Comparison and evaluation of different gis software tools to estimate solar irradiation

Giannelli, D.; León-Sánchez, C.; Agugiaro, G.

DOI
10.5194/isprs-Annals-V-4-2022-275-2022

Publication date
2022

Document Version
Final published version

Published in
ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences

Citation (APA)
Giannelli, D., León-Sánchez, C., & Agugiaro, G. (2022). Comparison and evaluation of different gis software tools to estimate solar irradiation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *5*(4), 275-282. https://doi.org/10.5194/isprs-Annals-V-4-2022-275-2022

COMPARISON AND EVALUATION OF DIFFERENT GIS SOFTWARE TOOLS TO ESTIMATE SOLAR IRRADIATION

Denis Giannelli, Camilo León-Sánchez, Giorgio Agugiaro*

3D Geoinformation group, Delft University of Technology, Faculty of Architecture and the Built Environment,
Department of Urbanism, Julianalaan 134, Delft 2628BL – The Netherlands
denisgiannelli@live.it, {c.a.leonsanchez, g.agugiaro}@tudelft.nl

Commission IV, WG IV/10

KEY WORDS: Solar irradiation analysis, GRASS GIS, ArcGIS, SimStadt, CitySim, Ladybug, CityGML.

ABSTRACT:

In this paper, five commonly used software tools to estimate solar radiation in the urban context (GRASS GIS, ArcGIS, SimStadt, CitySim and Ladybug) are run on the same test site and are compared in terms of input data requirements, usability, and accuracy of the results. Spatial and weather data have been collected for an area located in the Brazilian city of São Paulo, in the district of Santana. The test area surrounds a weather station, for which meteorological data of the last 15 years have been collected and used as ground truth when analysing and comparing the simulation results. In terms of spatial data, raster- and vector-based models of the study area have been generated in order to comply with the different input requirements. More specifically, in the case of the vector-based tools (SimStadt, CitySim and Ladybug), a common 3D model based on CityGML and containing buildings, vegetation (trees) and terrain has been generated and used as a common urban model. The paper presents the findings and discusses the results not only from a numerical point of view, but also from the perspective of the overall usability of the software in terms of data requirements, simulation time and task automatisation.

1. INTRODUCTION

The path to decarbonisation dictates that renewable energy sources be exploited more and more. In terms of solar energy, there has seen a drastic increase in terms of technological evolution and adoption of devices for its conversion and usage, e.g. by means of solar thermal or photovoltaic panels. Especially in the urban context, roof and (to a lesser extent) wall surfaces of buildings have become a popular subject of research worldwide in order to estimate *a priori* their suitability and potential in terms of solar energy yield. Accurate knowledge on both the intensity of solar irradiation and its spatial distribution plays, therefore, a major role when exploring or planning the installation of new solar panels. For this reason, in the past years several methodologies and tools have been developed covering both the geometric reconstruction of the urban scene (building stock, vegetation, etc.) and the estimation/simulation of the energy yield. Several cities have applied them and created so-called solar atlases/cadastres, which may however differ greatly from each other also considering the available data used for the purpose. Examples are found all over the world and range from Geneva (Desthieux *et al.*, 2018), to North Morocco (Echlouchi *et al.*, 2017) and Rio de Janeiro (RioSolar, 2016). An overview is provided by Bieda and Cienciała (2021).

In the past years, GIS tools have been employed to solve problems associated with placing solar PV systems in the urban contexts (e.g. Agugiaro *et al.*, 2012; Ramirez Camargo *et al*., 2015; Viana-Fons *et al*., 2020). Further links to existing literature can be found in Hassan *et al.* (2020). Very often, the use of GIS in solar energy is not restricted only to suitability analysis for placing solar PV systems but it also includes the estimation of the potential electric power that can be generated from these systems.

Still, each GIS software package often has its own data requirements, it may implement different radiation models, and can deliver different types of results when it comes to spatial and temporal resolutions, and – most importantly – their accuracy. To the knowledge of the authors, a literature research has not yielded any meaningful results concerning this aspect. For this reason, this paper focuses on comparing 5 different common GIS software packages for solar irradiation simulation when it comes to:

- Input data requirements (spatial data, climate data, data formats, etc.)
- Output data (data formats, spatial and temporal resolution of results, etc.)
- Accuracy, compared to data collected by a weather station used as ground truth.

The software tools compared and described in this paper are: a) GRASS GIS and b) ArcGIS when it comes to raster-based analyses, and c) SimStadt (SimStadt, 2021), d) CitySim (CitySim, 2021) and e) Ladybug (Ladybug, 2021) when it comes to vector-based analyses. Please note that all software applications are free and open-source, except for ArcGIS and Ladybug, the latter being itself open-source, but generally used together with the commercial Rhinoceros 3D/Grasshopper. In both cases, the decision to include them was due to their popularity in the GIS and architecture/urban planning domains, respectively. As a study case, the area surrounding the meteorological station “Mirante de Santana”, in the northern part of the city of São Paulo, Brazil (Figure 1), was chosen, so that a comparison between the simulated and the measured data could be carried out. Various datasets were collected in order to generate the required input datasets for each GIS package. Then simulation results of each software were collected, harmonised

* Corresponding author

and compared. The following sections of the paper deal with each one of these steps and provide the reader with further details and insight.

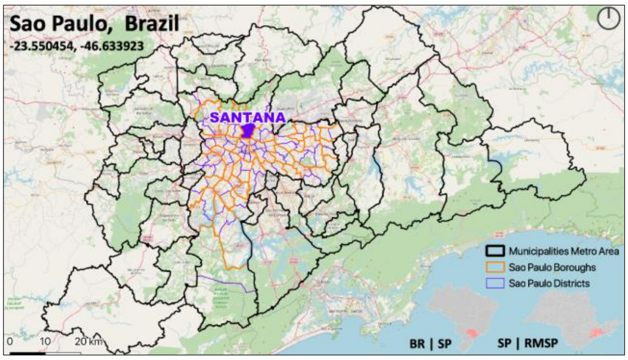

Figure 1. Overview of the study area and its location in São Paulo, Brazil.

