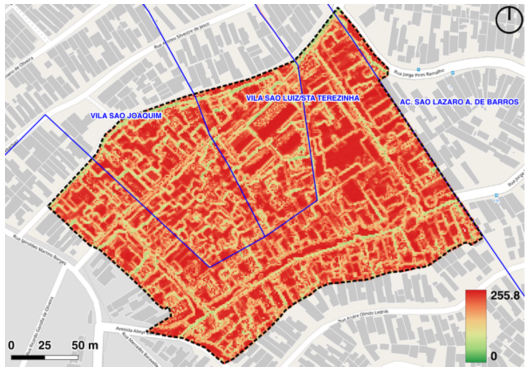


Figure 5-149 Brasilândia GRASS raster DEC.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **GRASS vector** – kWh/m²

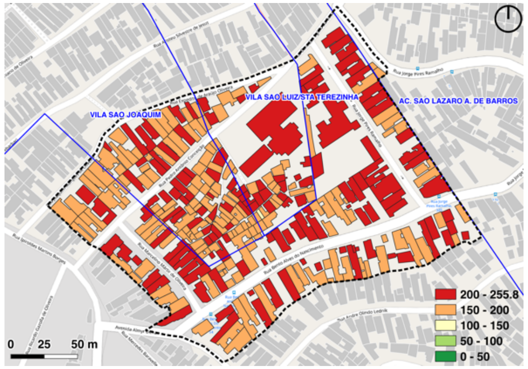


Figure 5-150 Brasilândia GRASS vector JAN.

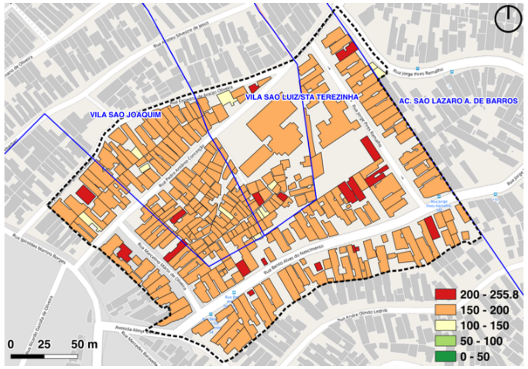


Figure 5-151 Brasilândia GRASS vector FEB.

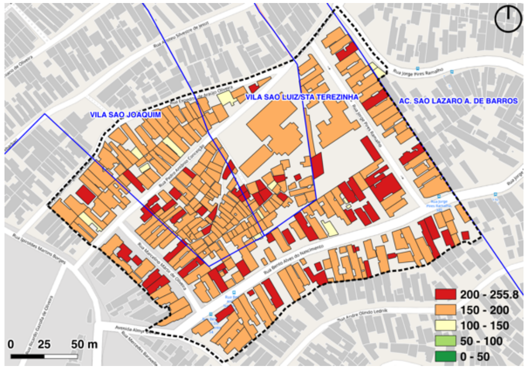


Figure 5-152 Brasilândia GRASS vector MAR.

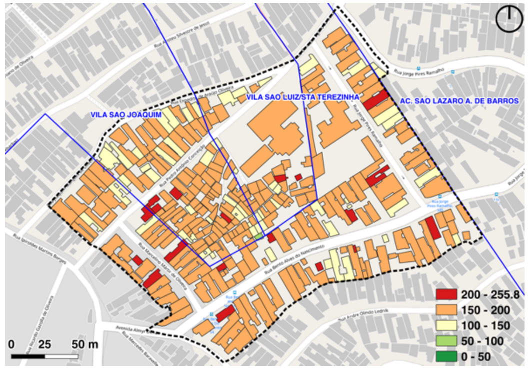


Figure 5-153 Brasilândia GRASS vector APR.

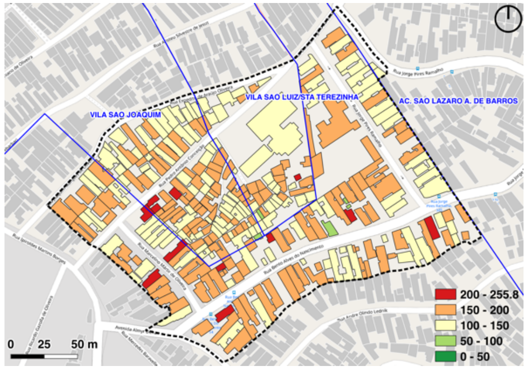


Figure 5-154 Brasilândia GRASS vector MAY.

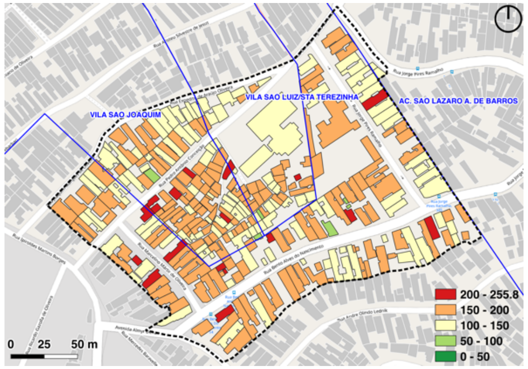


Figure 5-155 Brasilândia GRASS vector JUN.

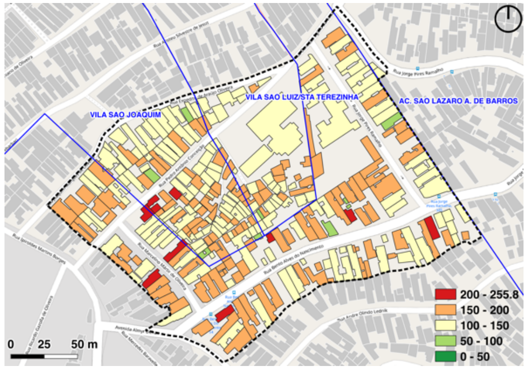


Figure 5-156 Brasilândia GRASS vector JUL.

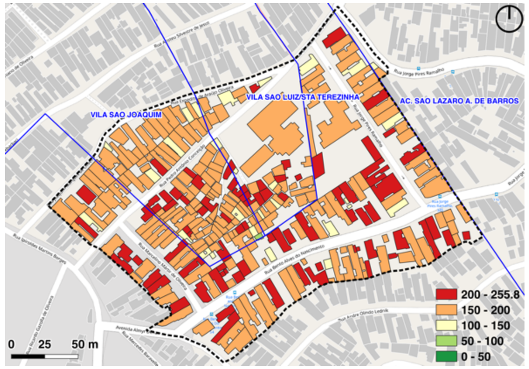


Figure 5-157 Brasilândia GRASS vector AUG.

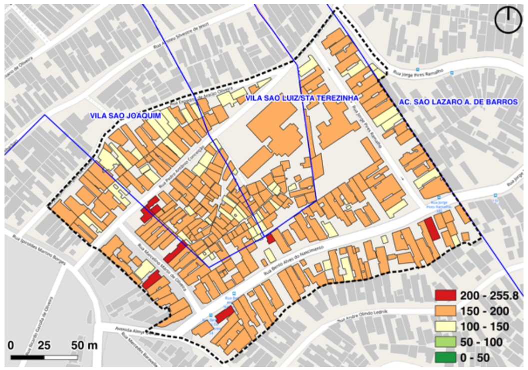


Figure 5-158 Brasilândia GRASS vector SEP.

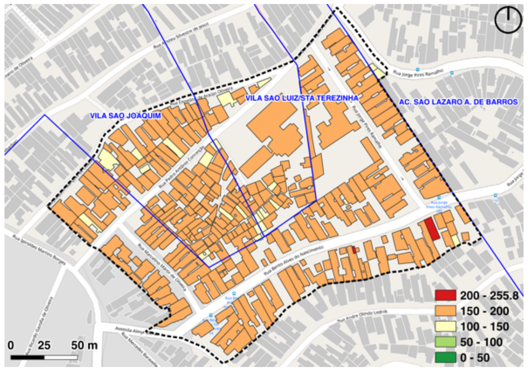


Figure 5-159 Brasilândia GRASS vector OCT.

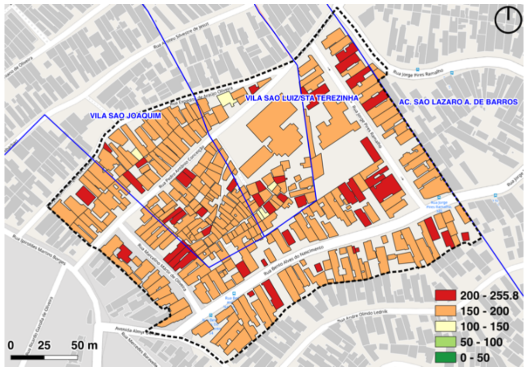


Figure 5-160 Brasilândia GRASS vector NOV.

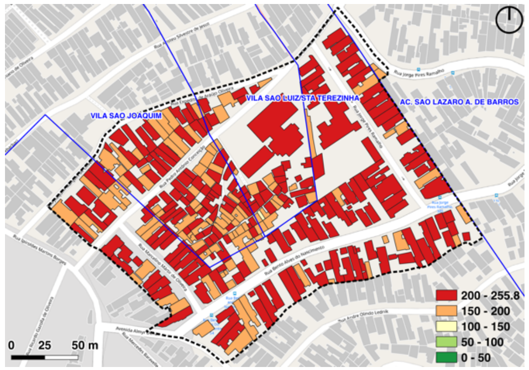


Figure 5-161 Brasilândia GRASS vector DEC.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **ArcGIS raster** – kWh/m²

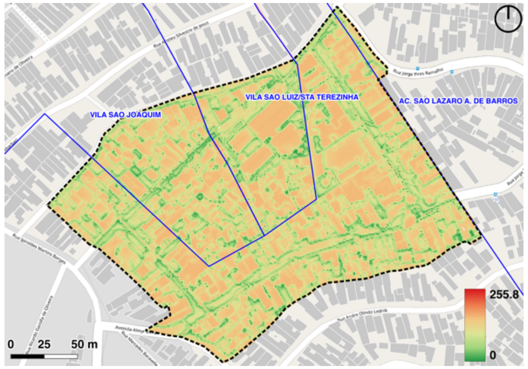


Figure 5-162 Brasilândia ArcGIS raster JAN.

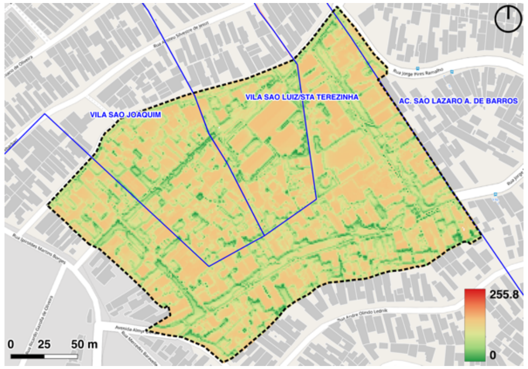


Figure 5-163 Brasilândia ArcGIS raster FEB.

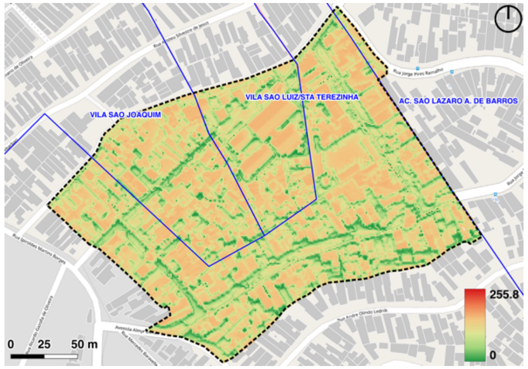


Figure 5-164 Brasilândia ArcGIS raster MAR.

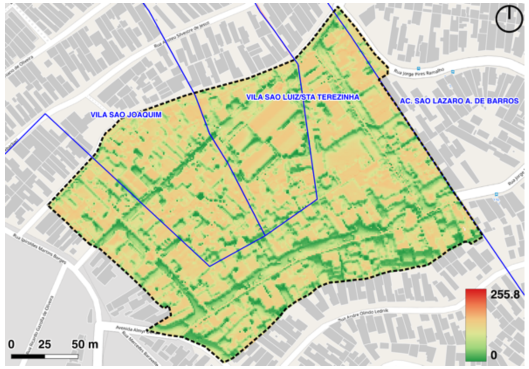


Figure 5-165 Brasilândia ArcGIS raster APR.

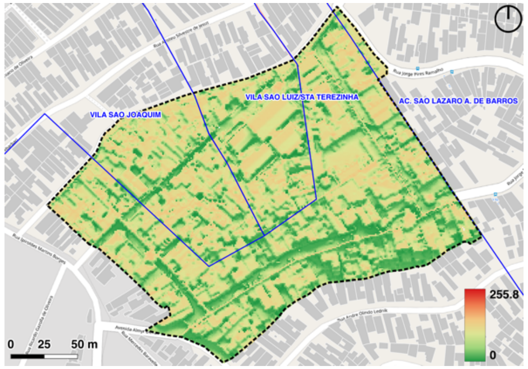


Figure 5-166 Brasilândia ArcGIS raster MAY.

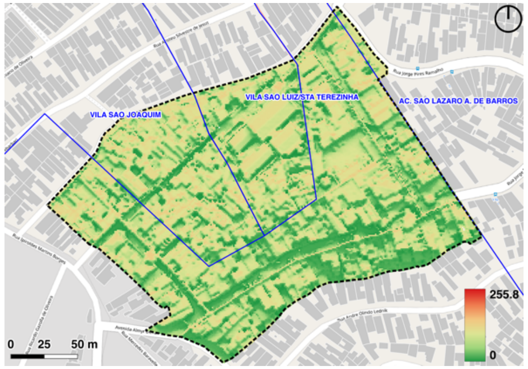


Figure 5-167 Brasilândia ArcGIS raster JUN.

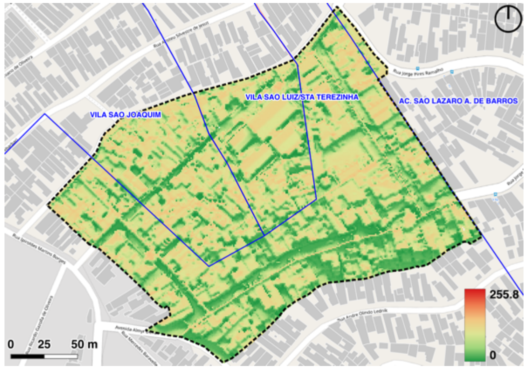


Figure 5-168 Brasilândia ArcGIS raster JUL.

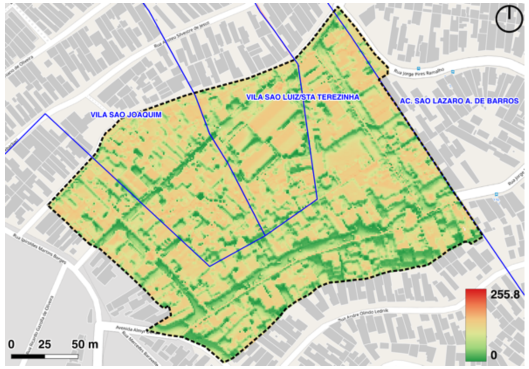


Figure 5-169 Brasilândia ArcGIS raster AUG.

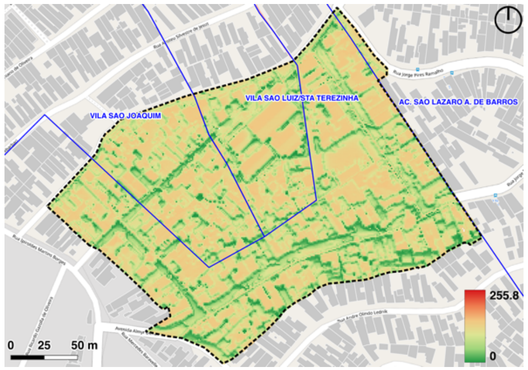


Figure 5-170 Brasilândia ArcGIS raster SEP.

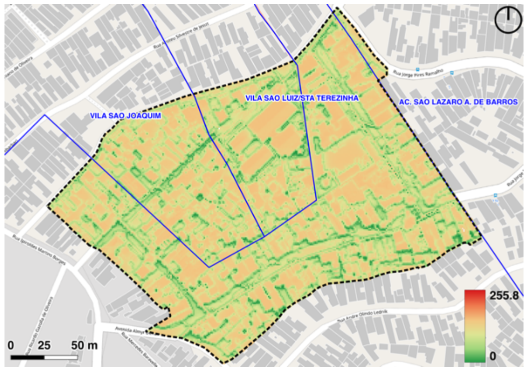


Figure 5-171 Brasilândia ArcGIS raster OCT.

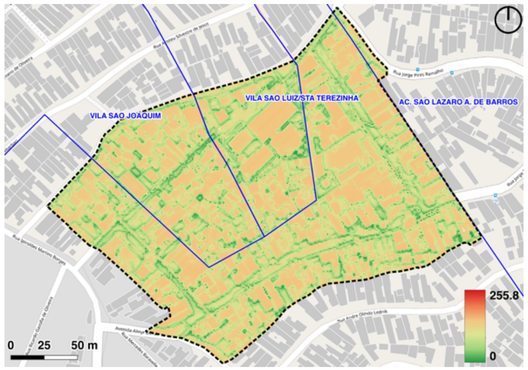


Figure 5-172 Brasilândia ArcGIS raster NOV.

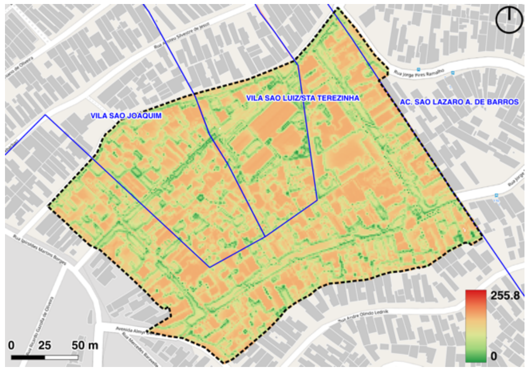


Figure 5-173 Brasilândia ArcGIS raster DEC.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **ArcGIS vector** – kWh/m²

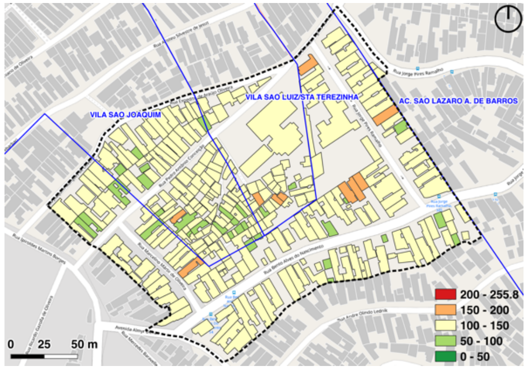


Figure 5-174 Brasilândia ArcGIS vector JAN.

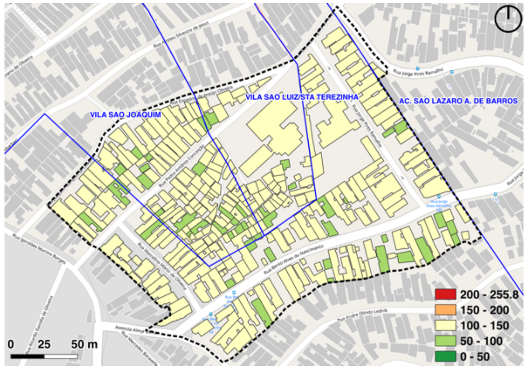


Figure 5-175 Brasilândia ArcGIS vector FEB.