Please note that the work presented in this paper is actually part of a larger study that aims to investigate tools and methodologies to estimate the PV potential in the so-called informal settlements (e.g. favelas) (Salazar Miranda *et al.*, 2021) of that city. It must be noted that the lack of (geospatial) information in those areas represents an additional challenge in understanding the urban processes that take place, as such informal settlements are, for many application domains, a sort of "*terra incognita*". For example, in the case of the solar cadastre of Rio de Janeiro, all favelas have *not* been considered and eventually left out of the solar map (Feitosa *et al.*, 2020). Therefore, knowing which tools to use, their data requirements, and their overall accuracy (also without the availability of a reference station nearby for calibration) play a major role when dealing with these urban areas. Further details can be retrieved from Giannelli (2021).

2. DATA COLLECTION AND PREPARATION

Two are the principal data sources for this work: the Mapa Digital da Cidade de São Paulo (Portuguese for "Digital Map of the City of São Paulo") (Geosampa, 2021) – popularly known just as *GeoSampa* – and the database of the Instituto Nacional de Meteorologia (National Institute of Meteorology) (INMET, 2021), also known as INMET. Geosampa is the geoportal from which all publicly available municipal geodata and metadata can be visualised, queried, and downloaded. It hosts a collection of 353 spatial datasets, which are clustered in thematic groups and sub-groups. For this research, data covering the test area were retrieved from these two datasets:

- Building 2D: A 2D polygon shapefile file with building attributes: id, area, and height. From a simple visual inspection, it is possible to verify that the height information is not always reliable. For example, 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 ASPRS15 lidar class values, namely: relief, vegetation, buildings, and other horizontal / vertical features.

The meteorological datasets, in turn, were retrieved from the INMET database, 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. With the exception of cloud nebulosity, all necessary weather data were extracted from station "A701 São Paulo Mirante", whereas the former was collected from station "83781 Mirante de Santana", which lies adjacent to it.

Given the heterogeneity of the GIS software used, the above-mentioned datasets were used to generate different models of the study area (e.g. raster-based and vector-based), although the main purpose was to prepare input datasets as similar as possible, in order to reduce possible sources of deviation in the simulation results. Likewise, where applicable, particular care was taken to prepare and format weather data input files using as much as possible the same measured data from the weather station. The next subsection summarises the data preparation process.

2.1 Preparation of the raster spatial data

GRASS GIS and ArcGIS have raster-based solar irradiation modules. Their minimum spatial data requirements are rather simple and consist just of a DSM. In the case of Santana, from the lidar dataset all classes except relief, vegetation and buildings were filtered out. The resulting point cloud was rasterised at 1 m grid resolution. In the case of "holes", they were filled using bilinear interpolation. The final step was to edit the height value of the cell containing the weather station, which was set to the actual height of the station instead of the underlying terrain. From the resulting raster DSM (covering an area of circa 250×250 m), slope and aspect and horizon maps were derived to speed up (e.g. in GRASS GIS) the solar irradiation simulation time.

In order to include further possible occlusions from nearby objects and relief, two additional larger relief rasters were created. The first one, a DSM at 1 m grid resolution, covers an area of circa 2.9×2.7 km around the weather station, the second one, a DTM at 50 m grid resolution, covers an area of circa 30×20 km to include the hills located mostly north of the weather station. In both cases, horizon profiles consisting of 360 elevation angle values were computed using the 3D coordinates of the weather station, and then merged (Figure 2).

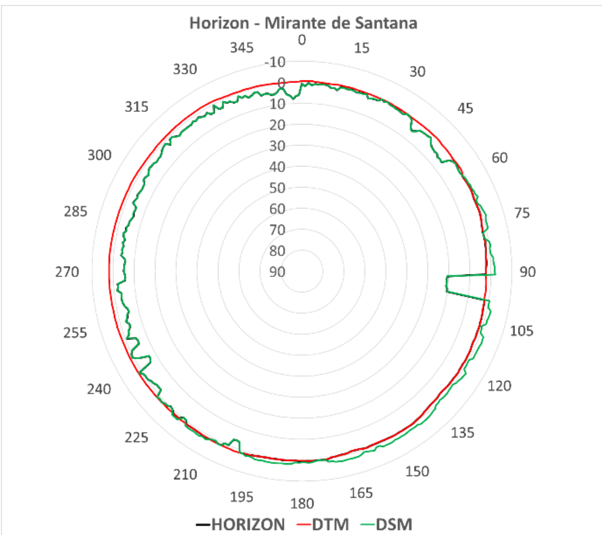

Figure 2. Horizon values from the weather station computed from the DTM and DSM. Azimuth and elevation angles in decimal degrees.

2.2 Preparation of the vector spatial data

In the case of CitySim, SimStadt and Ladybug, a common vector-based input dataset was prepared based on the OCG inter-