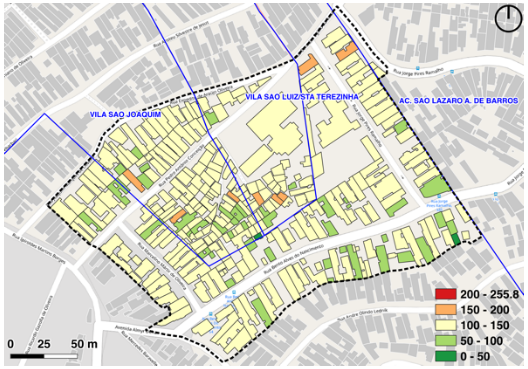


Figure 5-176 Brasilândia ArcGIS vector MAR.

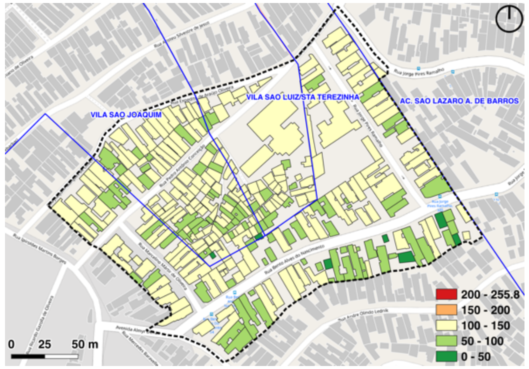


Figure 5-177 Brasilândia ArcGIS vector APR.

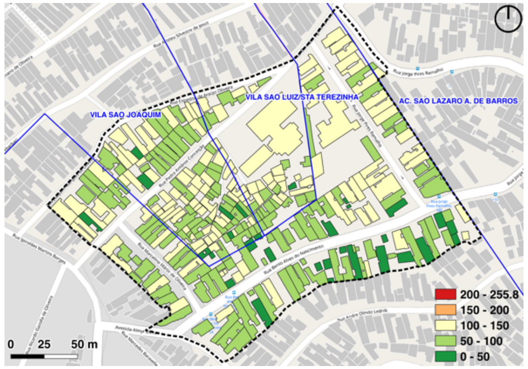


Figure 5-178 Brasilândia ArcGIS vector MAY.

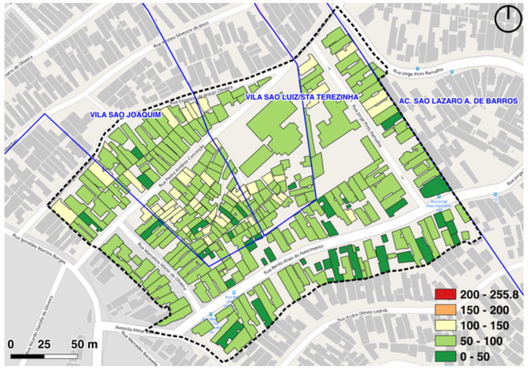


Figure 5-179 Brasilândia ArcGIS vector JUN.

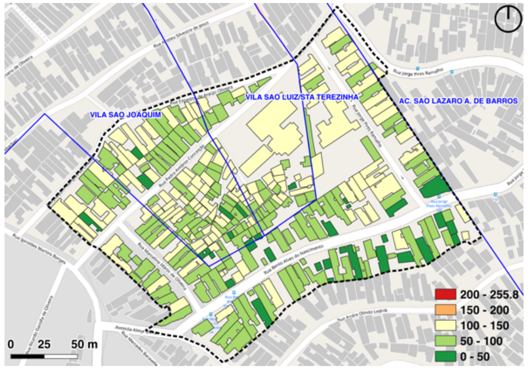


Figure 5-180 Brasilândia ArcGIS vector JUL.

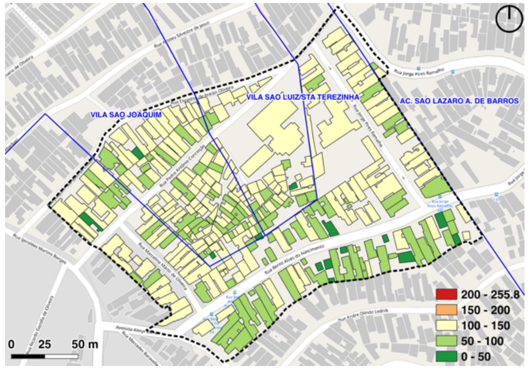


Figure 5-181 Brasilândia ArcGIS vector AUG.

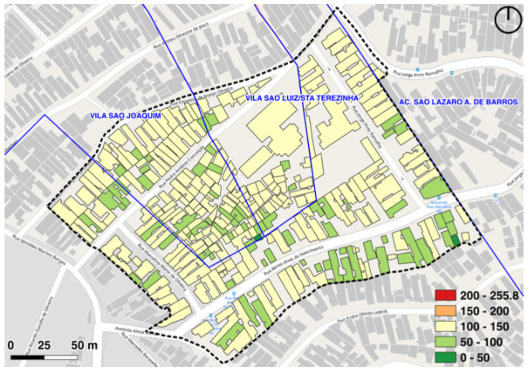


Figure 5-182 Brasilândia ArcGIS vector SEP.

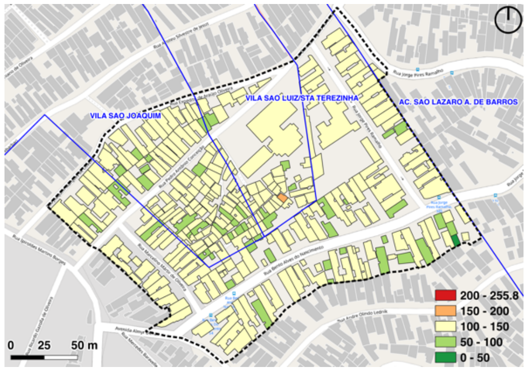


Figure 5-183 Brasilândia ArcGIS vector OCT.

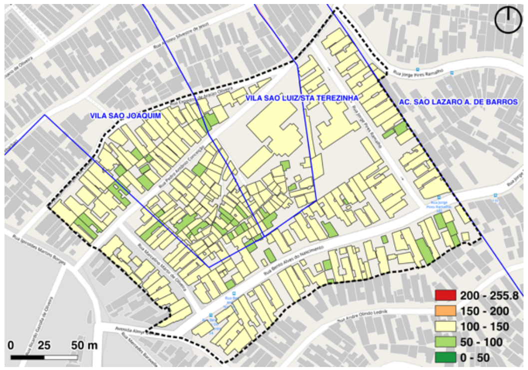


Figure 5-184 Brasilândia ArcGIS vector NOV.

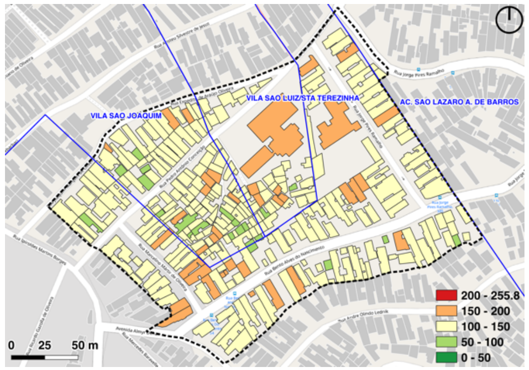


Figure 5-185 Brasilândia ArcGIS vector DEC.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **CitySim** – kWh/m²

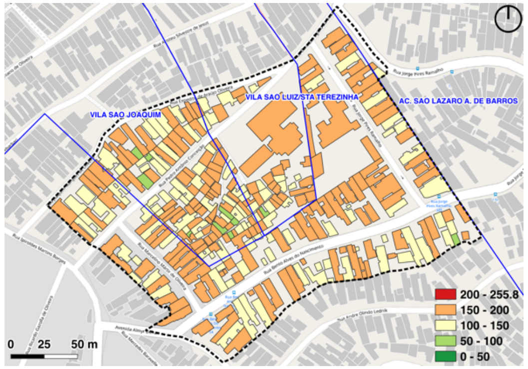


Figure 5-186 Brasilândia CitySim JAN.

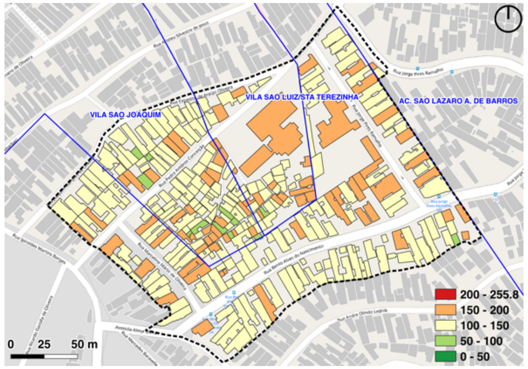


Figure 5-187 Brasilândia CitySim FEB.

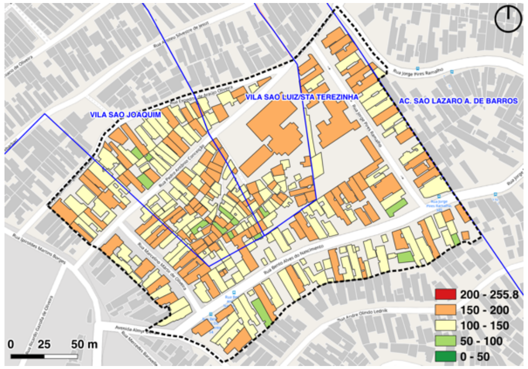


Figure 5-188 Brasilândia CitySim MAR.

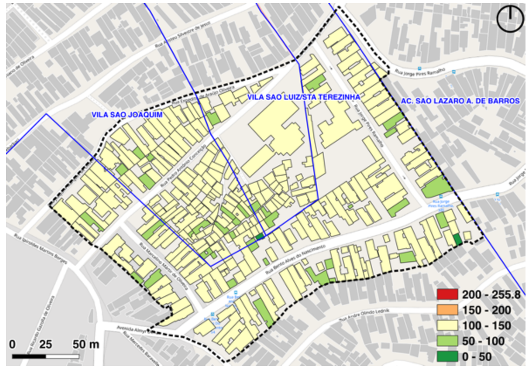


Figure 5-189 Brasilândia CitySim APR.

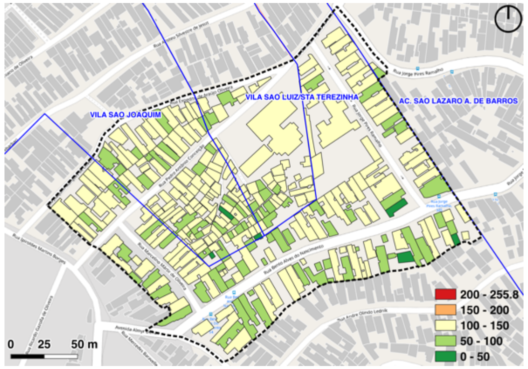


Figure 5-190 Brasilândia CitySim MAY.

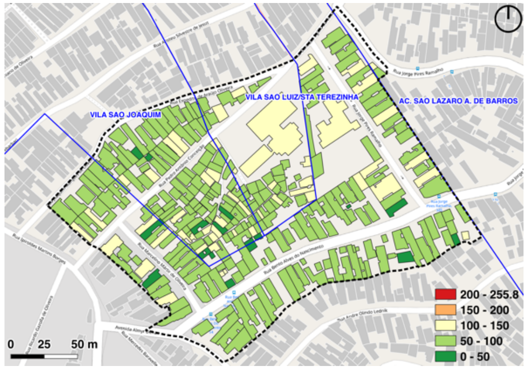


Figure 5-191 Brasilândia CitySim JUN.

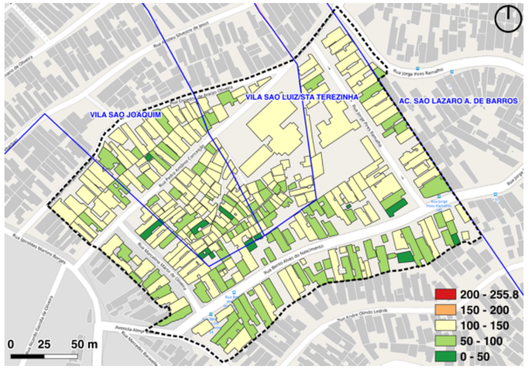


Figure 5-192 Brasilândia CitySim JUL.

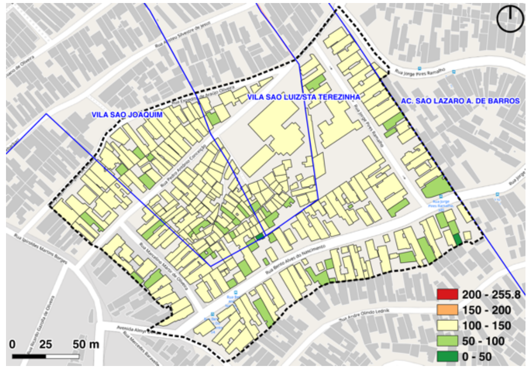


Figure 5-193 Brasilândia CitySim AUG.

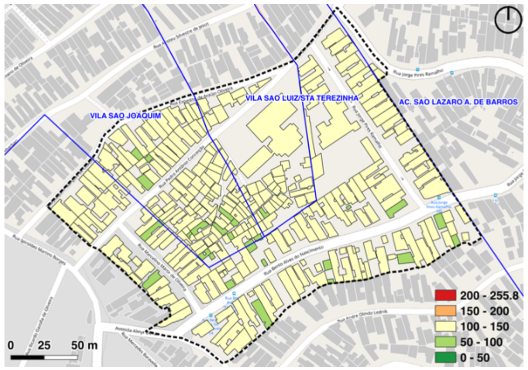


Figure 5-194 Brasilândia CitySim SEP.

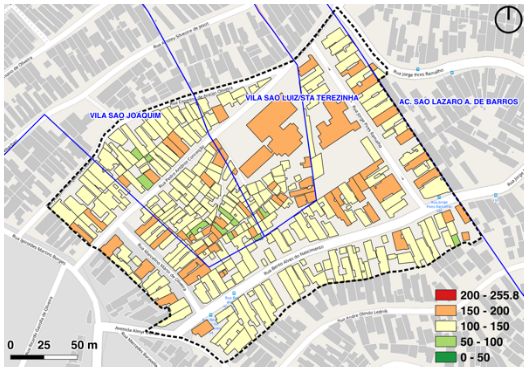


Figure 5-195 Brasilândia CitySim OCT.

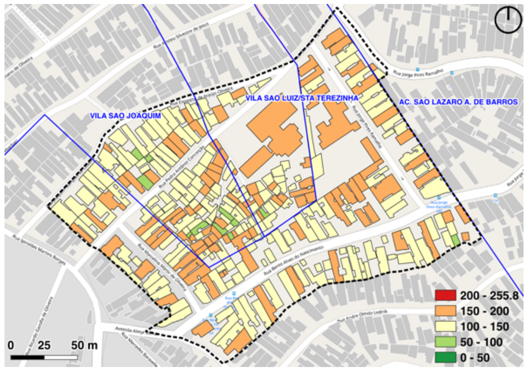


Figure 5-196 Brasilândia CitySim NOV.

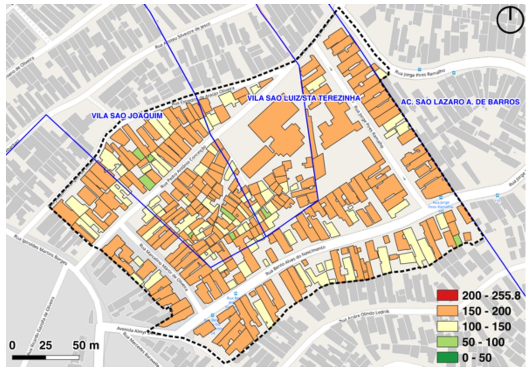


Figure 5-197 Brasilândia CitySim DEC.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **SimStadt** – kWh/m²

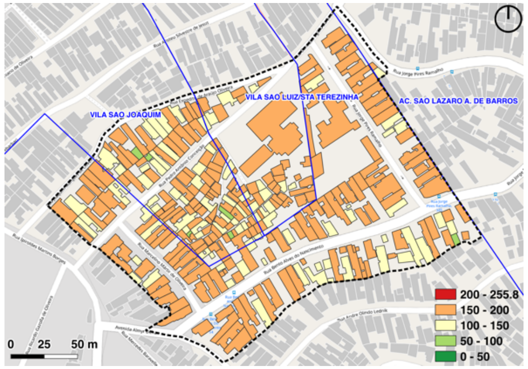


Figure 5-198 Brasilândia SimStadt JAN.

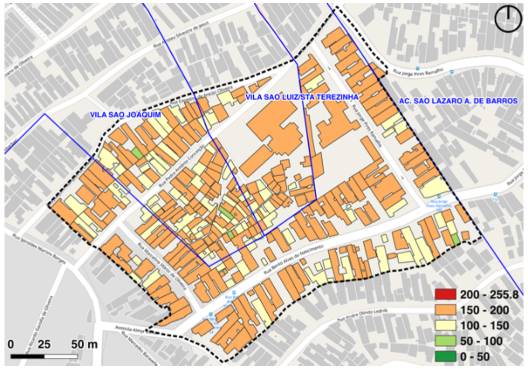


Figure 5-199 Brasilândia SimStadt FEB.

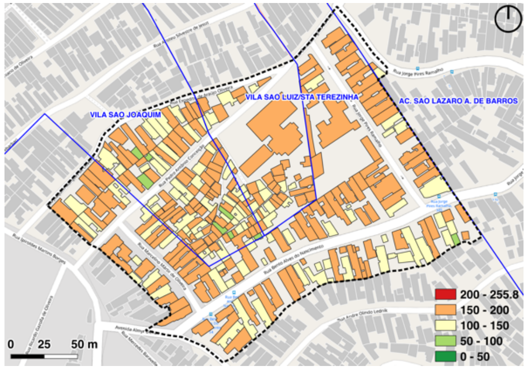


Figure 5-200 Brasilândia SimStadt MAR.

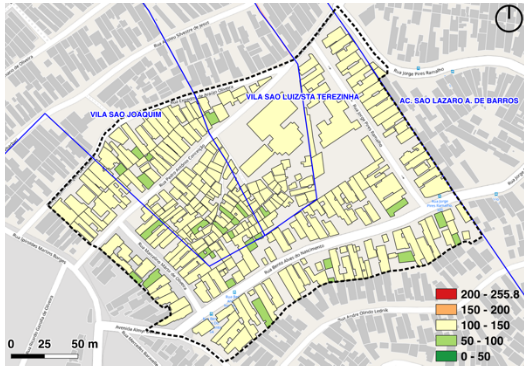


Figure 5-201 Brasilândia SimStadt APR.

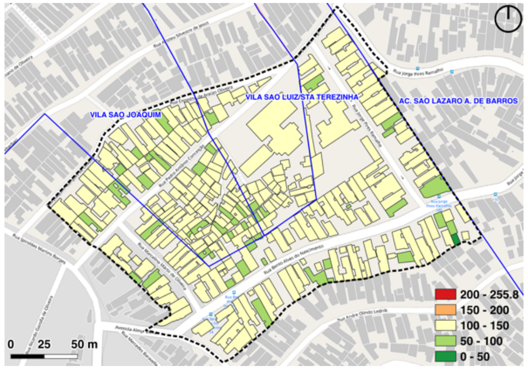


Figure 5-202 Brasilândia SimStadt MAY.

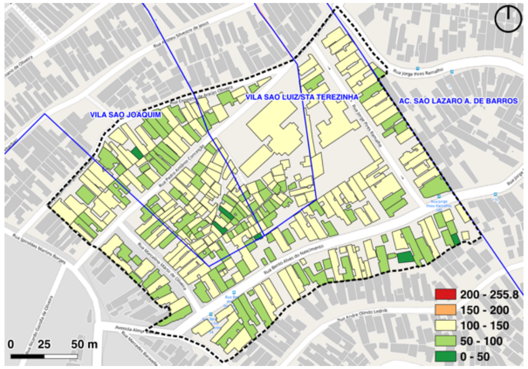


Figure 5-203 Brasilândia SimStadt JUN.

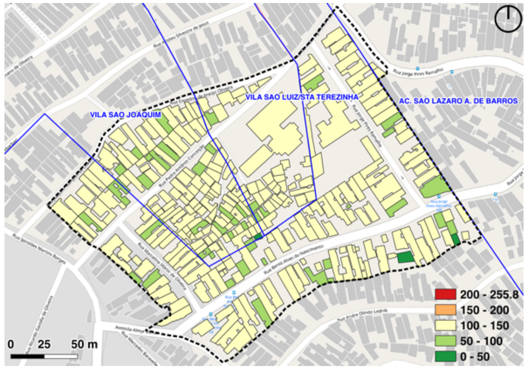


Figure 5-204 Brasilândia SimStadt JUL.

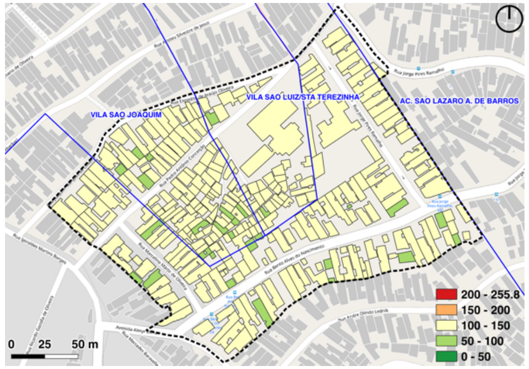


Figure 5-205 Brasilândia SimStadt AUG.

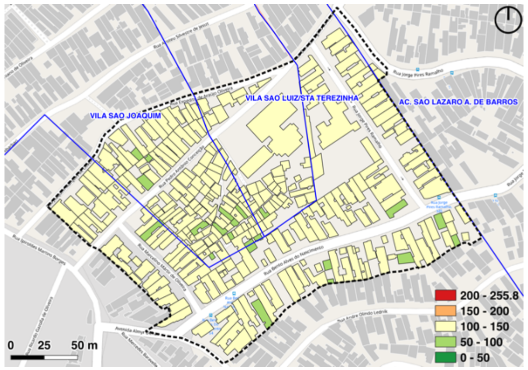


Figure 5-206 Brasilândia SimStadt SEP.

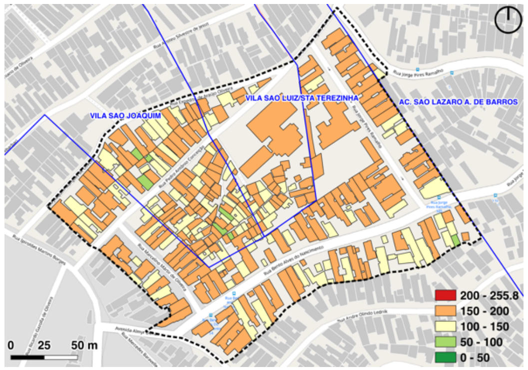


Figure 5-207 Brasilândia SimStadt OCT.

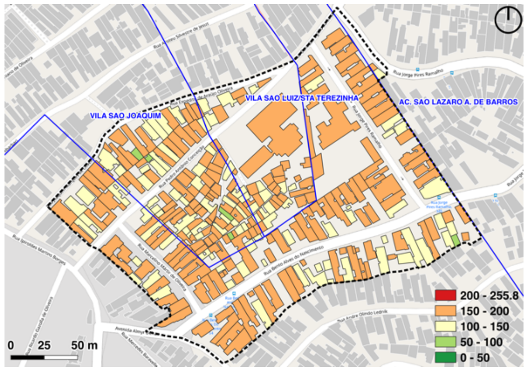


Figure 5-208 Brasilândia SimStadt NOV.

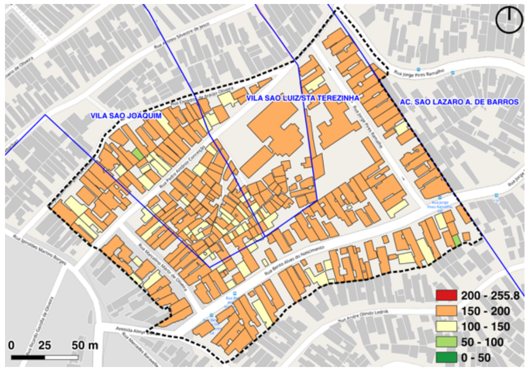


Figure 5-209 Brasilândia SimStadt DEC.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **Ladybug** – kWh/m²

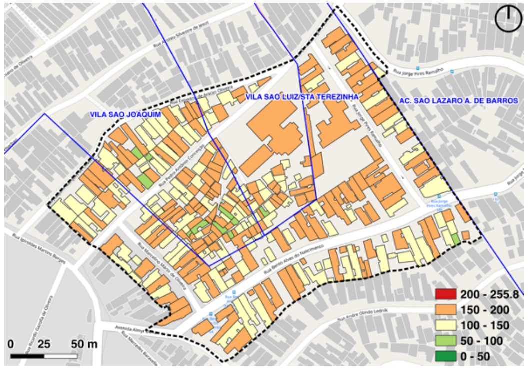


Figure 5-210 Brasilândia Ladybug JAN.

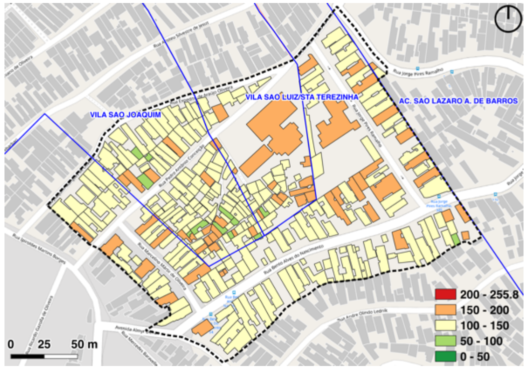


Figure 5-211 Brasilândia Ladybug FEB.

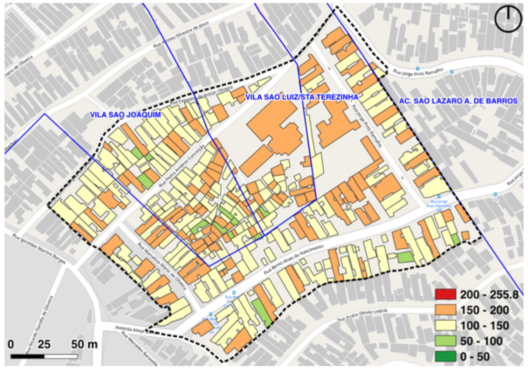


Figure 5-212 Brasilândia Ladybug MAR.

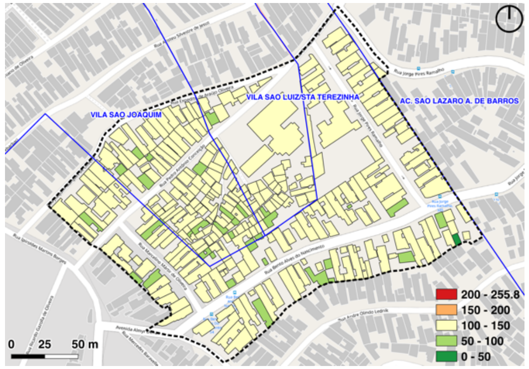


Figure 5-213 Brasilândia Ladybug APR.

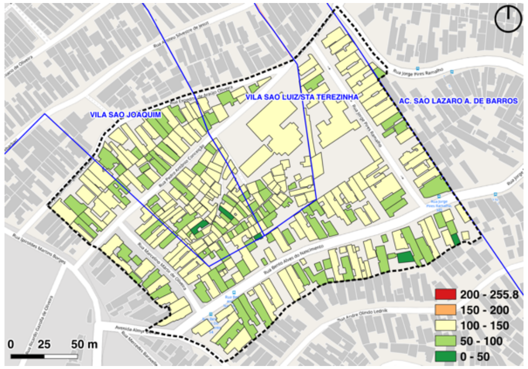


Figure 5-214 Brasilândia Ladybug MAY.

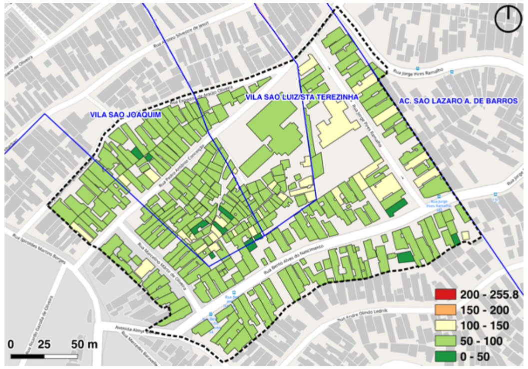


Figure 5-215 Brasilândia Ladybug JUN.

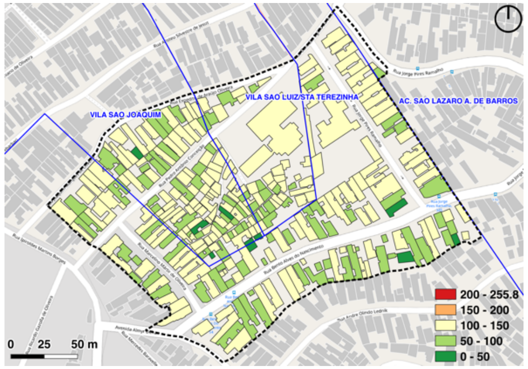


Figure 5-216 Brasilândia Ladybug JUL.

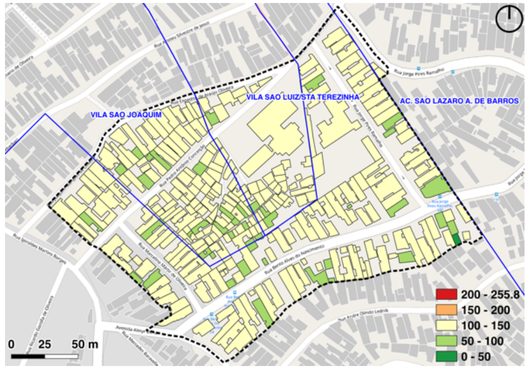


Figure 5-217 Brasilândia Ladybug AUG.

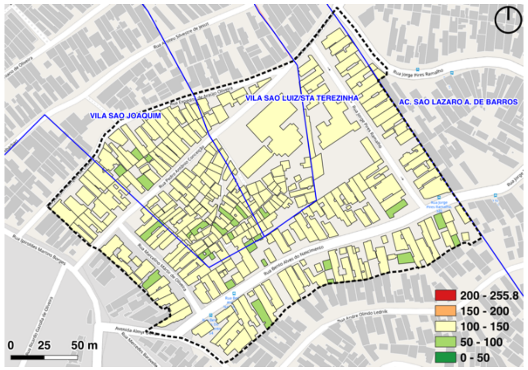


Figure 5-218 Brasilândia Ladybug SEP.

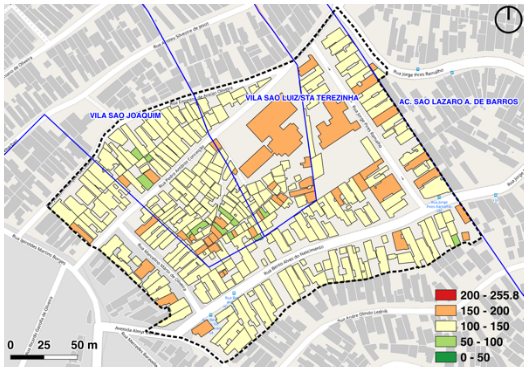


Figure 5-219 Brasilândia Ladybug OCT.

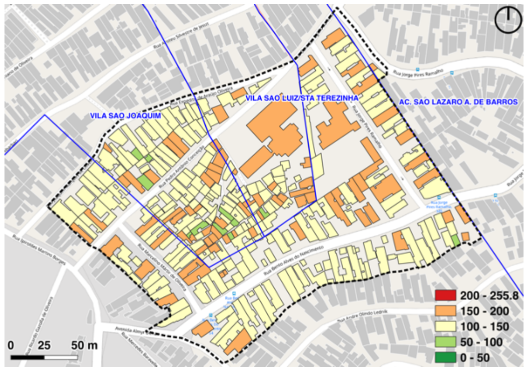


Figure 5-220 Brasilândia Ladybug NOV.

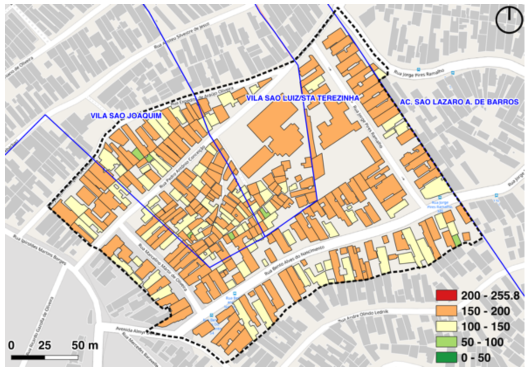


Figure 5-221 Brasilândia Ladybug DEC.

Brasilândia City Block – Corrected Monthly Global Solar Irradiation – **Virtual City Systems** – kWh/m²

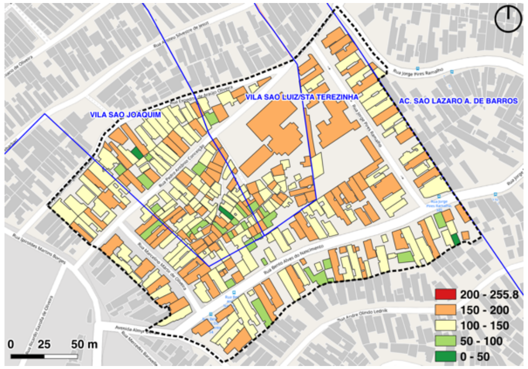


Figure 5-222 Brasilândia VCS JAN.

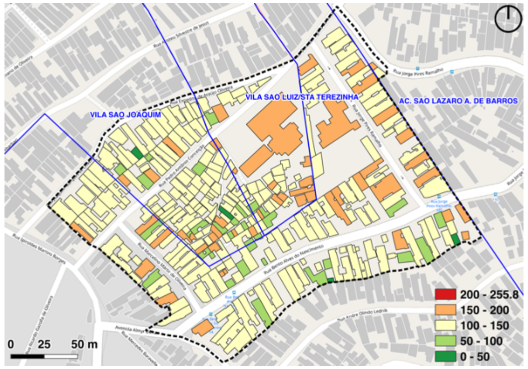


Figure 5-223 Brasilândia VCS FEB.

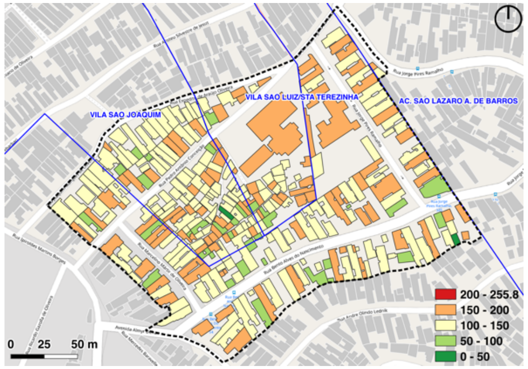


Figure 5-224 Brasilândia VCS MAR.

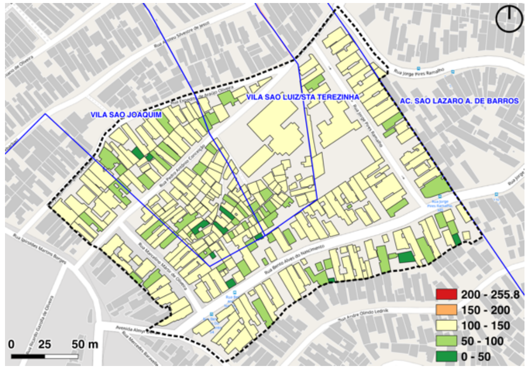


Figure 5-225 Brasilândia VCS APR.

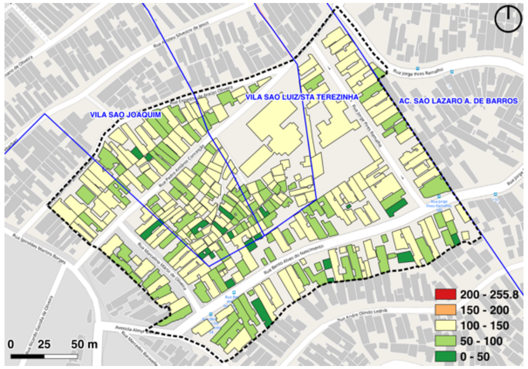


Figure 5-226 Brasilândia VCS MAY.

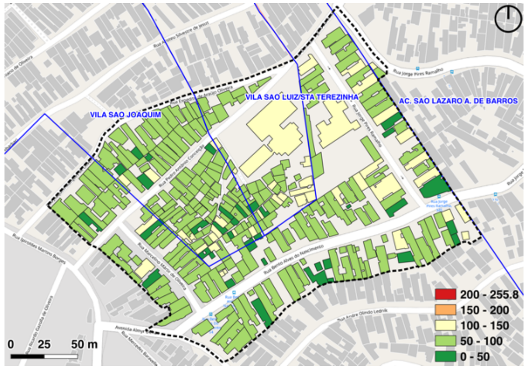


Figure 5-227 Brasilândia VCS JUN.

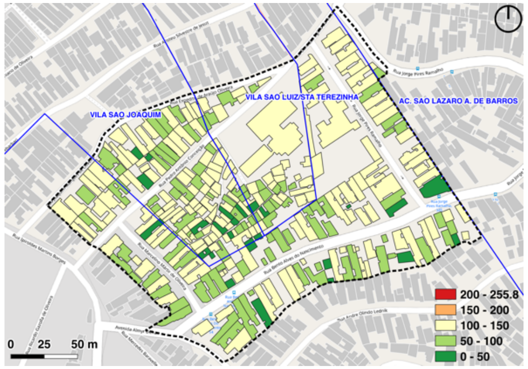


Figure 5-228 Brasilândia VCS JUL.

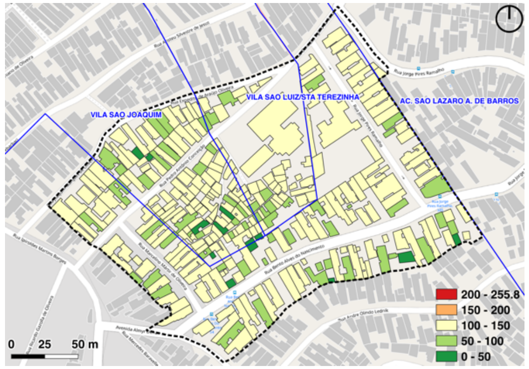


Figure 5-229 Brasilândia VCS AUG.