https://doi.org/10.5194/isprs-annals-V-4-2022-275-2022

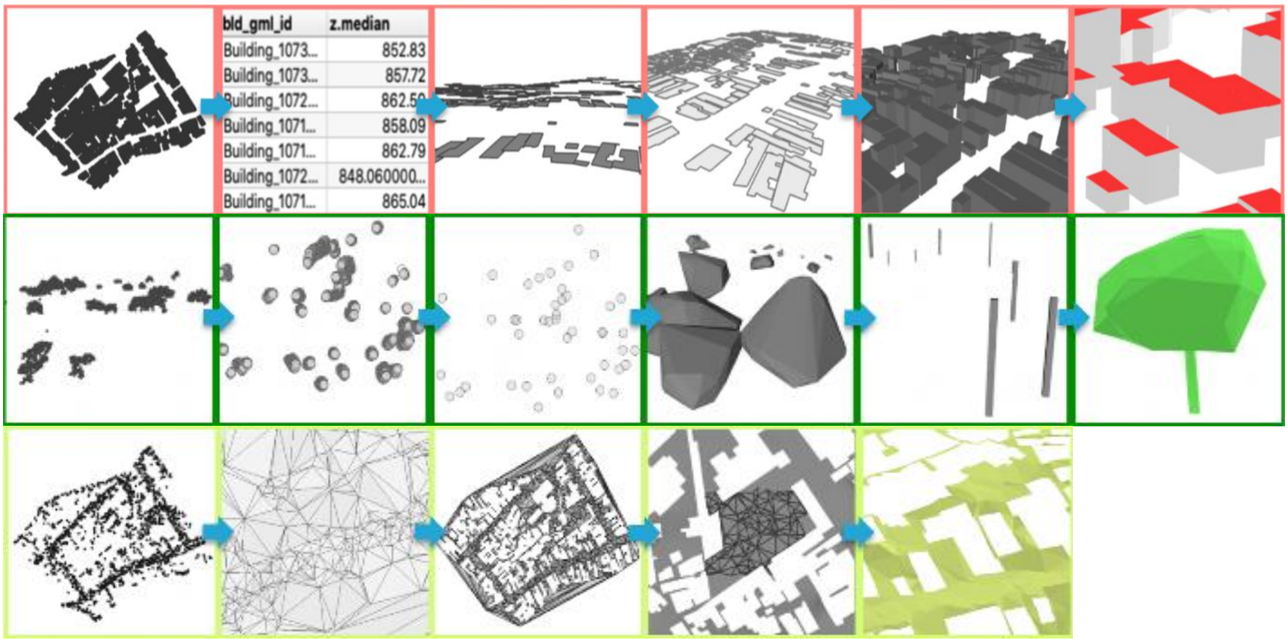

Figure 3. Visual overview of the 3D modelling workflow for buildings [top], trees [centre] and relief [bottom].

national standard CityGML. For the 3D urban scene, only 3 features classes were considered: buildings, vegetation (trees) and relief. The 3D modelling was carried out integrating FME and Python. The datasets used were again the Lidar point cloud and the 2D buildings dataset containing the building footprints.

For the buildings, the 2D footprints were draped on the DTM and assigned the height of the lowest vertex. The original point cloud was filtered and only those points classified as building were kept. They were then intersected with the building footprints dataset. LoD1 shapes were generated by extrusion using the mean height value computed for each footprint. A constraint of minimum 3 m height was added. A prismatic representation was deemed sufficient, as for simulation purposes the nearby buildings are used merely as shadow-casting objects. Besides, flat roofs are the most frequent roof shape in Brazil. Nevertheless, in order to comply with CitySim, all building surfaces were reclassified also into LoD2 Ground-, Wall- and RoofSurfaces. Similarly to the case of the raster, an additional “artificial” building with a footprint of 1x1 m was added in lieu of the weather station, the flat RoofSurface being at the height of the station. A graphical overview of the process is given in Figure 3 [top]. For the vegetation, only the Lidar points classified as vegetation and not overlapping the building footprints were kept. Using a clustering algorithm with a radius of 5 m, resulting tree points were used to reconstruct tree crowns computing their 3D convex hulls. Tree trunks were modelled as simple extruded geometries from a square-shaped cross-section of 25 cm side. The resulting geometries (crowns and trunks) were stored as objects classified as CityGML SolitaryVegetationObject. Further details can be found in Giannelli (2021). A graphical overview of the process is given in Figure 3 [centre]. For the relief, only those Lidar points classified as ground were kept. A TIN was generated spanning over the extents of the study area, however the building footprints were then used to cut “holes” in the TIN and therefore reduce the computation time as relief areas “under” buildings are not included into the simulation. A graphical overview of the process is given in Figure 3 [bottom]. The resulting 3d city model of the study area is presented in Figure 4.

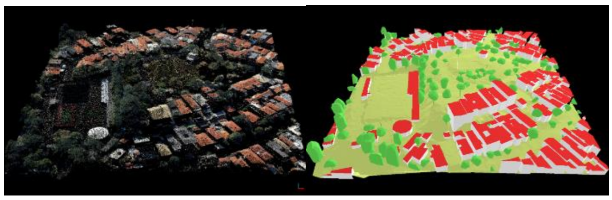

Figure 4. 3D representation of the input point cloud [left] and the resulting 3D scene modelled as Buildings, Relief and SolitaryVegetationObjects in CityGML [right].

2.3 Preparation of the weather data

Regarding the weather data, both meteorological datasets – the automatic and the conventional one – were downloaded with their records ranging from 26/07/2006 (the first non-null records of the automatic station) until 25/07/2021, therefore a period of fifteen years. The corresponding hourly values of solar irradiation were first averaged through the 15 years, in order to eventually obtain hourly values of a “typical year”. The average hourly values were then aggregated at daily and monthly level, as shown in Figure 5 and Figure 6, respectively. During the process, the timestamp information (in UTC) present in the datasets from INMET was considered. Since São Paulo lies in UTC -3, the records were shifted accordingly.

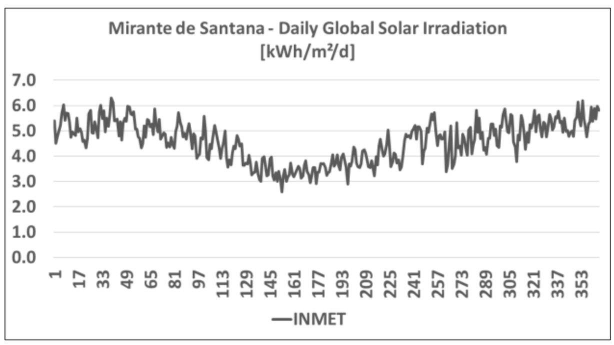

Figure 5. Average daily solar irradiation values from the weather station “São Paulo Mirante” (Data source: INMET).

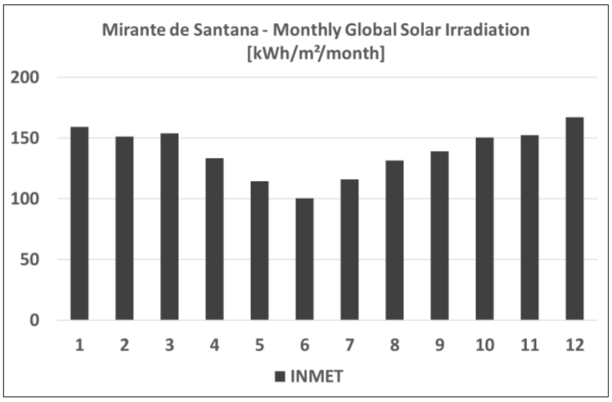

Figure 6. Average monthly irradiation values from the weather station “São Paulo Mirante” (Data source: INMET).

Finally, some simulation tools require global diffuse *and* direct normal surface irradiation values, but the data available from the weather station only provide the former, CitySim was used to pre-process the weather data and generate the diffuse and direct beam values to be then used for all simulations. CitySim implements the DISC-model (Maxwell, 1987).

3. SOLAR IRRADIANCE SIMULATION

Although it is beyond the scope of this paper to provide the reader with an in-depth explanation of how the aforementioned different tools carry out the solar irradiance simulations, this section will offer some details on data requirements, limitations and workarounds that were identified for each software package. Given the differences existing among the software tools, the input datasets have been prepared in order to provide, on the one hand, a set of input data as similar as possible, but also, on the other hand, to evaluate which characteristics and capabilities are common among the tools, and which are not, e.g. in terms of supported input features (buildings, terrain, vegetation, etc.). An overview of the findings is presented in Table 1. The reader can however find more details on the implemented radiation models following the provided links. In general terms, all simulation tools have some basic common requirements such as the location, the surface and atmospheric conditions. Besides the relief surface – which is mandatory –, for GRASS GIS and ArcGIS most of the other simulation input values are optional, user-adjustable parameters, otherwise calculations are carried out for clear-sky conditions. For Ladybug, hourly values of direct and diffuse irradiance are necessary. SimStadt and CitySim require as input more detailed weather data with properties such as air temperature, wind direction and speed, global and diffuse irradiance, relative humidity, to name a few.

3.1 GRASS GIS

In GRASS GIS the r.sun module computes direct (beam), diffuse and reflected solar irradiation raster maps for given day, latitude, surface and atmospheric conditions (r.sun, 2021). The model considers the shadowing effect of the local topography and computes beam, diffuse and reflected radiation for the clear sky conditions, i.e. not taking into consideration the spatial and temporal variation of clouds. The only required input is a raster-based DSM of the study area. To speed up the simulation time, slope, aspect and horizon maps can be precomputed. Additional, user-adjustable parameters include the ground albedo and the Linke atmospheric turbidity coefficients. In our case, all DSM-derived maps were pre-calculated, including the 360 horizon maps (one for each horizontal angle). The monthly Linke atmospheric turbidity coefficient values for the study area were retrieved from the open data available at the Solar and Radiation data web portal (SoDa, 2021). For the albedo, the default value of 0.2 was used. The simulation was automatised by means of a Python script. The output consists of 365 raster maps with daily values of solar irradiation.

3.2 ArcGIS

In ArcGIS, under the Spatial Analyst licensing, there are two simulation tools available: the Area Solar Radiation and the Points Solar Radiation module. The first one calculates the insolation across a region and the latter one for a specific position. Their fundamental requirement is a DSM. Optional, user-defined parameters allow to specify the time frame (e.g. range hours or days) and the use of a slope surface or the sky size. With the latter, the resolution of the viewshed, the sky map, and sun map rasters that are used in the radiation calculations are meant (ArcGIS Solar, 2021). These are upward-looking, hemispherical raster representations of the sky and do not have a geographic coordinate system. The tools can be executed manually via the GUI or automatised by means of Python scripts. In our case, a script to run the area simulation tool for the 365 days of the year was created. The computation is generally longer than with other tools described in the next sections, as it (re)calculates the slope and aspect maps for each iteration. The output consists of a set of 365 raster files with daily values of solar irradiation.

3.3 CitySim

CitySim is a free and open-source energy simulation software developed for dynamic simulation of clusters of buildings. It was developed to provide urban energy planners with decision support using 3D geometrical buildings at urban district scale. Although the CitySim solver works with its own data model (CitySim XML file format), the GUI offered by CitySim Pro can import CityGML files and export CityGML data with Energy ADE content, similarly to SimStadt. CitySim generates results with a much finer temporal resolution reaching hourly values. Besides CityGML, CitySim exports results as tsv (tab-separated-values) files. For solar irradiance simulations, there are two mandatory inputs: 3D geometry and a climate file. Optionally, an horizon file can be provided. The climate file is a tsv file that encodes the following parameters for every hour of a typical year (CitySim wiki, 2021): global horizontal irradiance, air and surface temperature, wind velocity and direction, relative humidity, total precipitation and cloud nebulosity. The tsv-based horizon file gives information on the far field obstruction of the skyline. It contains the values computed and shown in Figure 2. CitySim requires manual interaction from the GUI only to import the input datasets and to merge them into a unique XML-encoded CitySim file. From that moment on, the simulations can be run by means of command line scripts. The output data are written to a tsv file containing hourly values of solar irradiation. In our case, for the geometries, the above-mentioned CityGML file was read. Buildings, vegetation and relief geometries were all converted into the CitySim XML-based file format. Weather data from the Santana weather station was formatted as the requested tsv climate file. The optional horizon file was provided as well. The output results were aggregated from hourly to daily values of solar irradiation in order to obtain comparable results.