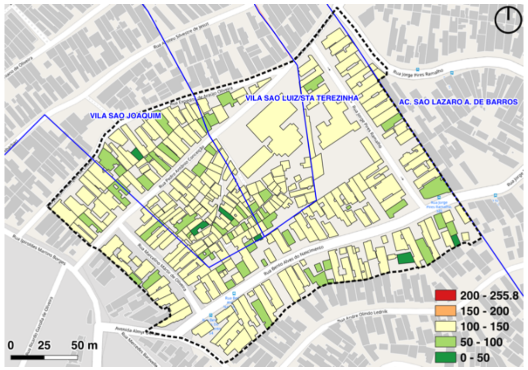


Figure 5-230 Brasilândia VCS SEP.

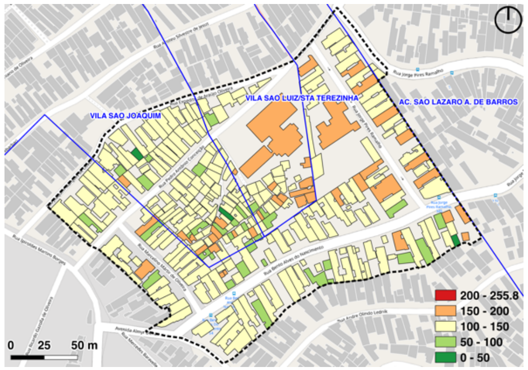


Figure 5-231 Brasilândia VCS OCT.

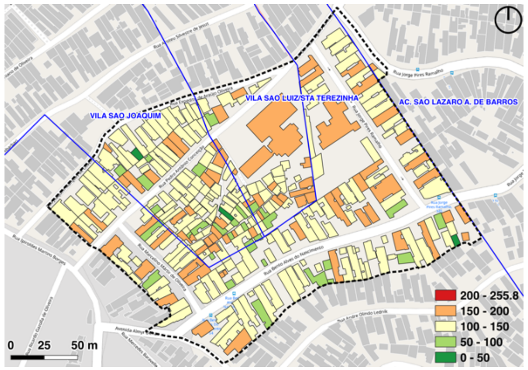


Figure 5-232 Brasilândia VCS NOV.

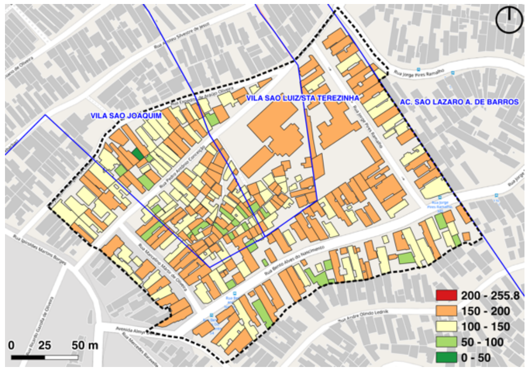


Figure 5-233 Brasilândia VCS DEC.

In a broad analysis of the spatial-temporal distribution of global solar irradiation across the building roofs in Brasilândia throughout the months of the year, some considerations are worth of highlight. Analogously to the analysis conducted on an annual basis, the following table and graph resume the monthly weighted averages of corrected global solar irradiation.

Scenario Brasilândia - Monthly Weighted Averages of Corrected Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	198.18	180.12	184.34	168.10	153.31	157.46	148.27	186.10	166.93	172.84	186.65	209.76
ARC GIS	128.28	122.40	124.98	108.18	92.32	80.63	90.76	106.18	112.93	122.11	122.80	134.42
CITYSIM	149.65	141.89	143.66	122.95	103.83	89.58	101.12	119.95	129.02	141.08	142.88	157.18
SIMSTADT	154.51	155.56	151.46	130.65	118.71	99.52	122.98	132.17	134.79	151.47	151.88	163.44
LADY BUG	147.55	140.53	142.42	122.04	103.42	89.98	101.22	119.86	128.21	139.72	141.57	155.71
VCS SOLAR	140.35	133.48	135.76	116.91	99.10	85.84	97.12	114.62	122.71	133.10	134.44	147.03

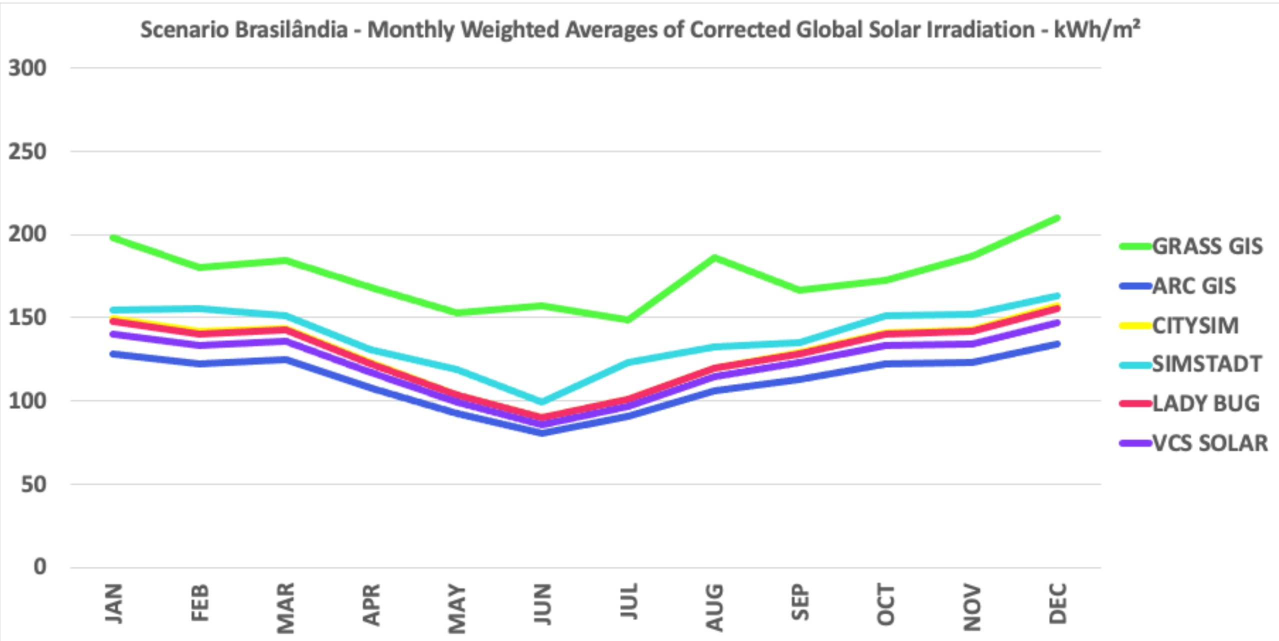


Figure 5-234 Scenario Brasilândia – Monthly Weighted Averages of Corrected Global Solar Irradiation – kWh/m².

- Regarding the results coming from GRASS GIS, even after applying the correction factors, the monthly values of global solar irradiation remain elevated throughout the entire year when compared to the other results. The weighted average curve for GRASS GIS has a slight peak in the month of August, which results from intense global solar irradiation values observed in some of the buildings roofs in Brasilândia.

- The reverse situation applies to the results coming from ArcGIS: in this case, throughout the entire year, there is a predominance of lower values with respect to the ones coming from other solar modules. During summertime, the discrepancy in between the curve of ArcGIS and the ones from other software is higher.

- CitySim results in relatively intense global solar irradiation values for the months of December and January, with many buildings roofs represented with values in between 150 and 200 kWh/m². Substantial differences are found in building roofs of the favela-like settlement, with values below 100 kWh/m² even during summer months. During winter months, the weighted average of global irradiation on buildings roofs decline to lower values, especially in July, in which it reaches 89.58 kWh/m².

- SimStadt also results in relatively intense global solar irradiation values, but for a longer period – in between October and March – when the weighted averages of solar irradiation on building roofs go above 150 kWh/m². During the rest of the year, the values are mild, with a more expressive decrease in the month of July. It is also interesting to notice the heterogeneity of values within the same month across buildings in the favela-like area, thus indicating that the buildings are also acting as shading elements to each other.

- Ladybug, in turn, presents a weighted average curve that is almost identical as the one of CitySim – in fact, they overlap almost completely in the graph. In the case of Ladybug, there are relatively higher solar irradiation values on certain buildings roofs during the months of December and January (in between 150 and 200 kWh/m²), and lower values during the month of June (with most values below 100 kWh/m²). The results coming from Ladybug for the month of June are, on average, the lowest throughout the entire atlas. Nevertheless, the weighted average curve results in moderate values, just as the ones resulting from CitySim.

- Finally, the results coming from the 3D Solar Potential Analysis tool of Virtual City Systems present very heterogeneous global solar irradiation values in an intra month perspective. Especially for the favela-like settlement, there is a huge discrepancy among adjacent buildings, with some values between 150 and 200 kWh/m², whereas other values go below 50 kWh/m². This trend demonstrates that the 3D Solar Potential Analysis tool of Virtual City Systems is truly able to detect nuances of solar irradiation values due to shading elements. In general, the curve of weighted averages for the VCS Solar tool presents values that are much like the ones present in the ArcGIS curve, with a difference slightly higher for summer months and minimal for the winter months.

In synthesis, it is possible to affirm that, apart from GRASS GIS, all solar modules result in similar weighted average curves of corrected global solar irradiation values. The curves built with ArcGIS-, CitySim-, Ladybug- and VCS Solar- derived values are very much parallel to each other, differing by some minor offset. Regarding the curve of GRASS GIS, there are some major differences, both in magnitude – the global solar irradiation values are considerably higher throughout the year – and well as in trend – the curve of GRASS GIS contains peaks in the months of June and August which are not present in none of the other curves.

With respect to the eight individual building comparison, the following pages present the corrected monthly global solar irradiation values derived for their roofs.

The building with GML id **“Building_1078833”** is the one presenting the lowest corrected **annual** global irradiation value (399 kWh/m²/year, for VCS Solar).

On a **monthly basis**, it is possible to observe that, except for the one representing GRASS GIS values, all five other irradiation curves are alike. Despite some offsetting among these curves, they follow a similar trend throughout the data series, with values always lower than 100 kWh/m²/month. GRASS GIS, in turn, results in a much higher curve (values around 151.09 kWh/m²/month), with notorious local peaks in the months of March, June and August, which is not the case of the results derived from other solar modules.

Building_1078833 - Monthly Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	164.04	153.37	159.21	146.06	135.17	140.84	131.63	162.26	144.64	148.19	156.17	171.46
ARC GIS	64.50	50.50	36.40	23.05	17.67	15.27	17.16	21.41	29.26	45.61	59.84	70.17
CITYSIM	82.44	72.70	63.00	45.14	31.65	22.91	26.72	39.15	51.55	66.33	74.98	87.46
SIMSTADT	89.90	84.73	72.77	55.61	44.55	35.74	39.02	53.00	61.04	77.67	87.89	99.62
LADY BUG	79.31	72.42	64.65	48.62	35.24	28.84	27.54	45.67	53.29	69.69	78.75	88.27
VCS SOLAR	48.33	41.89	38.74	26.32	18.23	14.38	16.44	22.85	31.43	39.00	44.59	56.61

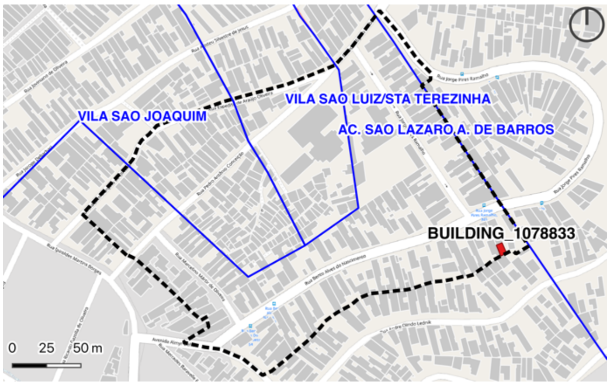


Figure 5-235 Building_1078833 in the city block of Brasilândia.

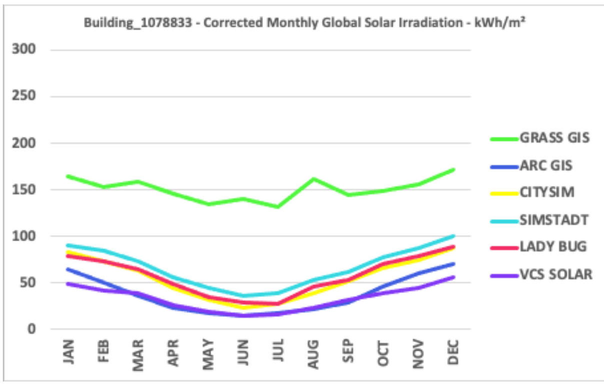


Figure 5-236 Building_1078833 – Corrected Monthly Global Solar Irradiation – kWh/m².

The building with GML id **“Building_1078195”** is the one presenting the highest corrected **annual** global irradiation value (2609 kWh/m²/year, for GRASS GIS).

On a **monthly basis**, it is possible to observe, once more, the same outstanding pattern for the GRASS GIS curve, which contains much higher global solar irradiation values than the other five counterparts, most of them above 200 kWh/m²/month. In the curve of GRASS GIS, it is also possible to notice the unique local peaks in the months of June and August.

Regarding the five other solar module curves, these follow a much closer trend in the data series, with even less offsetting among themselves when confronted to the curves in the previous case. It is also interesting to notice the relatively continuous trend for the period in between October and March, during which the curves do not vary significantly.

Building_1078195 - Monthly Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	208.36	199.84	222.07	223.32	221.10	236.87	218.60	255.81	208.80	197.38	199.27	217.19
ARC GIS	121.93	120.00	126.67	111.65	94.84	82.05	92.81	109.21	115.38	120.70	117.46	126.07
CITYSIM	141.76	136.84	140.41	121.69	102.11	86.89	99.06	119.24	127.41	136.38	135.94	148.51
SIMSTADT	145.40	148.57	145.84	126.04	114.76	95.65	120.08	127.92	130.13	145.16	144.44	154.47
LADY BUG	133.36	129.61	133.07	113.86	95.93	83.02	95.37	112.06	119.99	128.89	129.18	141.59
VCS SOLAR	124.68	122.86	127.31	110.67	93.93	79.82	91.93	109.45	115.34	124.05	122.13	129.71

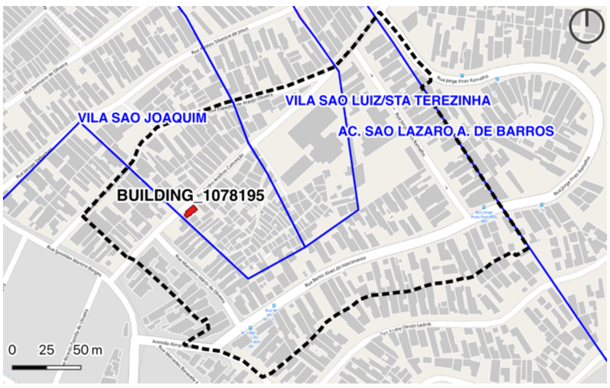


Figure 5-237 Building_1078195 in the city block of Brasilândia.

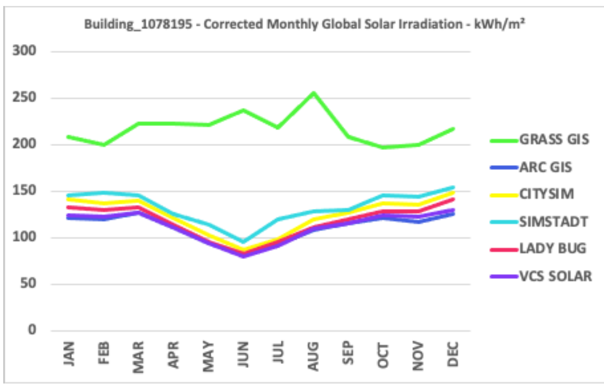


Figure 5-238 Building_1078195 – Corrected Monthly Global Solar Irradiation – kWh/m².

The building with GML id **“Building_1083095”** is the one presenting the smallest **annual** difference in corrected global solar irradiation values (129.7 kWh/m²/year).

On a **monthly basis**, the curves of corrected global solar irradiation are very close to each other, in this case including the one of GRASS GIS. GRASS GIS only presents distinct corrected global solar irradiation values for the months of January and December, which are slightly higher than their counterpart values. SimStadt, in turn, results in corrected global solar irradiation values for the months of May and July that are slightly higher than the ones derived from the other software.

Among all eight buildings that are detailed in this research, this building is the one with the most similar monthly curves among themselves.

Building_1083095 - Monthly Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	188.47	165.76	159.35	130.06	110.67	109.72	105.01	139.41	138.26	156.64	175.35	202.85
ARC GIS	155.52	148.58	152.77	134.12	116.46	102.63	115.06	132.51	138.66	148.45	148.93	162.99
CITYSIM	159.34	151.28	154.21	133.92	115.00	100.63	113.24	131.93	139.65	150.90	152.31	167.33
SIMSTADT	163.96	165.27	161.49	140.60	128.99	108.82	135.17	142.81	144.27	161.03	161.00	172.97
LADY BUG	158.79	150.81	153.61	133.24	114.32	100.12	112.63	131.15	138.95	150.32	151.88	166.70
VCS SOLAR	158.85	150.75	153.56	133.23	114.35	100.16	112.69	131.19	138.92	150.29	151.84	166.50

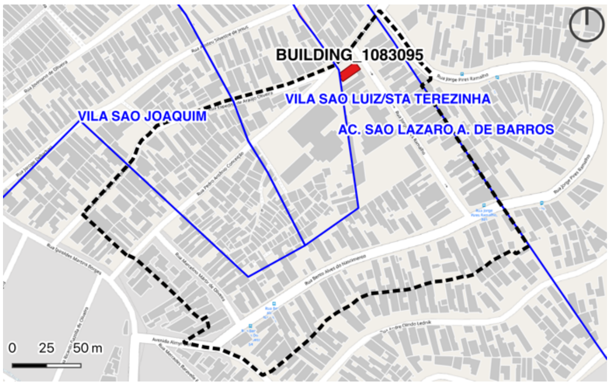


Figure 5-239 Building_1083095 in the city block of Brasilândia.

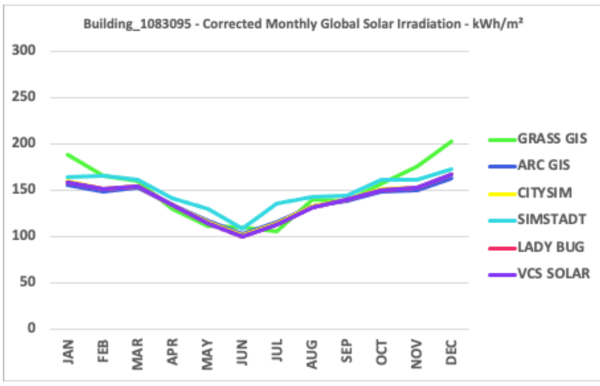


Figure 5-240 Building_1083095 – Corrected Monthly Global Solar Irradiation – kWh/m².

The building with GML id **“Building_1085317”** is the one presenting the largest **annual** difference in corrected global solar irradiation values (1593.44 kWh/m²/year).