3.4 SimStadt

SimStadt is an open-source, multi-platform, Java-based software. It performs different analyses on buildings, such as the energy-demand computation of buildings based on the energy-balance

	GRASS GIS	**ArcGIS**	**CitySim**	**SimStadt**	**Ladybug**
License	FOSS	Commercial	FOSS	FOSS	FOSS (but comm. Rhinoceros 3D – Grasshopper)
Minimum input data requirements	Raster-based DSM	Raster-based DSM	Vector-based 3D scene (Buildings in CityGML LoD2), weather data	Vector-based 3D scene (Buildings in CityGML LoD1/2)	Vector geometries as Brep/Mesh, weather data file
Optional input files	Slope, aspect, Linke turbidity, albedo maps	N/A	Horizon file	N/A	N/A
Interaction	GUI/Python/shell scripts	GUI/Python scripts	GUI/shell scripts GUI only (for data import/conversion)	GUI/shell scripts	GUI/Python scripts
Urban features	All features represented in the DSM	All features represented in the DSM	Buildings Vegetation Relief	Buildings	Buildings Vegetation Relief
Approx. running time (HH:MM)	08:40	03:18	16:44	00:06	00:26
Results: type	2.5D surfaces, i.e. no vertical surfaces	2.5D surfaces, i.e. no vertical surfaces	3D surfaces (roof, walls, etc.)	3D surfaces (roof, walls, etc.)	3D surfaces (roof, walls, etc.)
Results: format	Raster file	Raster file	TSV file	OUT file	Data tree
Results: minimum temporal resolution	Second (mode 1), daily (mode 2)	Hourly, generally daily	Hourly	Hourly	Hourly

Table 1. Comparison of the tested simulation tools for solar irradiance. Tests were carried out on a MacBook Pro equipped with a 2.5 GHz Dual-Core Intel Core-i5 processor, 8 GB RAM, a 1 TB SSD and running Windows 10 on Boot Camp.

method (i.e. monthly and yearly values), estimation of solar irradiation and PV potential, etc. Besides a csv file format, SimStadt can also export CityGML data enriched with the Energy ADE. This functionality is, at the moment of writing, still experimental, though. In terms of input, SimStadt accepts (only) CityGML buildings in LoD1 or LoD2. No other city objects are currently supported such as terrain, vegetation, or other shadow-casting urban objects. Regarding weather data, these are mandatory and they must be formatted as a tmy3 file. The mandatory values are: global horizontal irradiance, direct normal irradiance, diffuse horizontal irradiance, total sky cover, dry-bulb temperature, dew-point temperature, relative humidity, station pressure, wind direction and wind speed. The software requires manual interaction from the user through a GUI, but its operations can be otherwise scripted via command shell. The results consist in average yearly irradiation values per surface. Nevertheless, SimStadt stores hourly values in a cache file located in the same directory of the project, which can therefore be retrieved. The output data are written to a text file. In our case, the CityGML-based 3D model created before was read directly in SimStadt. Only buildings were loaded. For the weather data, a tmy3 file was generated *ad hoc* using the data from the weather station in Santana (see section 2.3). The radiation model chosen was Perez (1987) in order to include the mutual shadowing effect of nearby buildings.

3.5 Ladybug

The Ladybug Tools are a collection of free and open-source applications for environmental design. Although they are free software, they run best embedded in Grasshopper, a visual programming language and environment that runs within the (commercial) software Rhinoceros 3D. The Ladybug Tools connect 3D CAD interfaces to several simulation engines. In particular, the Ladybug module allows to perform solar radiation studies, view analyses, sunlight-hours modelling, etc. Since LadyBug builds upon Rhinoceros 3D, which itself is a CAD program and not *strictly speaking* a GIS software, the geometries need to be prepared beforehand as a DWG file. Additionally, Ladybug requires the creation of a so-called sky-matrix file, which can be generated from one Ladybug location, constructed with the weather station's name, latitude, longitude, altitude and timezone, and two Ladybug Data Collections, one containing hourly diffuse irradiation values and the other the direct beam ones. In our case, the geometries of the CityGML file were first triangulated and then converted into a layer-based DWG file, storing each surface mesh as a single layer named after its GMLID to keep track of the geometries and associated results. Buildings, trees and the relief could be imported. Additionally, Ladybug allows to classify the imported city objects into objects to be simulated and simple shadow-casting objects, therefore reducing the simulation time. Regarding the weather data, the hourly values of direct and diffuse irradiation used also for the other simulation tools were loaded into the workbench in a simple text box, and further transformed to comply with the sky matrix inputs. Results are available as hourly values and were therefore aggregated into daily values in order to obtain comparable results.

4. RESULTS COMPARISON AND DISCUSSION

The results coming from each simulation software were collected and compared with those of the weather station in Santana, considered as ground truth. In order to obtain comparable results, daily values of global solar irradiation were retrieved from the simulation results of each tool at the position of the weather station, i.e. either the raster cell containing the weather station coordinates, or on the RoofSurface of the artificial building in the 3D city model which ideally "supports" the weather station. If needed, results were aggregated in order to obtain time series of daily values of global solar irradiation (in $kWh/m^2/d$). Figure 7 shows a plot of all time series and allows a first visual comparison of the results. The reasons for the "spike" in the ArcGIS results near the end of the year are not yet clear and need further investigation. In the following subsections, the main findings are presented. Please note that the meteorological station is identified as "INMET".

4.1 Comparison of aggregated yearly values

The first comparison was carried out in terms of total energy over the whole year. The value of the weather station was considered

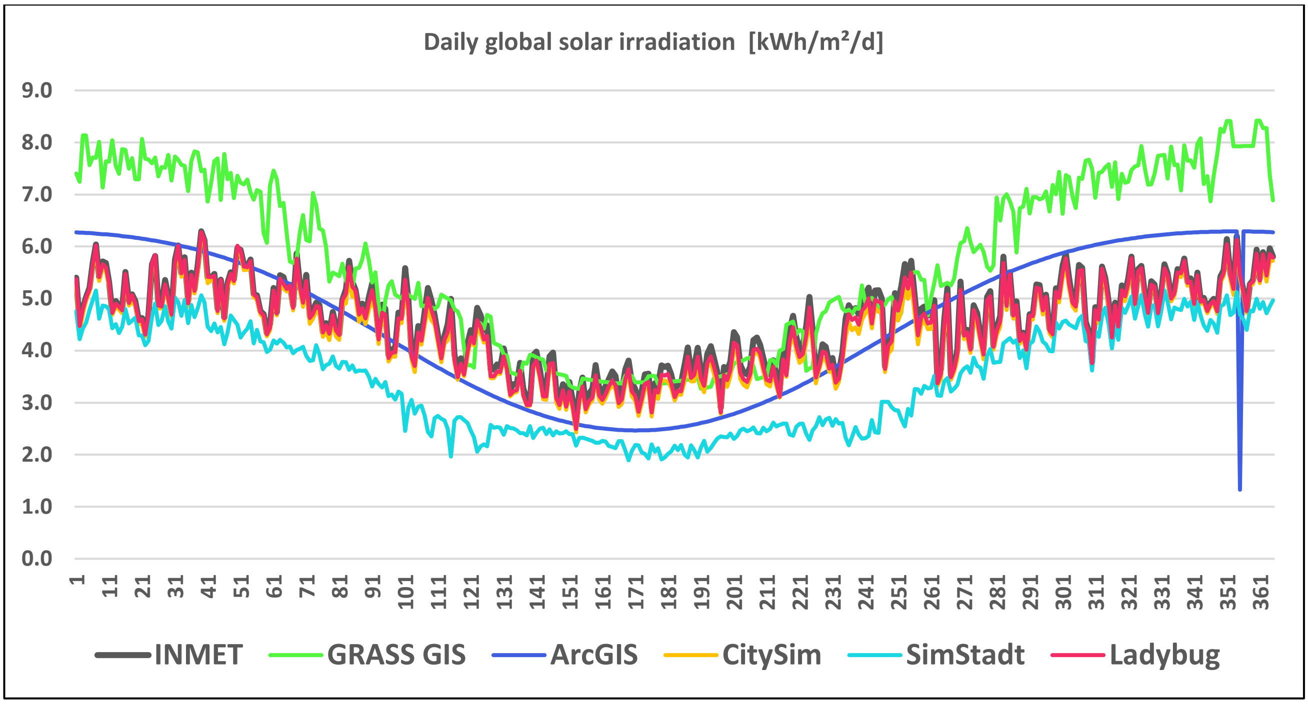

Figure 7. Yearly time series of daily global solar irradiation resulting from the 5 simulation tools and compared to the meteorological station (INMET, in black).

	GRASS	ArcGIS	CitySim	SimStadt	Ladybug
YGSR	2077	1672	1602	1268	1635
Diff.	24.5%	0.2%	-4.0%	-24.0%	-2.0%
RMSE	1.58	0.91	0.20	1.26	0.11

Table 2. Yearly global solar irradiation (YGSR) in [kWh/m^2/a], difference and RMSE compared to the INMET reference data.

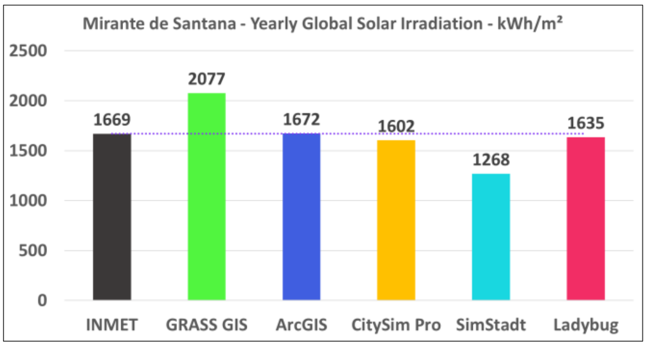

Figure 8. Yearly global solar irradiation values from the simulation tools and the meteorological station (INMET). Purple dot line is the reference annual value.

as reference (i.e. 1669 kWh/m^2/annum). Results are presented in Table 2. In order to better quantify the temporal variation over the whole year, the RMSE values were computed as well. Table 2 and Figure 8 present the results. In terms of yearly global solar irradiation, ArcGIS has the closest value with a difference of 0.2%, followed by Ladybug and CitySim with -2.0%and -4.0%, respectively. GRASS GIS and SimStadt yield much larger deviations, 24.5% the former and -24.0% the latter. However, looking at the RMSE, it is clear that Ladybug and CitySim deliver results that are much closer to the ground truth data – as their respective RMSE values of 0.11 and 0.20 reveal.

4.2 Comparison of aggregated monthly values

In order to better quantify the effect of the seasonal variation, the same type of analyses were carried out at a finer temporal resolution. All values were recomputed on a monthly basis. Results are presented in Table 3. In terms of monthly differences of global solar radiation, Ladybug always underestimates the ground truth values, but the maximum deviation is -4.5% in the winter month of June. CitySim has similar but slightly worse results, with the biggest deviation of 7.1% during the same month. SimStadt, ArcGIS and GRASS GIS score much worse, whereas the latter two have negative *and* positive deviations, reaching the 48.8% of GRASS GIS in the summer month of January. Looking at the RMSE values, Ladybug has the absolute lowest value in February (0.03), and its maximum (0.16) in the July-August months. CitySim, again, scores similar but slightly worse values (0.09 in February, and 0.29 in August-September). ArcGIS, SimStadt and finally GRASS GIS all have generally higher values of RMSE during all months, with GRASS GIS reaching the highest value of all (2.55) in January.