On a **monthly basis**, it is possible to observe that GRASS GIS still results in elevated global solar irradiation values throughout the year – in between 150 and 200 kWh/m²/month), with some local peaks in the months of March, June, and August.

The curves of CitySim, SimStadt and Ladybug, in turn, are very much alike, both in monthly amount of energy and in trend throughout the data series.

The curves of ArcGIS and VCS Solar, in turn, contain the lower values among all in the data series. During the months of spring and summer (from October to March), ArcGIS results in higher values, which eventually reach the mark of 100 kWh/m²/month. In wintertime, however, the ArcGIS values drop significantly, reaching the ones in the lower trend of VCS Solar.

Building_1085317 - Monthly Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	189.36	178.17	187.69	173.97	160.64	166.64	156.06	193.49	171.29	173.16	180.41	198.32
ARC GIS	99.58	92.36	81.40	52.58	28.41	19.66	24.20	45.70	67.42	89.23	95.07	105.02
CITYSIM	116.82	104.48	104.42	89.69	74.30	63.86	71.55	89.20	98.42	108.58	111.98	124.90
SIMSTADT	117.72	111.23	105.38	88.90	79.39	66.16	76.84	90.23	96.75	111.63	114.67	126.59
LADY BUG	113.31	103.10	102.58	86.76	71.80	62.90	68.26	85.90	95.08	104.87	108.89	122.36
VCS SOLAR	51.24	49.76	51.55	45.38	33.63	18.25	26.17	38.73	50.15	55.81	57.20	57.88

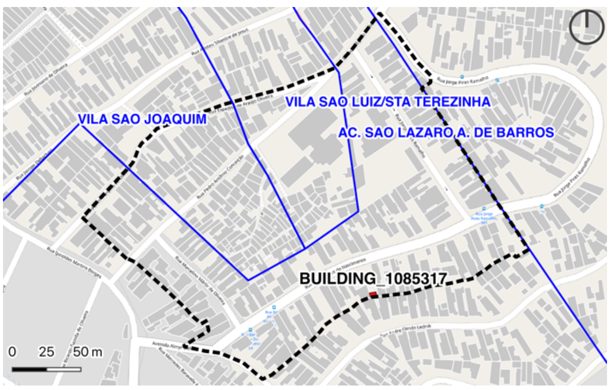


Figure 5-241 Building_1085317 in the city block of Brasilândia.

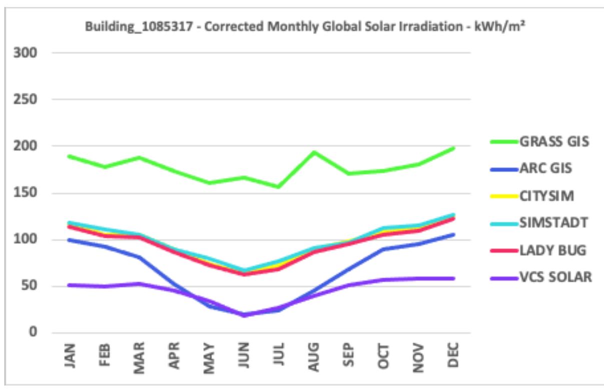


Figure 5-242 Building_1085317 – Corrected Monthly Global Solar Irradiation – kWh/m².

The building with GML id **“Building_1080661”** is the one with the smallest footprint, which is only 7.52 m² large.

On a **monthly basis**, the global solar irradiation curves are offset from each other but tend to follow the same temporal trend, with an exception for the summer values derived for GRASS GIS, which are much higher than their counterpart values for the same period. In the case of GRASS GIS, the months of January and December present global solar irradiation values that go beyond 200 kWh/m²/month.

It is also interesting to notice how similar the curves of CitySim and Ladybug are, and so is the case in between the curves of ArcGIS and VCS Solar, which are composed of much similar values, except for the months of November, December and January.

Building_1080661 - Monthly Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	207.87	170.46	149.04	111.10	83.52	78.68	76.94	113.72	125.28	154.86	190.12	226.48
ARC GIS	118.25	104.25	91.04	58.40	37.08	28.83	33.99	51.12	76.68	100.57	111.64	126.78
CITYSIM	134.15	127.73	121.67	88.22	58.67	41.19	48.57	74.89	99.39	121.28	126.25	141.10
SIMSTADT	139.44	141.71	132.25	102.38	81.33	63.43	73.25	97.89	111.42	134.66	137.63	149.70
LADY BUG	131.29	125.87	121.87	93.00	64.82	52.44	52.88	86.72	104.28	122.36	126.31	140.81
VCS SOLAR	130.76	109.46	96.41	66.16	40.44	34.11	40.17	56.12	80.88	103.07	123.80	139.98

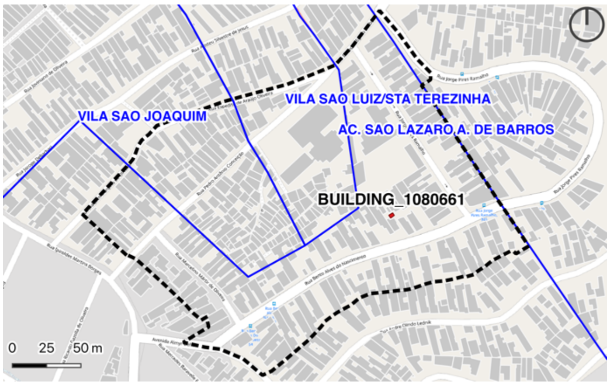


Figure 5-243 Building_1080661 in the city block of Brasilândia.

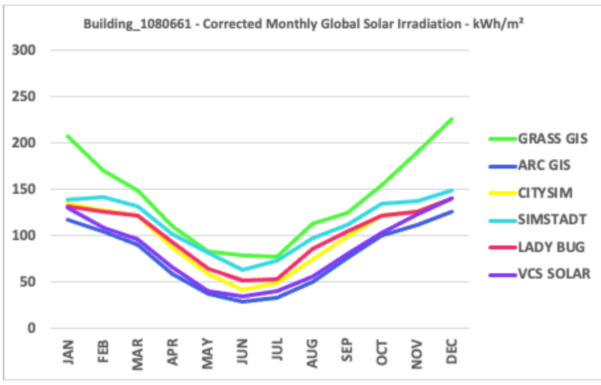


Figure 5-244 Building_1080661 – Corrected Monthly Global Solar Irradiation – kWh/m².

The building with GML id **"Building_1088141"**, in turn, is the one with the largest footprint (1308.1 m²)

On a **monthly basis**, the curve built with GRASS GIS corrected values is once more outstanding, with higher values throughout the year and reaching beyond 200 kWh/m²/month in January and December. The local peak of August is, once more, in evidence.

The curves of CitySim, Ladybug and VCS Solar are almost identical, and indeed the first two are overlayered in the graph by the uppermost purple VCS solar curve. The curve of ArcGIS is also very much alike the previous three, with some exceptions for the summer months, in which the ArcGIS values are slightly lower than the counterpart values derived from other solar modules. The curve of SimStadt, in turn, is relatively similar but with some local peaks, especially in the months of February and July.

Building_1088141 - Monthly Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	208.79	187.21	187.57	166.47	147.93	149.39	141.83	182.48	168.16	178.33	195.89	221.76
ARC GIS	147.28	141.18	145.82	128.58	111.93	98.74	110.64	127.13	132.54	141.22	141.12	154.20
CITYSIM	159.34	151.22	154.13	133.80	114.85	100.47	113.05	131.77	139.57	150.87	152.31	167.32
SIMSTADT	164.09	165.36	161.78	140.80	129.32	109.06	135.73	143.02	144.48	161.16	161.11	173.17
LADY BUG	158.48	150.52	153.25	132.88	114.00	99.83	112.28	130.80	138.62	150.03	151.61	166.42
VCS SOLAR	158.87	150.77	153.58	133.25	114.36	100.17	112.69	131.20	138.94	150.30	151.86	166.52

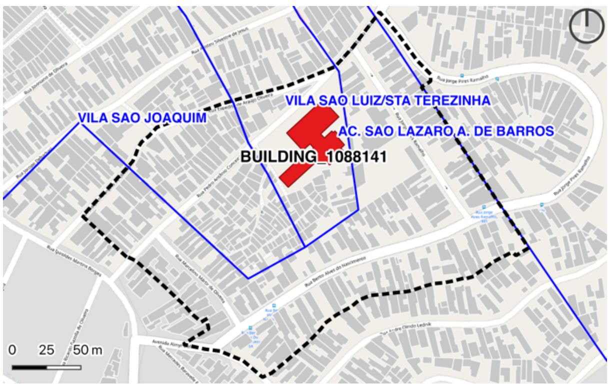


Figure 5-245 Building_1088141 in the city block of Brasilândia.

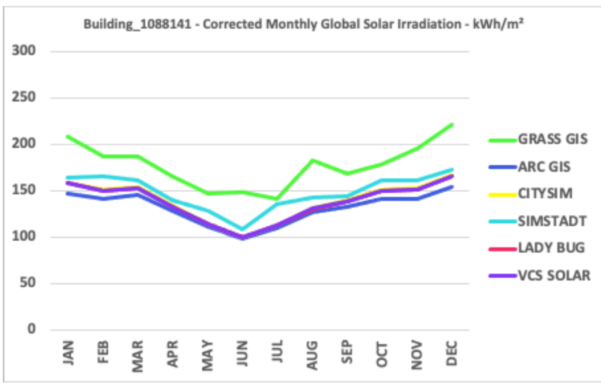


Figure 5-246 Building_1088141 – Corrected Monthly Global Solar Irradiation – kWh/m².

The building with GML id **“Building_1077433”** is the one with the lowest roof, 839.22 m above sea level.

On a **monthly basis**, it is interesting to notice the relatively flatness of all six curves of corrected global solar irradiation values.

The derived results from GRASS GIS are, one more, higher than the counterparts derived from the other solar modules, fluctuating around 145.34 kWh/m²/month. The local peak of August is again in highlight.

The curve with derived values from VCS Solar, in turn, has an interesting shape with peaks in March (late summer) and September (late winter), whereas in the other months the corrected global solar irradiation values decrease. The curve containing ArcGIS derived values, in turn, presents constant values with slightly lower values in the winter period. Once more, it is interesting to notice how similar the curves of CitySim, SimStadt and Ladybug are.

Building_1077433 - Monthly Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	155.78	142.87	148.72	142.42	134.84	141.29	131.96	159.64	137.19	137.80	146.96	164.62
ARC GIS	61.59	59.37	60.54	51.07	43.98	39.33	43.56	49.77	54.28	59.19	59.19	64.36
CITYSIM	74.45	73.74	76.90	68.11	58.21	49.92	56.91	69.17	71.53	73.24	70.94	77.30
SIMSTADT	78.15	81.95	79.79	66.37	57.73	46.56	56.39	67.47	72.31	80.44	80.02	85.72
LADY BUG	74.32	74.68	75.57	61.94	50.29	41.25	47.07	62.03	68.93	73.94	75.65	81.07
VCS SOLAR	31.36	42.25	50.09	44.49	34.77	27.40	32.65	43.79	46.60	44.79	37.56	33.25

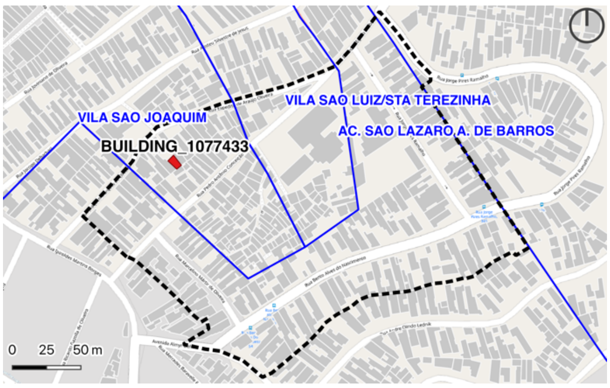


Figure 5-247 Building_1077433 in the city block of Brasilândia.

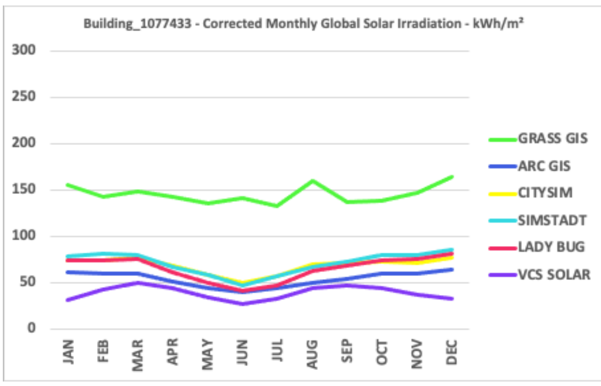


Figure 5-248 Building_1077433 – Corrected Monthly Global Solar Irradiation – kWh/m².

The buildings with GML id **"Building_1071008"** is the one with the highest roof, 867.41 m above sea level.

On a **monthly basis**, it is possible to observe that all the curves are much alike, albeit with some local minima and maxima that differentiates them, especially in the case of GRASS GIS and SimStadt. The curves of CitySim, Ladybug and VCS Solar, in turn, completely overlap in the graph, with monthly differences at decimal levels. ArcGIS, in turn, also results in a similar curve, but in this case the summer values are slightly lower than their counterparts derived from other software.

Building_1071008 - Monthly Global Solar Irradiation - kWh/m²												
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
GRASS GIS	156.24	140.81	144.70	135.30	126.23	127.28	122.58	149.82	133.51	133.79	147.66	163.85
ARC GIS	138.65	133.41	138.84	123.87	109.32	97.17	108.43	123.00	126.59	133.68	132.93	145.07
CITYSIM	159.59	151.47	154.39	134.05	115.11	100.72	113.34	132.04	139.81	151.10	152.55	167.59
SIMSTADT	164.09	165.36	161.77	140.81	129.24	108.94	135.33	143.05	144.53	161.27	161.11	173.02
LADY BUG	159.07	151.06	153.86	133.47	114.53	100.31	112.84	131.37	139.19	150.57	152.15	167.00
VCS SOLAR	159.20	151.06	153.85	133.44	114.52	100.28	112.83	131.35	139.18	150.58	152.15	166.86

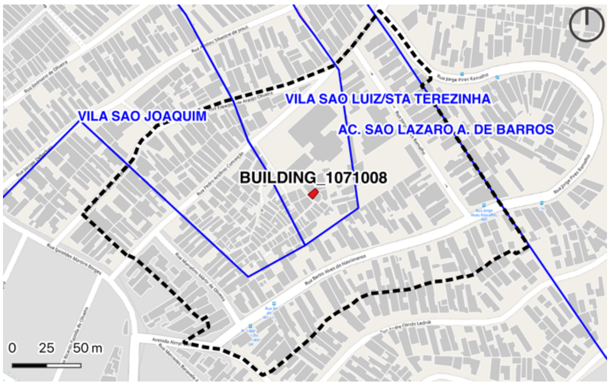


Figure 5-249 Building_1071008 in the city block of Brasilândia.

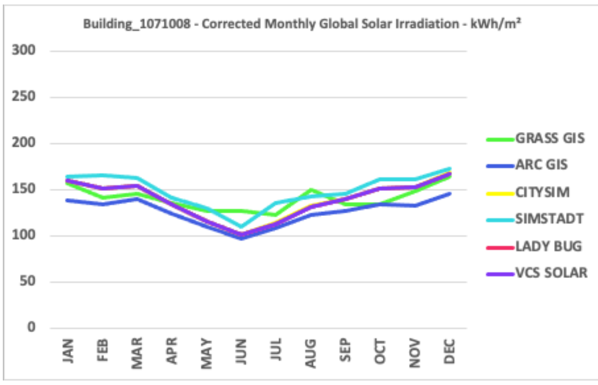


Figure 5-250 Building_1071008 – Corrected Monthly Global Solar Irradiation – kWh/m².

5.3 Roof Mapping

From the roof mapping experiments conducted on a FME workbench, the resulting post-processed building footprints are presented below and confronted with their corresponding originally reconstructed building footprints, as well as with the ground truth dataset coming from the official cadastre of the City Hall of São Paulo. Additionally, the orthophoto dataset – which is contemporary to the Lidar point cloud one – reveals the actual morphology of these buildings in the city block of Brasilândia.

The post-processed building footprints are passed as input data to generate a CityGML model, adopting the same criteria as the ones used for the Santana and Brasilândia scenarios. The resulting model is also illustrated below.

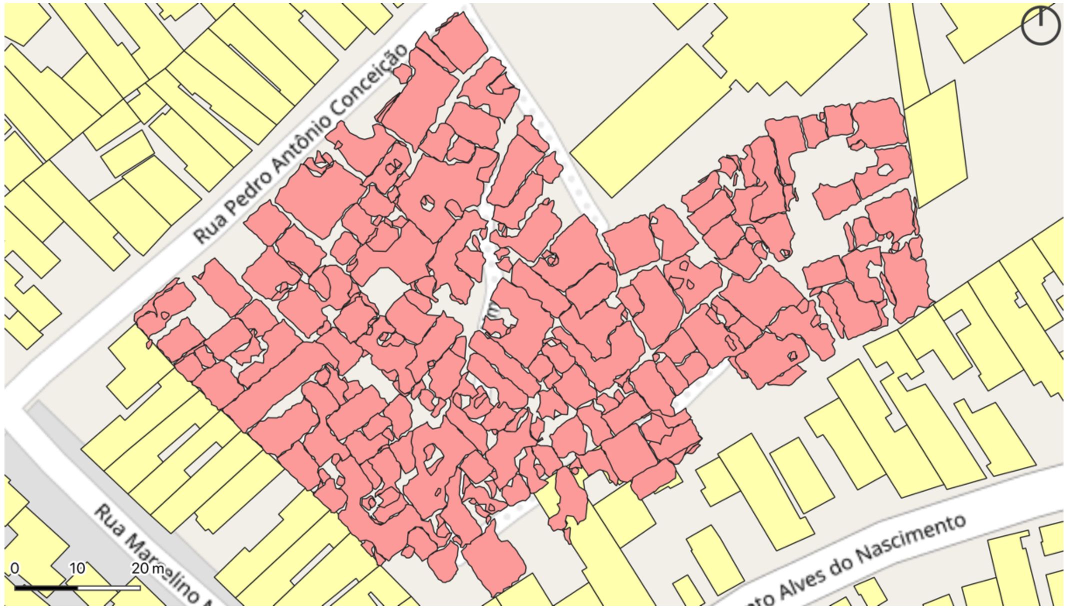


Figure 5-251 Roof Mapping – Reconstructed building footprints.