4.3 Discussion

From a purely numerical point of view it can be affirmed that the results from GRASS GIS and ArcGIS might be due to the fact that they use their own radiation model and no weather data file can be provided by the user. If, on the one hand, this limited flexibility has the advantage of being simple to use, as the user does not need to provide too many input datasets, on the other hand, this comes at the cost of accuracy of simulation results. In the case of the vector-based simulation tools, SimStadt, CitySim and Ladybug use in our tests all the same weather data, albeit in different file formats. If CitySim and Ladybug deliver very similar results, this is not the case of SimStadt, which tends to underestimate the irradiation values all over the year. The reasons need further investigation (and, possibly, further discussions with the developers). One of the major differences, at least in terms of input data, is that SimStadt uses only buildings and no other urban features during the simulation process. Therefore, looking at both Table 2 and Table 3, Ladybug delivers the best estimation of the solar irradiance in our study area, closely followed by CitySim. On the opposite side, GRASS GIS and SimStadt deliver the results with the largest differences on yearly basis, and overall

	GRASS GIS		ArcGIS		CitySim		SimStadt		Ladybug	
	Diff.	**RMSE**	**Diff.**	**RMSE**	**Diff.**	**RMSE**	**Diff.**	**RMSE**	**Diff.**	**RMSE**
January	48.8%	2.55	20.4%	1.13	-1.9%	0.10	-10.4%	0.57	-0.6%	0.04
February	34.5%	1.92	7.1%	0.57	-1.6%	0.09	-16.4%	0.95	-0.4%	0.03
March	22.4%	1.34	0.2%	0.45	-3.3%	0.18	-22.2%	1.17	-1.1%	0.07
April	10.4%	0.71	-12.8%	0.75	-4.8%	0.23	-35.7%	1.71	-2.6%	0.13
May	4.6%	0.43	-20.6%	0.85	-5.3%	0.20	-34.7%	1.41	-3.8%	0.15
June	1.9%	0.27	-25.2%	0.89	-7.1%	0.24	-35.6%	1.25	-4.5%	0.15
July	-5.5%	0.40	-26.9%	1.05	-6.9%	0.26	-39.1%	1.51	-4.2%	0.16
August	6.8%	0.69	-15.9%	0.80	-6.5%	0.29	-40.9%	1.85	-3.6%	0.16
September	14.0%	1.07	0.5%	0.81	-5.8%	0.29	-32.5%	1.72	-2.7%	0.14
October	36.8%	1.89	15.2%	0.87	-3.2%	0.16	-15.8%	0.83	-1.4%	0.08
November	46.4%	2.41	20.6%	1.16	-2.4%	0.13	-9.4%	0.53	-1.0%	0.06
December	44.9%	2.47	13.6%	1.24	-2.0%	0.12	-11.7%	0.68	-0.8%	0.05

Table 3. Differences and RMSE of simulated monthly solar irradiation values with respect to the INMET weather station.

high values of RMSE both on yearly and monthly basis. Somewhere in-between lies ArcGIS, both in terms of yearly and monthly cumulative global solar irradiation values, and corresponding RMSE values.

Choosing the "best" software tool is, however, more complex whenever other complementary aspects are also considered, such as the type and quantity of input data requirements, the possibility to automatise the process, and, finally, the computation time. For example, regarding both raster-based software ArcGIS and GRASS GIS, their relatively basic input data requirements make them "easier" to use as they need just a DSM to be run. However, one of their intrinsic limitations resides in the capability to simulate only 2.5D surfaces, which exclude vertical walls and similar geometries. This limitation might be acceptable at latitudes close to the equator, but at higher latitudes the lower height of the sun during the winter months makes the installation of solar panels on walls particularly interesting, therefore it is valuable to be able to include them in the simulation as well. Finally, it is worth mentioning the computation time difference between GRASS GIS and ArcGIS, the former taking more than twice as long as the latter.

When it comes to the vector-based simulation tools, Ladybug and CitySim offer nearly comparable and accurate results, as already observed. From the perspective of a GIS user, Ladybug has an initial steeper learning curve as it requires to learn a "new" environment (e.g. Grasshopper and how Ladybug is embedded in it) and to prepare/convert the input data. This is of course understandable because the underlying Rhinoceros 3D/Grasshopper platform is actually designed for the CAD world and not for GIS data. However, given the overlapping "urban scale" of both CAD and GIS communities, an easier way to convert and import GIS data (especially in the case of 3D city models) would be very welcome from the user's perspective. On the other hand, SimStadt and CitySim share common properties from the usability point of view: both support CityGML, both are provided with a GUI, and both can be run also via shell scripts. The main disadvantage of SimStadt resides in the temporal resolution of the results, i.e. monthly. Having access to hourly results – as we did for this work – requires opening the temporary files created during the simulation. Since no GMLID of the simulated surfaces is stored in such temporary files, identifying the corresponding RoofSurface "supporting" the weather station and retrieving the associated simulation values was done manually thanks to the very simple geometry of the "artificial" building. However, this is not a reliable *modus operandi* in case of more complex geometries. Regarding the simulation time, SimStadt is the fastest software in our tests, closely followed by Ladybug. CitySim is instead by far the slowest. The reason could be that the simulation considers *all* input surfaces, unlike Ladybug that computes solar irradiance only on chosen objects and skips all others, using them just as shadow-casting objects.

5. CONCLUSIONS AND OUTLOOK

This paper has presented a comparative study of five commonly used software tools to estimate solar radiation in the urban context. Two tools work on a raster base (GRASS GIS, ArcGIS), the remaining three require a 3D vector-based representation of the urban scene. Care was taken to collect spatial and non-spatial data (mostly weather data from the reference meteorological station) and to prepare them in a way that all software tools use – as much as possible – the same input data. Besides the accuracy of the results, the software tools have been evaluated also in terms of data requirements, simulation time, and overall usability. Although their different capabilities, for example in terms of supported features (buildings, terrain, vegetation) play surely a role in the accuracy of the overall results, this work has intentionally tried to evaluate the tools also from a user perspective, in that the capabilities of each software tool have been tested and compared starting from the same available datasets.

From the discussion presented in section 4.3, it can be concluded that the decision upon which software tool to use depends on several factors, as accuracy of the results is indeed fundamental but must be weighed with several other parameters. Depending on available data we may suggest a raster-based approach for preliminary studies, best however when supported by availability of ground truth data to calibrate the simulation results. For more accurate studies, and depending on the size of the study area, we'd suggest Ladybug or CitySim. In this case, however, acquiring some basic knowledge of CityGML might be beneficial, although both software tools allow importing other vector formats as well. This is however not the case of SimStadt, which works exclusively with CityGML data, but has the fastest simulation time.