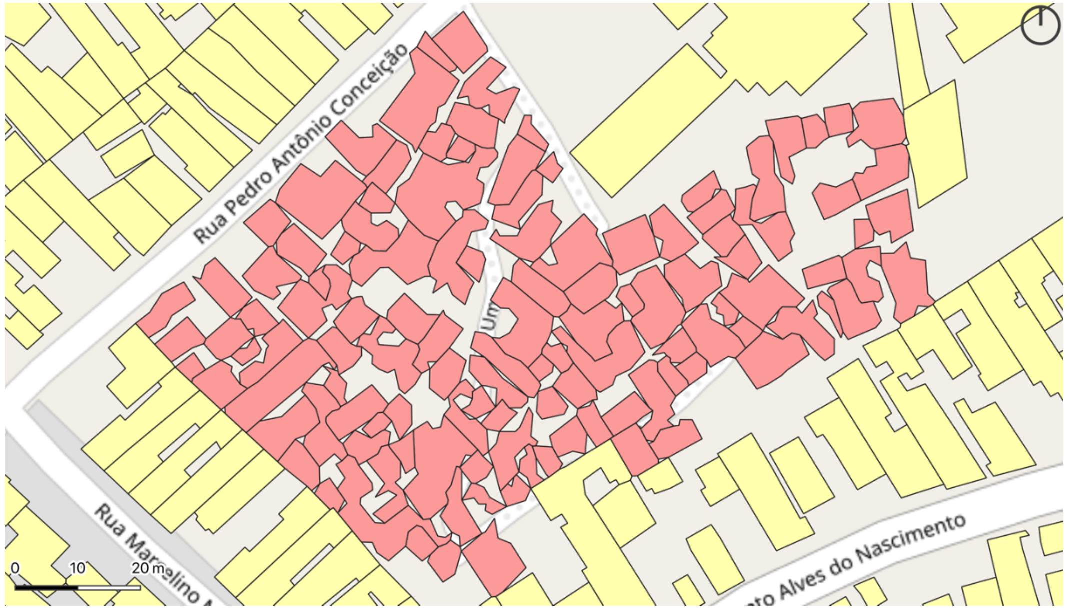


Figure 5-252 Roof Mapping – Post-processed building footprints.

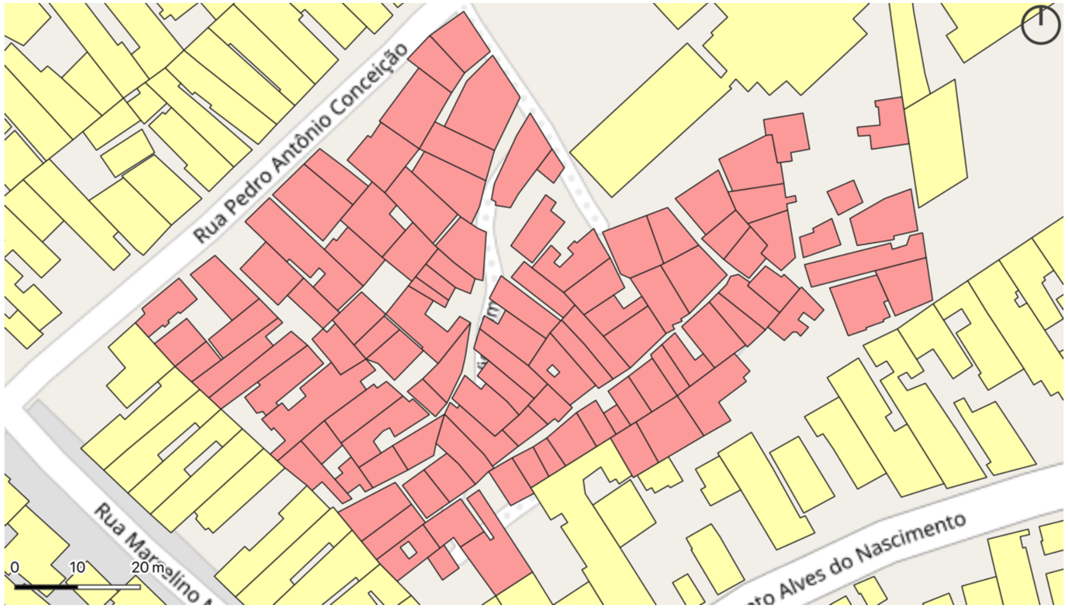


Figure 5-253 Roof Mapping – Ground truth building footprints.

Figure 5-254 Roof Mapping – Orthophoto.


Figure 5-255 Roof Mapping – CityGML model from post-processed building footprints, view to NW.


Figure 5-256 Roof Mapping – CityGML model from post-processed building footprints, view to NE.

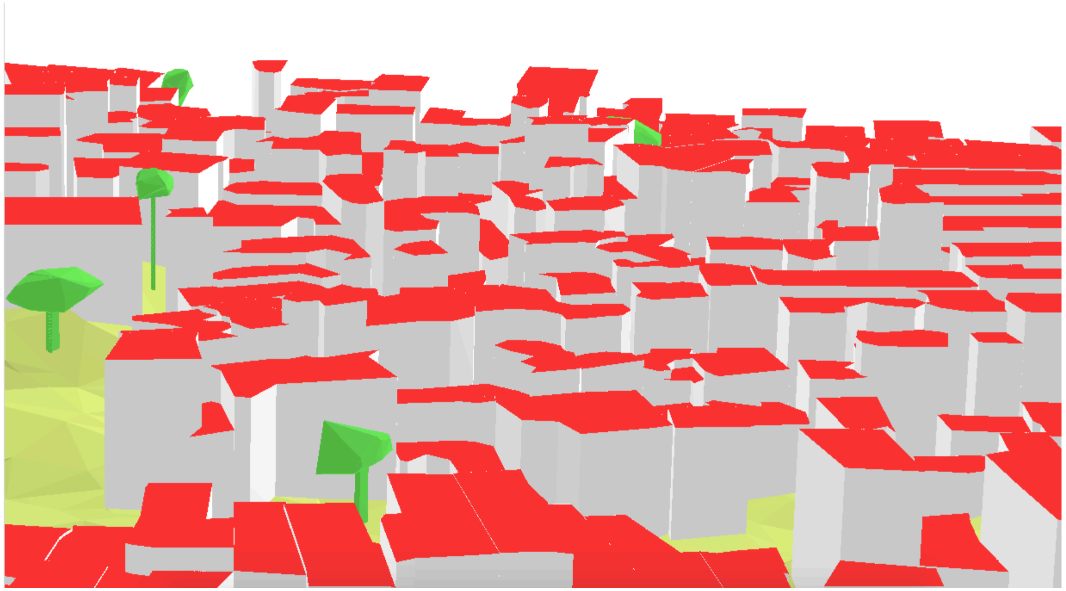


Figure 5-257 Roof Mapping – CityGML model from post-processed building footprints, view to SE.

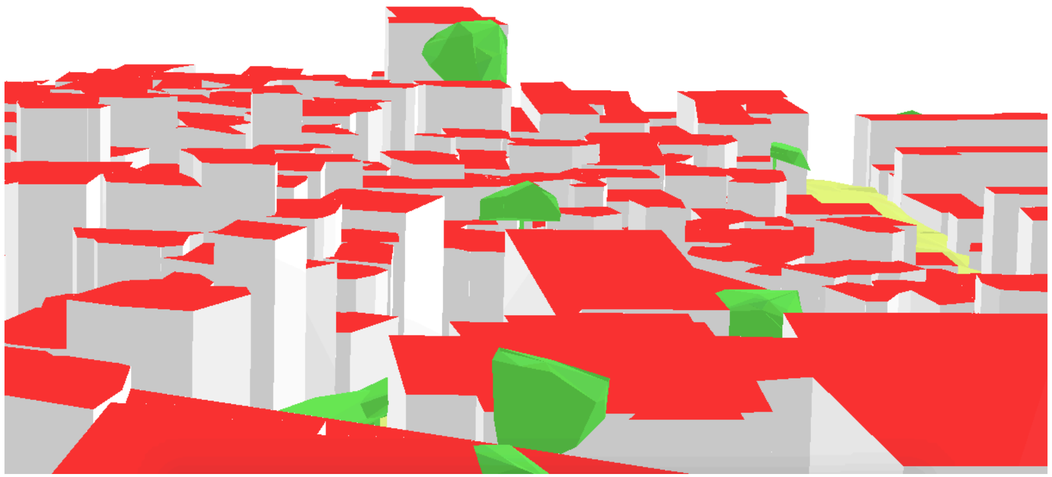


Figure 5-258 Roof Mapping – CityGML model from post-processed building footprints, view to SW.

From the map comparison, it is possible to observe that the post-processing workbench removes several features of insignificant dimensions present in the reconstructed dataset. These small features frequently correspond to water tanks on top of the building roofs, and so their removal is positive for the final dataset. Moreover, the newly generated footprints are much more uniform and do not rely on so many edges, as it is the case in the original reconstructed dataset.

When it comes to covering area and overall disposal, the post-processed dataset offers a great resemblance with the ground truth dataset. It is also possible to recognise many individual buildings by their form and by their correspondence to the ones available as ground truth. The non-built areas in between constructions, resulting in small courts, are also recognisable in both datasets. Preserving this general morphology is positive for solar analysis since the shading effect from adjacent buildings affects the results significantly.

Nevertheless, it is also worth mentioning that the post-processed building dataset still present nonrectangular shapes, and often undesired spiky edges. This is due to a limitation in the Douglas Peuker generaliser, which is set with a threshold such that these building footprints do not become triangular polygons.

In this sense, a drawback of this approach is the fact that it relies on a user-driven process to check if the threshold produces satisfactory results for the specific dataset. Therefore, it would be worth testing this methodology in other informal settlements in the City of São Paulo, so that its validity is based on multiple datasets, and not exclusively on this one.

Still, an important point to be addressed is that the present research assesses global solar **irradiation** values on building roofs of favelas. As irradiation is an average of radiation (kWh) per square meter (m^2), the overall disposal of the reconstructed building roofs and their heights play a much more significant role than the precision of their edges. In other words, even if the resulting building roofs are not as neat as the ones coming from the official dataset, the solar modules will likely result in irradiation values that are representative of that built environment.

In addition to the morphological overview, a statistical analysis is carried out to compare the ground truth dataset to the final post processed building dataset. The results are reported in the following synthesis table and histograms.

	BUILDINGS #	AREA SUMMATION m²	AREA AVERAGE m²	AVG. GROUND ALTITUDE m	AVG. ROOF ALTITUDE m	HEIGHT AVERAGE m	VOLUME SUMMATION m³	VOLUME AVERAGE m³
GROUND TRUTH DATASET	107	4556.29	42.58	851.59	859.88	8.29	37664.23	352.00
POST PROCESSED DATASET	120	4401.48	36.68	851.75	859.52	7.77	35457.58	295.48
ABSOLUTE DIFFERENCE	13	-154.81	-5.9	0.16	-0.36	-0.52	-2206.65	-56.52
RELATIVE DIFFERENCE	12.1%	-3.4%	-13.9%	0.0%	0.0%	-6.3%	-5.9%	-16.1%

With respect to the **number of footprints**, the original ground truth dataset presents 107 buildings within the informal settlement perimeter (pink features on the map), whereas the post processed dataset results in 120, an increase of 13 buildings or 12.1%.

Considering only these buildings within the informal settlement, the **area summation** of ground truth footprints is 4556.29 m², whereas the post processed dataset covers 4401.48 m², i.e., a decrease of 154.81 m² or 3.4% with respect to the original dataset. In terms of **average area**, the ground truth dataset presents a value of 42.58 m², whereas the post processed dataset results in 36.68 m², a decrease of 5.9 m² or 13.9% with respect to the original dataset. The histogram of building area reveals that, in the post-processed dataset, many buildings fall under 20 m², which contributes for this average decrease.

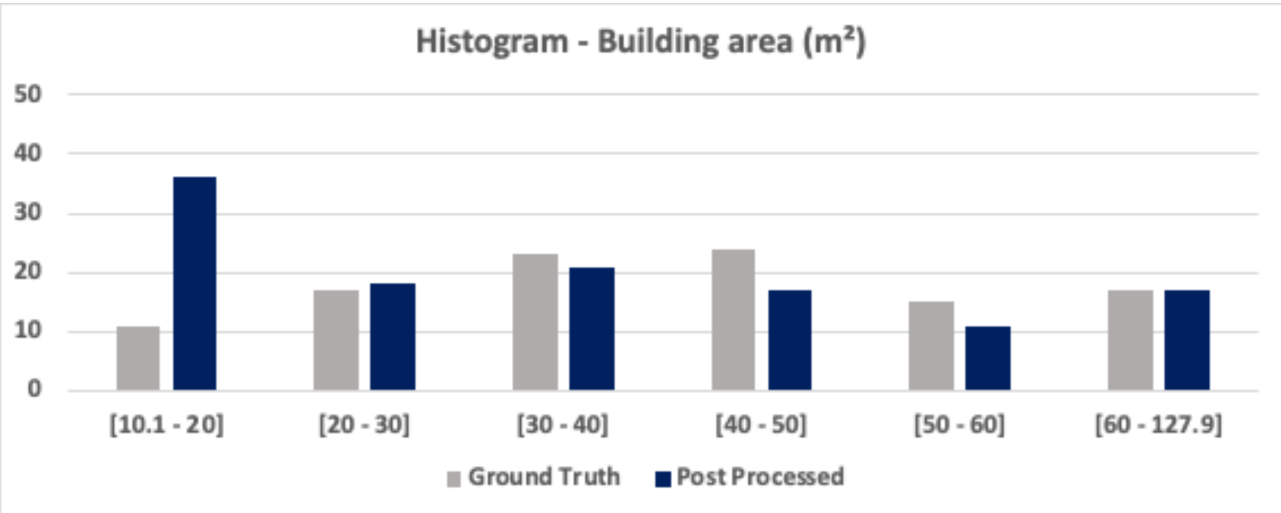


Figure 5-259 Roof mapping – Histogram – Building area (m²).

When it comes to **building ground altitude**, small differences also occur due to the interactive process that designs the CityGML model terrain based also on the building footprints. The CityGML buildings that are designed with the ground truth dataset are, on average, elevated to an altitude of 851.59 m above the sea level, therefore a small difference to the average of 851.75 m in the CityGML model designed with post-processed data, i.e. 0.16 m higher. This is due to the accommodation of the footprints on the relief surface. The histogram of ground altitude also reveals minimal differences.

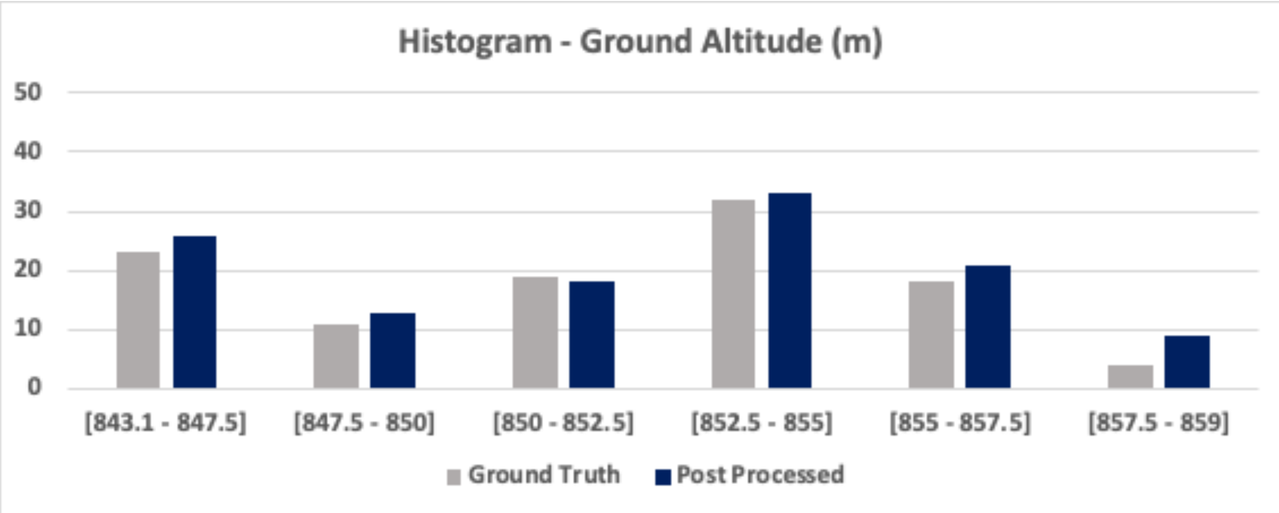


Figure 5-260 Roof mapping – Histogram – Ground Altitude (m).

Regarding the average **building roof altitude**, the difference is also small. The ground truth-based CityGML model results in an average roof altitude of 859.88 m above sea level, whereas the post-processed-based CityGML model results in an analogous average of 859.52, i.e., 0.36 m lower. Since there are more footprints in the post processed dataset, the histogram of roof altitude accounts for more features, which are classified in between 848.8m and 885.0 m above sea level.

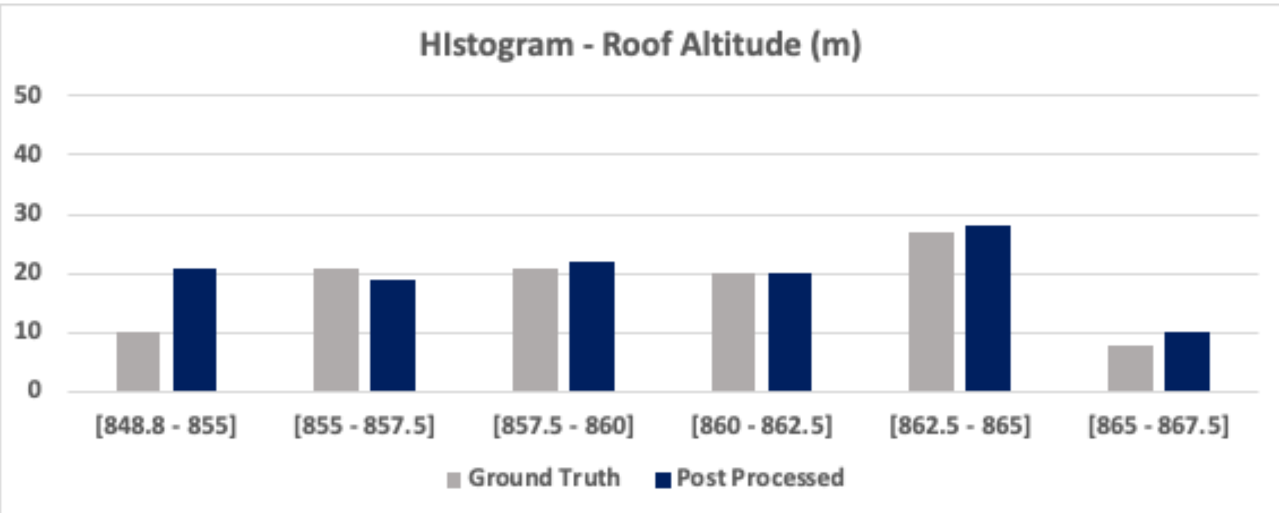


Figure 5-261 Roof mapping – Histogram – Roof Altitude (m).

In terms of **average height of the buildings**, the ground-truth based CityGML model results in a value of 8.29 m, whereas the CityGML model based on the post processed building footprints results in an average height of 7.77 m, i.e., a decrease of -0.52 m or 6.3% from the original average height. Nevertheless, the histogram of building height demonstrates a similar distribution of buildings, with most of them falling in between 6 and 10 meters high in both datasets.

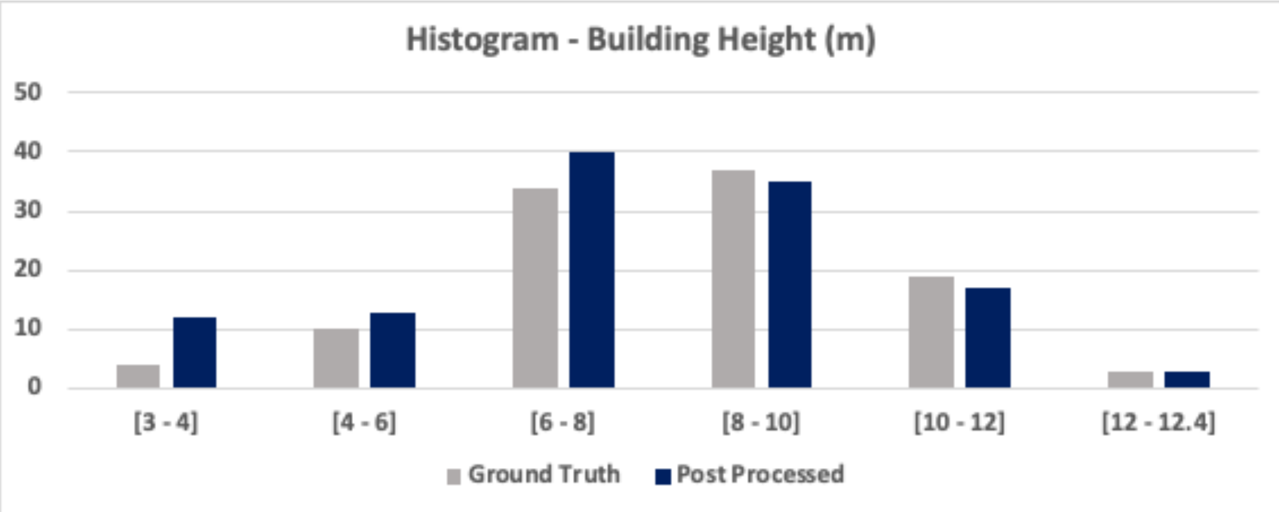


Figure 5-262 Roof mapping – Histogram – Building Height (m).

Finally, with respect to the **volume summation**, the ground-truth based CityGML model results in 37664.23 m^3, whereas the CityGML model based on the post processed building footprints results in a global volume of 35457.58 m^3, a decrease of 2206.65 m^3 or 5.9% of the original volume summation. In terms of **average volume** of these buildings, the ground truth dataset presents a value of 352.00 m^3, whereas the post processed dataset results in 295.48 m^3, a decrease of 56.52 m^3 or 16.1% with respect to the original dataset. In the histogram, post-processed-derived buildings often result in much smaller volumes.

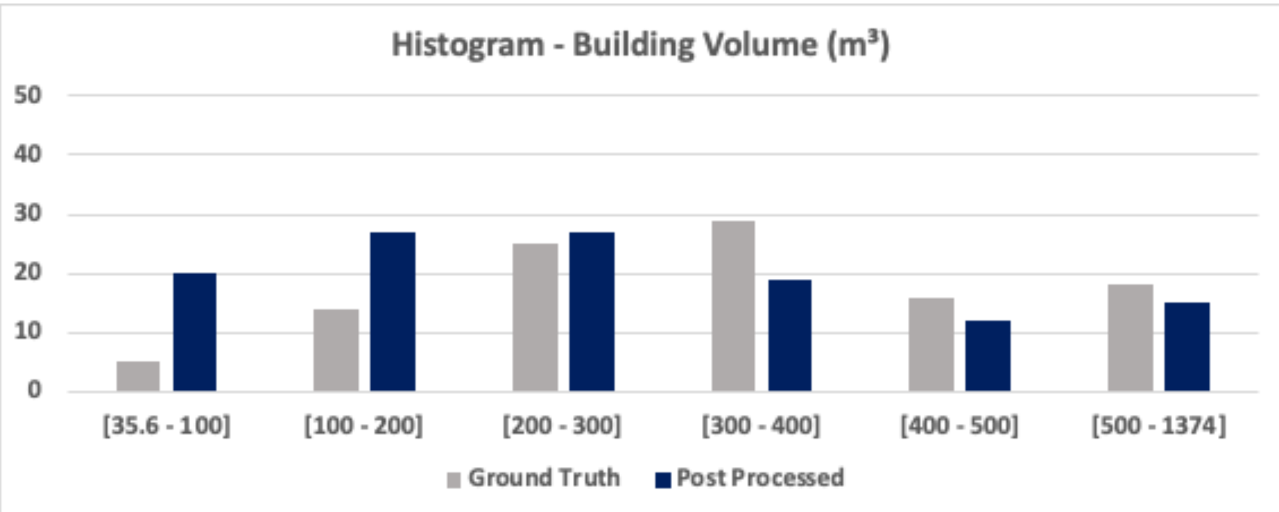


Figure 5-263 Roof mapping – Histogram – Building Volume (m^3).

From the abovementioned metrics, it is possible to affirm that the post-processed footprint dataset results in a CityGML model with buildings roofs that are generally lower and smaller when confronted to the roofs present in the informal settlement portion of the original CityGML model of Brasilândia. Nevertheless, the footprint average area difference is compensated by the existence of more buildings in the post-processed scenario, which mitigates the difference in the summation of footprint (and therefore roof) area. When it comes to the height difference, the corresponding histogram reveals a similar distribution in both datasets, which indicates that both CityGML models will have similar design for solar analysis purposes.

Finally, the present chapter is concluded by responding the sub-research questions of the roof mapping topic:

- **Considering the absence of cadastral data, what are the minimum geodata required to map buildings in a favela, in particular building roofs?**
  Based on the experiments conducted with the building reconstruction algorithm from the 3D geoinformation group and based on the further post-processing workbench, the only necessary spatial dataset to map buildings roofs in a favela is a classified, georeferenced, and dense lidar point cloud. The classification is necessary to filter only points that represent buildings; georeferencing is fundamental to position the dataset in the projection system correctly; and a sufficient point density is fundamental to extract planes from the point cloud.

- **What are the necessary algorithmic steps to achieve such goal?**
  The necessary algorithm steps are:
  - ‘Region growing’, to delimit 3D planes from the points.
  - An ‘alpha shape’ modeller, to refine these computed 3D planes
  - An ‘area filtering’, to remove small features that do not correspond to building roofs, such as the water tanks on top of these.
  - A ‘generaliser algorithm’, such as Douglas Peucker, to simplify the edges of the resulting 2D polygons.
  - A ‘donut hole extractor’, to mitigate the effect caused by the water tanks. It is important to emphasise that, by applying this transformer, holes that are

intrinsic to the building design will also be removed. However, when it comes to an environment of a favela, the simplicity of buildings – often designed in an auto construction process – compensates this drawback, since it is unlikely that single buildings are designed with internal courts or open spaces.

- An 'area gap and overall cleaner', to eliminate overlapping among the reconstructed building footprints and the ones in the surroundings.

- **Considering the urban morphology of favelas, do existing algorithms satisfy the digital reconstruction task?**

  Based on the implementation of the 3D budling reconstruction algorithm of the 3D geoinformation group, it is possible to affirm that it satisfies the digital reconstruction task partially. The reconstruction detects features that do correspond to the ones of ground truth. Nevertheless, a significant post-processing data pipeline is necessary to simplify the resulting features and recreate a topological consistent dataset.

- **If not, what are the important additional challenges to observe when setting up a specific methodology adapted to the environment of a favela?**

  As further work, it is suggested that specific methodologies of roof mapping, adapted to the environment of favelas, should consider that:

  - all roofs as flat, which is a valid assumption for most features,
  - reconstructed small features within other features should be eliminated.

CONCLUSION, DISCUSSION & FUTURE WORK

6 CONCLUSION, DISCUSSION AND FUTURE WORK

6.1 Summarised results

Qualitative Comparison

Regarding the qualitative comparison among all solar irradiation modules, it can be affirmed that each one of these offers potentialities but also limitations. Therefore, it can be concluded that a straightforward choice is not possible, since the optimal solution will be derived from a data driven approach that considers, among other factors: the scale of the favela, its topographical characteristics, the presence/absence of urban features other than buildings (such as vegetation), a possible pre-selection of buildings of interest that reduces the processing time, and so on. Nevertheless, a synopsis of these modules is offered:

GRASS GIS is an open-license software that only requires a raster DSM as the input data for the solar irradiation analysis. With a couple of intermediate steps, the self-contained **r.sun** module derives daily solar irradiation values in a raster data model. These results are computed in relatively short running time when compared to other software.

ArcGIS is a commercial/educational-license software that also requires only a raster DSM as the input data for the solar irradiation analysis. With the support of a Python script, the **Area Solar Radiation** module is run multiple times and computes daily solar irradiation values in a raster data model. Processing these data, however, takes a longer running time.

CitySim is a commercial/educational-license software that imports CityGML files containing LoD2/LoD3 building, tree and relief features. With the support of an additional climate file (and optionally a horizon one), the software runs multiple modules on a one-click basis, among which **short-wave irradiation** analysis. These hourly solar SW records are delivered in a spreadsheet data model. The running time is relatively longer due to the complexity of all simultaneous simulations, which cannot be run in self-contained modules.

SimStadt is an open-license software that also imports CityGML files, but in this case only LoD1/LoD2 buildings are considered, whereas other urban features are neglected. The weather dataset may be provided in a TMY3 file directly by the user, or via the PVGIS or INSEL databases. The self-contained **Solar Potential Analysis** workflow delivers, in a relatively short running time, hourly records of solar irradiance per building surface in a spreadsheet data model, albeit without the GML id of the assessed surfaces, and this absence requires additional assumptions for mapping these features.

Ladybug is a collection of applications mounted on the dual Rhino / Grasshopper environment. It requires an extensive user-controlled data pipeline in a Grasshopper workbench. The **Incident Radiation** module demands both an EPW weather file and Rhino CAD geometries as its input data and delivers, in a relatively short running time, radiation values in a data tree structure according to a pre-defined time interval. The data pipeline can be complemented with the **Real Time Incident Radiation** and other auxiliary components in order to extract interval-defined irradiation values automatically.

The **3D Solar Potential Analysis** from **Virtual City Systems** was run directly by the company given the restrict access to the module. It requires a CityGML file as its geometrical input, whereas the weather data was handled directly by the company. Only building walls and building roofs are assessed, and monthly radiation results are delivered in a valid and enriched CityGML output file.

Qualitative Comparison – Santana

When it comes to the level of accuracy delivered by these six solar irradiation modules, the comparison with the ground truth dataset of Mirante de Santana (INMET) reveals different trends on annual and monthly basis.

It is possible to affirm that ArcGIS outputs an annual summation of irradiation values that is the closest to the one of ground truth values, the former only 0.89% greater than the latter. Nevertheless, from an accuracy perspective, CitySim is the software that outputs a curve that has the best correspondence to the one of ground truth. The monthly Root Square Mean Error (RMSE) values are much lower than the RMSE values coming from all the other software, and so is the case with the annual RMSE value of CitySim.

GRASS GIS presents an annual summation of solar irradiation that is, in modulus, almost as distant as the one coming from CitySim, but the monthly and yearly RMSE values coming from CitySim are still lower compared to the ones of GRASS GIS.

SimStadt and Ladybug present much lower summations of global irradiation values since the summation of direct and diffuse values in the TMY3 and EPW files do not reach the global irradiation values present in the INMET dataset. Their monthly and annual RMSE values are also the highest among all software.

The 3D Solar Potential Analysis tool from Virtual City Systems, in turn, offers relatively low monthly and yearly RMSE values – but not as low as the ones coming from CitySim. The yearly summation of global solar irradiation is 6.55% underestimated, and therefore further away from ground truth with respect to GRASS GIS, ArcGIS and CitySim.

Therefore, considering both the monthly/annual amounts of energy delivered by the solar modules and their monthly/annual RMSE values, CitySim is the optimal choice under the criterium 'level of accuracy'. Albeit the solar module of ArcGIS results in an annual irradiation value closer to the one offered by the meteorological station, the distribution of irradiation throughout the year in CitySim is more representative of ground truth.

Quantitative Comparison – Brasilândia

When it comes to the solar analysis in the building roofs of Brasilândia, the comparison of corrected annual/monthly irradiation values among software reveals that, apart from GRASS GIS, all other solar modules result in relatively similar weighted average curves of global irradiation values. These weighted average curves built with data derived from ArcGIS, CitySim, Ladybug and VCS Solar values (and from SimStadt values to some extent) are very much parallel to each other, differing by some minor offsets. Such similarity among results from these five solar modules is also verified in most of the eight buildings that are studied in detail.

Regarding the curves of GRASS GIS, in most of the analysed cases, there are major differences, both in magnitude – the global solar irradiation values are often much higher throughout the year – and well as in trend – the curves of GRASS GIS often contain local peaks which are not present in any other curves.

In addition, it is also possible to affirm that the results coming from the 3D Solar Potential Analysis tool of Virtual City Systems present very heterogeneous global solar irradiation values in an intra month perspective. Especially for the favela-like settlement, there is a considerable discrepancy among adjacent buildings. This trend demonstrates that the 3D Solar Potential Analysis tool of Virtual City Systems is truly able to detect nuances of solar irradiation values due to shading elements, which is positive for solar analysis in favelas.

Roof Mapping

Finally, when it comes to the roof mapping, it is possible to conclude that the 3D building reconstruction algorithm can deliver a collection of 2D footprints that, in general, covers the extension of the cadastral footprints. Nevertheless, these reconstructed footprints are very fragmented and complex with respect to their official counterparts. The post-processing steps can mitigate these undesired characteristics, but the shapes of the final building footprints are often much more complex than the official footprints. Still, since irradiation is an average of radiation (kWh) per square meter (m^2), the overall disposal of the reconstructed building roofs and their heights play a much more significant role than the precision of their edges. In other words, even if the resulting building roofs are not as neat as the ones coming from the official dataset, the solar modules will likely result in irradiation values that are representative of the built environment.

From the statistical comparison carried out between buildings present in the ground truth-derived CityGML model and in the post-processed-derived CityGML model, it is possible to affirm that the roofs in the latter are generally lower and smaller when confronted to the roofs present in the former. Nevertheless, the footprint average area difference is compensated by the existence of more buildings in the post-processed scenario, which mitigates the difference in the summation of footprint (and therefore roof) area. With respect the roof height, there is a similar distribution of buildings in both ground truth and post-processed histograms, which indicates that these two CityGML models will offer similar designing conditions for solar analysis purposes.

6.2 Choosing a Solar Module

In conclusion to this MSc Thesis, it is opportune to reason about the choice of a software for solar analysis on buildings of favelas in São Paulo. The quality of such analysis is determined both by the representativeness of the geometrical model with respect to the existing built environment, as well as by the similarity to the adopted weather dataset with respect to the ground truth weather conditions.

Geometrical Model

Considering the geometrical model, a special discussion is addressed throughout this MSc research with respect to the specific environment of favelas.

In a traditional approach, a combination of the Lidar point cloud and building footprint datasets of São Paulo offers sufficient conditions for the 3D modelling of the so-called formal city. This is the case, for instance, of the vector-based model of Santana. The resulting CityGML model can be directly or indirectly input in all vector-based solar modules under analysis, i.e., CitySim, SimStadt, Ladybug and VCS Solar.

Nevertheless, the scarcity – or inexistence – of building cadastral data that would represent the environment of a favela introduces the original problem: the 3D modelling pipeline cannot be input with building footprints. In this sense, the roof mapping topic of this research verifies to what extent it is possible to automatically reconstruct the built environment of the scenario of Brasilândia based exclusively on the Lidar point cloud dataset.

The resulting CityGML model, designed with a post-processed footprint dataset, reveals that it is possible, for this specific site, to perform solar analysis on the building roofs in the informal settlement perimeter. However, this might not be the case when analysing other favelas, depending on the complexity and quality of the results derived from the roof mapping task.

In this sense, the raster-based solar modules – namely GRASS GIS and ArcGIS – request much simpler geometrical data as their input, since only a point-cloud derived DSM is sufficient to run the simulations. The resulting raster dataset contains irradiation values at a pixel level, but the problem arises once more if unique values are required per building roof.

Nevertheless, the advantage of the raster-based solar modules over the vector-based ones relies on the fact that it is possible to assess at least a panorama of solar irradiation values distributed throughout the extents of the raster model. In this case, not only visual interpretation but also computational queries can already determine what portions of the favela are more suitable for the installation of PV panels.

Therefore, considering the geometrical aspect, a general criterium can be established:

- If the roof mapping task results in a CityGML model that is either too unrealistic or too complex, the raster-based solar modules are preferred over the vector-based modules.
- In case both raster and vector models result in satisfactory representations of the built environment of a favela, the choice of a solar module must also take into consideration the similarity to the weather dataset.

Weather dataset

Under the context of the simulations conducted throughout this research, it is also meant to provide the six solar modules with weather datasets that are as similar as possible, so that the results will only vary due to differences in the way that the irradiation algorithms are implemented.

Nevertheless, such weather data homogeneity is almost never possible to reach. Each solar modules requires different weather data structure and attributes, and therefore assumptions are made to mitigate such differences, when possible.

GRASS GIS, ArcGIS and the 3D Solar Potential Analysis of VCS adopt their own weather datasets, and the author does not input these solar modules with ground truth data coming from the meteorological station of Mirante de Santana. In this sense, these three modules are completely independent from the ground truth dataset. From the one hand, the autonomy that the software is given to handle the weather data facilitates the user experience, so that they do not need to prepare a weather data prior to the simulation. On the other hand, the user also loses control of the simulation since the input weather data may diverge substantially from the ground truth reference.

SimStadt and Ladybug, in turn, require complex input weather data files, namely a TMY3 and a EPW file, respectively. When designing these files, the author needs to implement data series of direct and diffuse solar irradiance, and these segregated irradiance attributes do not exist in the ground truth dataset of São Paulo. Therefore, choosing SimStadt or Ladybug results in a major drawback of deriving direct and diffuse hourly irradiance values from the global irradiance values, which compromises the verisimilitude of the input weather data with respect the weather ground truth data.

Finally, still under the weather dataset criterium, CitySim is the software which offers the best solution. The climate file, necessary for CitySim, is much simpler than its TMY3 and EPW counterparts. It does not necessarily demand direct and diffuse irradiance attributes, but only global irradiance in case the other two are missing[30]. For the case of São Paulo – which is the focus of the present research – this flexibility is substantially positive for the final irradiation results, which are indeed very close to the ground truth ones, as it can be observed by the qualitative comparison carried out for Santana.

Therefore, based on the two criteria – geometrical model and weather dataset – the author expresses his preference for a raster-based solar module if, on the one hand, the CityGML model results in unrealistic or excessively complex buildings. On the other hand, if the resulting CityGML model is simple enough and representative of the built environment of the favela under analysis, the author suggests the adoption of CitySim. This suggestion is based on the facts that all the weather data attributes required in the climate file can be retrieved from the ground truth dataset, and the achieved results demonstrate a great level of accuracy with respect to ground truth data. The major drawback of CitySim, however, is its

[30] In a meeting with the developer of CitySim, it was explained that a simulation that adopts direct and diffuse hourly irradiance values results in more accurate values, if compared to an analogous simulation that adopts only global irradiance values. As the present research does not dispose of means to verify such difference, this information is hereby addressed only from a theoretical perspective.

running time, which could be compensated by an increase of computing efficiency – if this is possible – or fragmenting the area of study into multiple tiles and running multiple simulations. This last option is out of the scope of this MSc research, but could be further investigated.

6.3 Research Question

Based on the summarised results, it is possible to respond the main research question of the present MSc Thesis: **"How far is it possible to perform solar analysis on buildings of favelas in São Paulo, with the goal of estimating PV Potential?"**

From a roof mapping perspective, the absence of a cadastral dataset represents a complex technical challenge. The building reconstruction algorithm can deliver approximate 2D footprints, which are in turn further processed into more reliable and useful results for the purpose of 3D modelling. These results should be used as a first approximation.