As mentioned at the beginning of this paper, this work is part of a larger study that aims to investigate tools and methodologies to estimate the PV potential in so-called informal settlements (e.g. favelas). Given the additional challenge represented by the general lack or scarcity of digital information in those areas, the insights provided by this work may contribute to better evaluate how feasible this type of applications are in such urban contexts – or to find a reasonable trade-off between accuracy of the results and overall feasibility of such studies.

In the near future, some improvements and further tests are planned. First, we would like to add other tools to the comparison (Saga GIS, for example, being on the top of our list). We may optionally extend the test to other similar (proprietary) software solutions, but for the time being the focus is on free and open-source solutions. We also plan to conduct the tests on other, more recent machines, testing the software also on Macintosh and Linux configurations (as far as supported). At the same time, we plan to carry out the same set of tests in (at least) another country/place ideally at higher latitudes. A possible candidate could be a test area in the Netherlands, where some preliminary work has already been carried out (León-Sánchez *et al.*, 2021).

ACKNOWLEDGEMENTS

The authors would like to thank Jérôme Kämpf, Giuseppe Peronato (Idiap), and Eric Duminil (HFT Stuttgart) for the fruitful discussions and their hints on Citysim and SimStadt, respectively.

REFERENCES

Agugiaro, G., Nex, F., Remondino, F., De Filippi, R., Droghetti, S., Furlanello, C., 2012, Solar radiation estimation on building roofs and web-based solar cadaster. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, I(2), pp. 177-182.

ArcGIS Solar, 2021, https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/area-solar-radiation.htm (accessed on 10 January 2021)

CitySim, 2021, http://www.kaemco.ch/download.php (accessed on 10 January 2021)

CitySim wiki, 2021, https://github-wiki-see.page/m/kaemco/CitySim-Solver/wiki/The-Climate-File (accessed on 10 January 2021)

Bieda, A., Cienciała, A., 2021, Towards a Renewable Energy Source Cadastre – A Review of Examples from around the World. *Energies*, *14*, 8095. https://doi.org/10.3390/en14238095

Desthieux, G., Carneiro, C., Susini, A., Abdennadher, N., Boulmier, A., Dubois, A., Camponovo R., Beni, D., Bach, M., Leverington, P., Morello, E., 2018. Solar cadaster of Geneva: a decision support system for sustainable energy management. From Science to Society. Progress in IS, Springer, pp. 129-137.

Echlouchi, K., Ouardouz, M., Bernoussi, A.S., 2017, Urban Solar Cadaster: Application in North Morocco. *I*nt. Renewable and Sustainable Energy Conference (IRSEC), pp. 1-7, https://doi.org/10.1109/IRSEC.2017.8477342

Feitosa, P., de Lima, A.J.S., Cukierman, H., 2020. Digital Exclusion of Favelas from the Solar Map of Rio de Janeiro. *IEEE Technology and Society Magazine*, *39*(1), pp.30-85.

Geosampa, 2021, Mapa Digital da Cidade de São Paulo, http://geosampa.prefeitura.sp.gov.br/PaginasPublicas/_SBC.aspx (accessed on 10 January 2021)

Giannelli, D., 2021, Solar analysis on buildings of favelas in São Paulo to estimate PV potential. MSc thesis of Delft University of Technology, http://resolver.tudelft.nl/uuid:05ff635a-8113-4d65-aae1-bc0c727831d8

Hassaan, M.A., Hassan, A., Al-Dashti, H., 2021, GIS-based suitability analysis for siting solar power plants in Kuwait, The Egyptian Journal of Remote Sensing and Space Science, 24/3/1, pp. 453-461, ISSN 1110-9823, https://doi.org/10.1016/j.ejrs.2020.11.004

INMET, 2021, Instituto Nacional de Meteorologia https://www.gov.br/agricultura/pt-br/assuntos/inmet (accessed on 10 January 2021)

Ladybug, 2021, https://www.ladybug.tools/ladybug.html (accessed on 10 January 2021)

León-Sánchez, C., Giannelli, D., Agugiaro, G., Stoter, J., 2021, Testing the new 3D BAG dataset for energy demand estimation of residential buildings. *ISPRS Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVI-4/W1-2021, pp. 69–76.

Maxwell, E. L., 1987. A quasi-physical model for converting hourly Global Horizontal to Direct Normal Insolation. Solar Energy Research Institute, SERI/TR-215-3087, pp. 35–46.

Perez, R., Seals, R., Ineichen, P., Stewart, R., Menicucci, D., 1987. A new simplified version of the Perez diffuse irradiance model for tilted surfaces. Solar Energy, 39(3), pp. 221–231. https://doi.org/10.1016/S0038-092X(87)80031-2

r.sun, 2021, GRASS GIS r.sun module, https://grass.osgeo.org/grass78/manuals/r.sun.html (accessed on 10 January 2021)

Ramirez Camargo, L., Zink, R., Dorner, W., Stoeglehner, G., 2015, Spatio-temporal modeling of roof-top photovoltaic panels for improved technical potential assessment and electricity peak load offsetting at the municipal scale, *Computers, Environment and Urban Systems*, 52, pp. 58-69.

RioSolar, 2021, Solar Map do Rio de Janeiro, https://pcrj.maps.arcgis.com/apps/webappviewer/index.html?id=029ccb875ca24c209a8572e70a4b6b19 (accessed on 10 January 2021)

Salazar Miranda, A., Du, G., Gorman, C., Duarte, F., Fajardo, W., & Ratti, C. (2021). Favelas 4D: Scalable methods for morphology analysis of informal settlements using terrestrial laser scanning data. ArXiv, abs/2105.03235.

SimStadt, 2021, https://simstadt.hft-stuttgart.de (accessed on 10 January 2021)

SoDa, http://www.soda-pro.com/help/general-knowledge/linke-turbidity-factor (accessed on 10 January 2021)

Viana-Fons, J.D., Gonzálvez-Maciá, J. , Payá,J., 2020, Development and validation in a 2D-GIS environment of a 3D shadow cast vector-based model on arbitrarily orientated and tilted surfaces, Energy and Buildings, 224, 110258.