From a solar irradiation assessment perspective, it can be affirmed that morphological characteristics of favelas demand solutions that considers: the great number of building surfaces, the often-complex surrounding relief, and the shading effect that these multiple adjacent buildings produce into each other.

6.4 Suggestions and Future Work

A couple of suggestions are hereby offered to the developers of the solar irradiation modules that are assessed in this research.

- Regarding ArcGIS, the Area Solar Radiation module could be improved to accept pre-computed aspect, slope and horizon maps as its input, thus reducing the computing running time in a for loop statement.

- Regarding CitySim, the developers of the software could improve the Guided User Interface, allowing the user to select which simulations will be run and which will not, thus reducing the computing running time.

- Regarding SimStadt, the importing mechanism could be modified to accept relief and tree features at least as shading elements for the solar potential analysis.

Finally, as a recommendation of future work for other researchers: the methodology hereby implemented could be extended to multiple favelas in the city of São Paulo. In this case, the intrinsic morphological characteristics of these other informal settlements could reveal potentialities and limitations that the present research did not encounter during the available timeline of the MSc thesis.

6.5 Personal Reflection

When I applied for the MSc Geomatics programme, the intention was to enrich my developing career as an Urbanist with theoretical and technical knowledge coming from the geo-information field. The perspective of an urban planner – a professional who understands cities also from a socio-economic perspective – in combination with a geodata approach would increase my capacity of assessing urban phenomena and therefore to propose solutions for real world applications.

This MSc thesis period was a great opportunity to achieve such ambition within the context of the Geomatics programme at TU Delft. Since the early discussions with my supervisors – who were always open for cross-domain debates –, I oriented the preliminary research towards scientific topics within Geomatics that could contribute to the social problems in my home country, Brazil.

Therefore, working on the theme of solar energy in favelas was very exciting, not only because of the social responsibility that comes with the topic, but also because of the previous experience that I have had on energy applications during the MSc Geomatics programme. In addition, I felt constantly stimulated to investigate solar energy in favelas precisely because of the hands-on experiences that REVOLUSOLAR is carrying out in Rio de Janeiro and, therefore, because I could offer concrete solutions for their working pipeline.

Nevertheless, as I had brought 'my own topic' for the MSc thesis, and especially regarding a non-European context, there was an intrinsic additional challenge of defending the importance of working with solar energy in favelas. Albeit it consumed me a longer time to contextualise and to justify the topic in a Dutch University, the experience was very satisfactory, and I arrived at the Graduation Plan assessment convinced about the importance of this topic.

During the learning process, an additional challenge was to consolidate, in one single thesis, subjects of investigation coming from the domains of Urbanism, 3D Modelling and Solar Energy. From this cross-domain field of research, the thesis was structured as it is, i.e., with a city modelling pipeline followed by the 'solar modules' and 'roof mapping' topics.

The city modelling pipeline demanded an extensive and continuous research on how to design the building, relief, and tress in CityGML models according to the OGC standards. This was a really enriching experience for me, and it also motivates me to develop my career by working with digital twins and their applications for urban environments.

The 'solar modules' topic was also very demanding, and by far the most time consuming one. For every new software under analysis, understanding its data requirements and debugging simulation errors took a considerable amount of time throughout the year. For instance, performing solar analysis on Ladybug required me a two-week preparation time to learn Rhino and Grasshopper, excluding the solar analysis itself. Still, the resulting comparison among six solar modules is interesting and I hope that it stimulates further research on the matter.

When it comes to the 'roof mapping' topic, however, I had much less time to investigate the 3D building reconstruction algorithm or even to develop one myself. This was due to the fact that, within the limited time of the thesis, I had to handle six different solar modules. If I was to modify my work pipeline in this past year, I would have analysed only three or four solar modules and paid much more attention to the roof mapping pipeline.

Nevertheless, the results and conclusions respond the research questions that were proposed in the first place, and it is my ambition to continue working on the topic of roof mapping as a personal project of mine. Finally, I conclude the MSc Geomatics programme achieving my initial goal: to become a professional enriched with technical knowledge and who is able to plan and design cities with data for people.

LITERATURE

7 LITERATURE

AGUGIARO G., NEX F., REMONDINO F., DE FILIPPI R., DROGHETTI S., FURLANELLO C.. Solar radiation estimation on building roofs and web-based solar cadaster. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. I-2. 10.5194/isprsannals-I-2-177-2012. (2012)

BILJECKI, F., LEDOUX, H., & STOTER, J. (2016). An improved LOD specification for 3D building models. Computers, Environment and Urban Systems, 59, 25-37.

BRAZIL, Instituto Brasileiro de Geografia e Estatística (IBGE). Geoportal BGE Cidades. Available at: https://cidades.ibge.gov.br/ Accessed on 14 February 2021.

BONDUKI, Nabil, O modelo de desenvolvimento urbano de São Paulo precisa ser revertido. Estudos Avançados nº 25 Vol. 71, 2011. Available at: http://dx.doi.org/10.1590/S0103-40142011000100003 Accessed on 14 February 2021.

CHATURVEDI, K.; WILLENBORG, B; SINDRAM, M.; KOLBE, T. H.: Solar Potential Analysis and Integration of the Time-dependent Simulation Results for Semantic 3D City Models Using Dynamizers. Proceedings of the 12th International 3D GeoInfo Conference 2017 (ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences), ISPRS, 2017, 25-32

DOAN Truc Quynh. Testing and extension of a GIS-supported design too for new urban development areas. Case study: Sloterdijk I, Amsterdam. MSc thesis in Geomatics for the Built Environment. Delft University of Technology, 22nd June 2021. Available at: http://resolver.tudelft.nl/uuid:b38588b1-1854-4ffa-976e-d87db3036b71 Accessed on 16th September 2021.

DUARTE F., FAJARDO, W., & RATTI, C. (2021). Embracing the informal. MIT Technology Review. Available at: https://senseable.mit.edu/papers/pdf/20210424_Duarte-etal_EmbracingInformal_MIT-Review.pdf Accessed on 12 May 2021.

DIRLIK, Arif. Global South. In: Concepts of the Global South – Voices from around the world Global South Studies Center, University of Cologne, Germany – Accessed on 14 February 2021.

ERIKSSEN, Thomas Hylland. What's wrong with the Global South and the Global North? In: Concepts of the Global South – Voices from around the world. Global South Studies Center, University of Cologne, Germany – Accessed on 14 February 2021.

FEITOSA P., SILVA DE LIMA A. J., CUKIERMAN H., "Digital Exclusion of Favelas from the Solar Map of Rio de Janeiro," in *IEEE Technology and Society Magazine*, vol. 39, no. 1, pp. 30-85, March 2020, doi: 10.1109/MTS.2020.2967491. Available at: https://ieeexplore.ieee.org/document/9035548 Accessed on 18 March 2021

FERREIRA, João Sette Whitaker, **MOREIRA**, Tomás Antônio. Governança Urbana no Contexto das Cidades Subdesenvolvidas. IX Congresso Ibero-americano de Urbanismo, Recife, 27 a 30 de novembro de 2000. Accessed on 14 February 2021.

GUSSON, Carolina dos Santos. Efeito da densidade construída sobre o microclima urbano: construção de diferentes cenários possíveis e seus efeitos no microclima para a cidade de São Paulo, SP / Carolina dos Santos Gusson. -- São Paulo, 2014. 152 p. : il. Available at: https://teses.usp.br/teses/disponiveis/16/16132/tde-15082014-143528/en.php Accessed on 8 February 2021

KADEN Robert, KOLBE Thomas. Simulation-Based Total Energy Demand Estimation of Buildings using Semantic 3D City Models. International Journal of 3-D Information Modeling. 3. 35-53. 10.4018/ij3dim.2014040103. (2014)

LANGE, W. (2015). Refinamento do mapeamento da área potencial de telhados de edificações residenciais no Brasil Projeto: Energias Renováveis e Eficiência Energética em Cidades (4ES). Available at: https://www.researchgate.net/publication/319162449_Refinamento_do_mapeamento_da_area_potencial_de_telhados_de_edificacoes_residenciais_no_Brasil_Projeto_Energias_Renovaveis_e_Eficiencia_Energetica_em_Cidades_4ES Accessed on 13 April 2021

MARICATO Ermínia, “Metrópole de São Paulo, entre o arcaico e a pós modernida pós modernidade”, in Souza, Maria Adélia A. de, e outros. Metrópole e globalização, São Paulo, Cedesp, 1999. ________, Metrópole na periferia do capitalismo, Hucitec/Série Estudos Urbanos, São Paulo, 1996.

MARQUES, Eduardo Leão César, A metrópole de São Paulo no inicio do século XXI, Revista da USP nº 102 Dossiê Metrópoles p 23-32. Available at: http://dx.doi.org/10.11606/issn.2316-9036.v0i102p23-32

MEYER, Regina Prosperi; **GROSTEIN**, Marta Dora; **BIDERMAN**, Ciro. “São Paulo Metrópole “, EDUSP, 2004, Capítulo 2, “A construção da metrópole paulista” p. 34 a 110.

MUNTANI G., COCCOLO S., KAEMPF J. BILARDO, M. CitySim Guide: Urban Energy Modelling (Kindle Locations 3-6). CreateSpace Independent Publishing Platform. Kindle Edition.

MURSHED Syed Monjur, **SIMONS**, Alexander & Lindsay, Amy & Picard, Solène & Pin, Céline. (2018). Evaluation of Two Solar Radiation Algorithms on 3D City Models for Calculating Photovoltaic Potential. 10.5220/0006789702960303.

OHORI K. A., LEDOUX H., PETERS R. 3D building reconstruction. Lecture notes for lesson 2.2 of the course GEO1004 3D Modeling of the Built Environment, offered by Delft University of Technology. Version 10 March 2021.

OHORI K. A., LEDOUX H., PETERS R. Introduction to 3D modelling of the built environment: reality, data models and data structures. Lecture notes for lesson 1.1 of the course GEO1004 3D Modeling of the Built Environment, offered by Delft University of Technology. Version 5 February 2021.

REVOLUSOLAR. Relatório Energia Elétrica na Babilônia: acesso, desafios e perspectivas / Organização de Aline Maia Nascimento; Juan Cuervo; Nicole Rosner. – Rio de Janeiro, 2018. 33 p.: Il.; Gráf.

RIBEIRO, S. C. L., JARZABEK-RYCHARD, M., CINTRA, J. P., and MAAS, H.-G.: Describing the vertical structure of informal settlements on the basis of lidar data – a case study for favelas (Slums) in São Paulo City, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-2/W5, 437–444. Available at: https://doi.org/10.5194/isprs-annals-IV-2-W5-437-2019 Accessed on 21 December 2020.

SALAZAR MIRANDA, A., Du, G., GORMAN, C., DUARTE F., RATTI C., & FAJARDO W. (2021). Favelas 4D: Scalable methods for morphology analysis of informal settlements using terrestrial laser scanning data. arXiv preprint. Available at: https://senseable.mit.edu/papers/pdf/20210424_Salazar-etal_Favelas4D_Arxiv.pdf Accessed on 12 May 2021.

SÃO PAULO, Governo do Estado. Empresa Metropolitana de Planejamento S.A (EMPLASA). Available at: https://emplasa.sp.gov.br/RMSP Accessed on 14 February 2021.

SÁNCHEZ C. L., AGUGIARO G., STOTER J. Testing the new 3D BAG dataset for energy demand estimation of residential buildings. Abstract submitted to the Commission IV, WG IV/10

SOMEK Nadia, **CAMPOS FILHO**, Candido Malta, A cidade que não pode parar São Paulo em quatro Planos. IV SHCU 2000, Natal. Available at: http://unuhospedagem.com.br/revista/rbeur/index.php/shcu/article/view/832/807 Accessed on 14 February 2021.

SOUTHWICK Natalie. The Importance and Challenges of Putting Favelas on the Map. Article published on October 11, 2016, in #Rio2016, *Highlight, by International Observers, Favela Qualities, Favela Tourism, Perceptions

PIZARRO, Eduardo Pimentel. Interstícios e interfaces urbanos como oportunidades latentes: o caso da Favela de Paraisópolis, São Paulo / Eduardo Pimentel Pizarro. – São Paulo, 2014. 370 p.: 255 il.

TEMBA Plinio, NERO Marcelo, BOTELHO Lucas, LOPES, Matheus. (2015). Building vectorization inside a favela utilizing lidar spot elevation. 960707. 10.1117/12.2187090. Available at: https://www.researchgate.net/publication/314925044_Building_vectorization_inside_a_favela_utilizing_lidar_spot_elevation Accessed on 12 May 2021.

UNITED NATIONS. A/RES/71/256* New Urban Agenda. English 2017 ISBN: 978-92-1-132731-1 Available at: https://habitat3.org/wp-content/uploads/NUA-English.pdf Accessed on 11 March 2021

UNITED NATIONS. Programa para el medio ambiente. “Jovem brasileiro é finalista de prêmio global da ONU com modelo de energia sustentável para favelas”. Report written on 25th November 2020. Available at: https://www.unep.org/es/node/28548 . Accessed on 14 February 2021.

VILLAÇA, Flavio. São Paulo: segregação urbana e desigualdade, Revista de Estudos Avançados nº 25 Vol 71, 2011 p.40-58. Available at: http://dx.doi.org/10.1590/S0103-40142011000100004 Accessed on 14 February 2021.

WALTER E., KÄMPF J.H. A verification of CitySim results using the BESTEST and monitored consumption values, 2nd IBPSA-Italy conference Bozen-Bolzano, 4th – 6th February 2015

WANG X. Mapping Between Data Inputs of CitySim and CityGML with Energy ADE. Delft University of Technology, unpublished. CitySim Guide Urban energy modelling G. Mutani, S. Coccolo, J. Kämpf and M. Bilardo

WATE Parag, COORS Volker. 3D Data Models for Urban Energy Simulation. Energy Procedia. 78. 10.1016/j.egypro.2015.11.753. 2015

APPENDICES

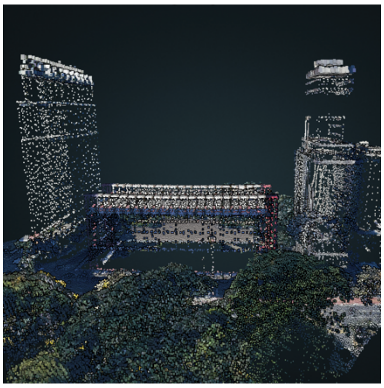
8 APPENDICES

8.1 Appendix A – List of datasets from Geosampa

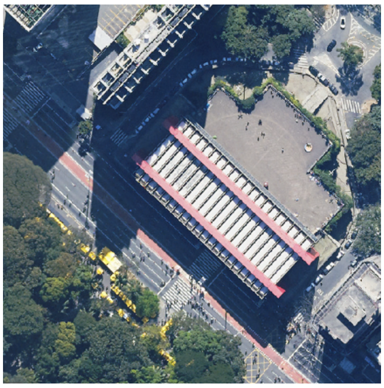
	Group: Housing / Cadastre Sub-group: Housing / Cadastre Name: **Favela – Habitasampa**
Description (PT)	“As favelas se caracterizam por assentamentos precários que surgem de ocupações espontâneas feitas de forma desordenada, sem definição prévia de lotes e sem arruamento, em áreas públicas ou particulares de terceiros, com redes de infraestrutura insuficientes, em que as moradias são predominantemente autoconstruídas e com elevado grau de precariedade, por famílias de baixa renda em situação de vulnerabilidade.”
Scale	1:2.000
Reference date	01/01/2016
Responsible	SEHAB/HABITASAMPA
Projection & datum (web)	UTM/SIRGAS 2000
Projection & datum (downl.)	UTM/SIRGAS 2000
Updating frequency	Daily
File format	Shapefile SAD69-96 and Shapefile SIRGAS2000.
Availability	Map and download.
Environment	Intranet, Internet, WMS and WFS.

Group	Housing / Cadastre
Sub-group	Housing / Cadastre
Name	**Irregular land parcelling – Habitasampa**
Description (PT)	“Os loteamentos irregulares se caracterizam por assentamentos em que a ocupação se deu a partir da iniciativa de um agente promotor e/ou comercializador, sem a prévia aprovação pelos órgãos públicos responsáveis ou, quando aprovados ou em processo de aprovação, implantados em desacordo com a legislação ou com o projeto aprovado. Do ponto de vista das condições urbanas, os loteamentos irregulares ocupados majoritariamente por população de baixa renda sofrem com algum tipo de desconformidade, como a largura das ruas, tamanho mínimo dos lotes, largura de calçadas e implantação de infraestrutura urbana, que configuram uma paisagem árida em que predomina o espaço construído, com alta densidade construtiva, carente de arborização e de espaços livres e de uso comum.”
Scale	1:1.000
Reference date	01/01/2016
Responsible	SEHAB/HABITASAMPA
Projection & datum (web)	UTM/SIRGAS 2000
Projection & datum (downl.)	UTM/SIRGAS 2000
Updating frequency	Daily
File format	Shapefile SAD69-96 and Shapefile SIRGAS 2000.
Availability	Map and download.
Environment	Intranet, Internet, WMS and WFS.

Group	Housing / Cadastre
Sub-group	Housing / Cadastre
Name	
Description (PT)	“As Edificações são polígonos fechados, vetorizados no processo de restituição a partir da observação, por exemplo, de topos de telhados de casas ou edifícios. A camada de Edificações 3D foi obtida através da estilização da camada estabelecendo-se valores.”
Scale	1:1.000 and 1:5.000
Reference date	01/01/2004
Responsible	SMUL
Projection & datum (web)	UTM/SIRGAS 2000
Projection & datum (downl.)	UTM/SAD69
Updating frequency	On demand.
File format	Shapefile SAD69-96 e Shapefile SIRGAS 2000.
Availability	Map and download.
Environment	Intranet, Internet, WMS e WFS.



Group	MDS
Sub-group	MDS
Name	MDS (DSM)
Description (PT)	“Arquivo de nuvem de pontos quem formam o Modelo Digital de Superfície (MDS), classificados em 5 categorias, Solo, Edificação, Vegetação, Obras viárias e Outras feições. Dados obtidos pela tecnologia LiDAR (Light and Detection Ranging) sendo o sensor acoplado a um helicóptero. Classificação feita pelo Consórcio Green-SP e validado pela prefeitura de São Paulo.”
Scale	1:1.000
Reference date	01/01/2017
Responsible	SF/SVMA/SMDU
Projection & datum (web)	SIRGAS2000 / UTM zone 23S (EPSG:31983)
Projection & datum (downl.)	SIRGAS2000 / UTM zone 23S (EPSG:31983)
Updating frequency	On demand.
File format	.laz
Availability	Map and download.
Environment	Internet.



Group	Aerial Imagery
Sub-group	Aerial imagery
Name	Orthophoto 2017
Description (PT)	“Arquivo tipo raster articulado. Sistema de projeção: UTM23S.”
Scale	0.12 m (resolution) e 0.24 m (resolution).
Reference date	01/01/2017
Responsible	PMSP
Projection & datum (web)	N/A
Projection & datum (downl.)	N/A
Updating frequency	N/A
File format	KMZ, DXF, Shapefile SAD69-96 e Shapefile Sirgas. (sic)
Availability	Map and download.
Environment	Intranet, Internet, WMS e WFS: