



HAL
open science

An integrated GIS and robust optimization framework for solar PV plant planning scenarios at utility scale

Benjamin Pillot, Nadeem Al-Kurdi, Carmen Gervet, Laurent Linguet

► To cite this version:

Benjamin Pillot, Nadeem Al-Kurdi, Carmen Gervet, Laurent Linguet. An integrated GIS and robust optimization framework for solar PV plant planning scenarios at utility scale. Applied Energy, 2020, 260, 10.1016/j.apenergy.2019.114257 . hal-02420300

HAL Id: hal-02420300

<https://hal.umontpellier.fr/hal-02420300>

Submitted on 19 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An integrated GIS and robust optimization framework for solar PV plant planning scenarios at utility scale

Benjamin Pillot^{a,*}, Nadeem Al-Kurdi^a, Carmen Gervet^b, Laurent Linguet^a

^aESPACE-DEV, Univ Guyane, IRD, Univ Montpellier, Univ Antilles, Univ Réunion, 275 route de Montabo, BP 165 - 97323 Cayenne Cedex, Guyane Française

^bESPACE-DEV, Univ Montpellier, IRD, Univ Antilles, Univ Guyane, Univ Réunion 500 rue Jean François Breton, 34090 Montpellier, France

Abstract

Today, the overall goal of energy transition planning is to seek an optimal strategy for increasing the share of renewable sources in existing power networks, such that the growing power demand is satisfied at manageable short/long term investment. In this paper we address the problem of PV penetration in electricity networks, by considering both 1) the spatial issue of site selection and size, and 2) the temporal aspect of hourly load and demand satisfaction, in addition with the investment and maintenance costs to guarantee a viable and reliable solution. We propose to address this spatio-temporal optimization problem through an integrated GIS and robust optimization model, that allows handling of the ubiquitous dependencies between resource and demand time variability and the selection of optimal sites of renewable power generation. Our approach contributes to the integration of the multi-dimensional and combinatorial aspects of this problem, gathering geographical layers (*regional* or *national* scale) and temporal packing (*hourly* time stamp) constraints, and cost functions. This model computes the optimal geographical location and size of PV facilities allowing energy planning targets to be met at minimal cost in a reliable manner. In this paper, we illustrate our approach by studying the penetration of large-scale solar PV in the French Guiana's power system. Among the results, we show for instance that: 1) our approach performs geographical aggregation with real contextual data, i.e. balances the intermittency of RE sources by spreading out the corresponding installations (location + size) across the territory; 2) the total installed PV capacity can be doubled by removing the 35 % penetration limit on intermittent power without exceeding hourly demand; 3) the safest investment scenario is below 30 MW of new PV facilities (≈ 45 M€ and 2 plants), though it is theoretically possible to install up to 45 MW (>120 M€ and 11 plants).

Keywords: GIS; robust optimization; spatiotemporal dimensions; solar PV; energy planning; site selection

Nomenclature

B_i	Boolean variable equal to 1 if site PS_i is selected, 0 otherwise
$Ccap_i$	Capital cost for implementation of a new PV power plant (€)
$Ccon_i$	Connection cost for each new PV plant, transmission lines and substation (€)
$Clan$	Transmission line unit cost (€/m)
Cop_i	Annual fixed operational cost per new PV plant (€)
$Csta$	Substation power unit cost (€/kW)
$Csta_i$	Capital cost for new substation (€)
Dem_h	Estimated hourly power demand (kWh)
Dg_i	Minimum distance from the grid to the centroid of a candidate (m)
$Eint_h$	Current hourly production from intermittent sources (kWh)

*Corresponding author

Email addresses: benjamin.pillot@ird.fr (Benjamin Pillot), nadeem.alkurdi@ird.fr (Nadeem Al-Kurdi), carmen.gervet@umontpellier.fr (Carmen Gervet), laurent.linguet@univ-guyane.fr (Laurent Linguet)

$E p_h$	Current hourly production from non-intermittent sources (kWh)
h	hour of the year
i	Site index
m	Final number of solar PV plants
n	Number of candidate sites derived from GREECE
P	Set of selected candidate sites
$P p v_{h,i}$	Estimated hourly production per PV unit for each site (kWh/m ²)
PS	Set of potential candidate sites
R_{it}	Resource time series corresponding to potential site PS_i
SA_i	Surface to consider for a candidate site PS_i that is selected (m ²)
$Smax_i$	Maximum area for each candidate site (m ²)
$Smin$	Minimum area for each candidate site (m ²)

1. Introduction and related work

Enhancing the share of renewable-based generation in power networks is today one of the main objectives in countries' energy planning policies [1]. The need for mitigating our carbon footprint in combating climate change has initiated and accelerated this growth over the past 30 years. Many countries, political and economic unions or organizations have therefore established their own renewable energy policy frameworks [2] (e.g. in the European Union [3]). Typically, various targets have been implemented by government policies in order to increase the share of renewable energy (RE) sources in power networks [2]. But integration of those facilities, such as solar PV [4] or wind turbines [5], is not straightforward: geographical dispersion and intermittency have often imposed grid adaptations in order to avoid causing instability [1], and forecasting the energy potential can be a modelling challenge on its own in the absence of expert knowledge [6]. It therefore requires to develop specific energy planning strategies and tools allowing objectives to be met without jeopardizing the existing infrastructures.

Most of RE sources vary with time (intermittency) and space (geographical dispersion) due to climate patterns, weather conditions, day-night alternation, terrain features, etc. Depending on the location where corresponding facilities are set up, resulting power supply will affect electrical network's dynamics [7], and thus the balance between production and demand. Strategic energy planning must therefore account for geographical restrictions as well as the resource potentials whose fine temporal patterns are location dependent. In other words, a key question for us is: what is the optimal *geographical distribution* of renewable power stations allowing specific energy and economic targets to be met given the *spatio-temporal variability* of the resources at hand? By geographical distribution we mean a set of power stations established in a region and defined by their location, type, size and capacity. In particular, for investors and decision makers, seeking the optimal land size is an essential component in terms of investment and energy production. Regarding energy planning targets, common ones may be to reach 100 % of renewable-based electricity within 10 years [8], or to balance power from intermittent energies at any time. However, we may also propose various scenario solutions of energy planning strategy by tuning a set of possible parameters and forecast data [1]. Those scenarios might take into consideration different combinations of resources, various land management strategies, growth of the population, climate change impacts, etc.

Today, strategic energy planning relies on system modeling tools with a strong growths of systems involving spatio-temporal modeling at different scales and with different objectives and study purposes [9]. In the case of identifying suitable sites for intermittent RE sources (solar, wind) parks, which we are interested in, there is a focus on local scale study approaches summarized below.

At local scale (urban, roof tops), different studies have combined geographic information systems (GIS) with the time dimension in order to retrieve the location and size of small RE power system units. Some approaches such as, Kucuksari et al. [10] combine a GIS approach with linear optimization in order to plan long-term PV installation on a campus environment. Potential sites are selected based on a GIS assessment of yearly radiation and building's rooftop suitability. Optimal location and size of the PV systems to be installed in the campus over 20 years are then retrieved so that yearly PV generation constraint target is reached and the total system long-term net profit is maximized. This approach considers an annual time scale, and does not take into account the temporality of satisfying the electricity demand. Also, Mavromatidis et al. [11]

propose a step-by-step framework in order to meet the local electricity demand of a village with solar PV and storage. In their problem definition, the generated energy could meet the aggregated hourly electricity demand, be stored in batteries or exported to the national grid. The hourly solar radiation profiles for each building are derived from a DEM approach (ArcGIS Solar radiation tools). They also create and aggregate synthetic electricity demand profiles to get the total hourly demand of the village. Finally, optimal building's rooftops for PV panels installation are determined by minimizing total costs and maximizing renewable share through a mixed-integer linear program (MILP). In this approach, there is no issue of land sizing nor of handling the data uncertainty in future costs, demand or resource potentials. A static rooftop planning approach is considered.

Similarly but at a municipal scale, Ramirez-Camargo and Stoeglehner [9] present a spatio-temporal GIS model for RE planning, including solar PV. Potential sites are selected based on rooftop areas also, that have been previously identified and classified according to PV usage. Size and location of those sites are therefore directly defined by the rooftop-based geometric constraints of the problem. The authors then use a decision tree in order to find the best match between corresponding RE generation and the electricity demand of the municipality. Relative costs are not considered nor uncertain data for forecasting trends in energy and demand.

Recently, Alhamwi et al. [12] have developed a GIS-optimization to match electricity demand by adding storage installation at urban level. The authors use a cellular approach whereby the city or district is divided into several interconnected *energy cells* in which demand and supply are balanced at hourly scale. For each energy cell, new power and storage are added so that total system costs are minimized and local electricity demand is met, according to different scenarios. Results show optimal storage capacity and configuration of power supply systems for each energy cell. The authors finally propose a post-processing GIS-based heuristic approach in order to achieve sizing and siting of storage elements. In that approach, sizing and siting do not directly intervene in the simulation and optimization. The authors rather propose a systemic method combined with a post-processing heuristic strategy that attempts to optimize placement of given RE capacity amounts within each sub-system (i.e. energy cell).

Regarding urban districts, Huang et al. [13] propose a GIS assisted method to assess optimal location and number of renewable powered electric vehicle charging stations, with respect to existing charging stations and solar PV generation. The authors first use GIS techniques to discretize a city district into a mesh, and then evaluate roof-mounted solar PV yearly generation within each grid cell according to available roof area. Partition theory and genetic algorithms are then used respectively to generate different alternatives for distributing charging stations among locations in the city district and to maximize global coverage area for each design. Feasible design alternatives meeting user-defined coverage ratio are finally chosen and the one with the minimum life cycle cost is selected as the optimal solution. The authors have applied their method to the Kowloon district of Hong Kong. In that case, the authors use static solar radiation data combined with mesh-based spatial optimization. Siting restrictions for implementing charging stations are not considered due to the very small size of the devices.

Another set of approaches seek to match power generation and energy demand at utility scale. For instance, applied to Malaysia, Sabo et al. [14] propose an analytical model for assessing implementation of utility-scale solar PV plants into the power grid. Their study provides a set of temporal and spatial factors such that matching of power generation from solar PV plants with the global energy demand is facilitated. After having first selected optimal sites (size + location) based on a GIS spatial selection, they compare power generation from those sites to electricity demand on an hourly basis through a set of static indicators. It virtually allows them to evaluate the impact of increasing levels of PV penetration onto the power network. In this case, matching static indicators are derived from spatially pre-selected sites, i.e. corresponding sites (spatial dimension) are not selected from their match (temporal dimension) with the electricity demand. Hence site selection does not take into account the existing interdependence between spatial and temporal variability of both the solar radiation and the energy demand satisfaction.

In a different line of work and objectives, bottom-up engineering approaches focus on time simulation and optimization of given energy system configurations. Their main objective is to guide *energy policy road map* often at a national scale and longer time horizon [1]. This holistic approach gives a significant insight into the potential contribution of RE sources [15], but does not aim at identifying physical parcel locations. Instead, these studies determine optimal RE capacity amounts for the whole system (i.e. region) or sub-systems (i.e sub-regions) according to given objectives (e.g. maximizing RE generation and minimizing net present cost) [16].

An insightful example of guiding energy policy road maps is the recently published study by Zappa et al. [17]. In this paper, the authors look at the feasibility of a 100 % renewable European power system by 2050. Europe is modeled as a single integrated power system including sharing of capacity between countries, and seven scenarios for a fully RE-based power

system are analyzed. Intermittent resources (wind, solar) are estimated using available profiles from ERA-interim. Hourly simulation is run at one week per month for year 2010 in order to mitigate CPU time. A MILP optimization approach is proposed to minimize the net present cost of the whole system over lifetime.

Luz and Moura [18] have used a multi-objective non-linear optimization framework to plan scenarios for the 100 % renewable transition road map for Brazil. Their model aims at minimizing the net present value of total system expansion costs and maximizing complementarity between demand and generation among regions and RE sources, at hourly scale. Solar irradiation and wind speed were taken from 27 automatic weather stations, and water inflow was retrieved from 10 hydropower. To assess complementarity between regions, the country's territory is beforehand divided into 3 to 5 regions for each resource. Their results show how wind and solar PV integration may reduce hydro dependence and reservoir requirements.

At a national scale, GIS-based studies, often combined with different multi-criteria decision making (MCDM) methods such as analytical hierarchy process (AHP) [19], fuzzy logic [20], etc. are particularly used to retrieve suitability maps. Those maps allow suitable areas for RE deployment to be identified and therefore give decision makers the necessary knowledge about where to invest. Typically, these studies first derive potential areas from an overlay analysis, by removing constraint and buffer layers from a given territory [21]. A refinement is then performed using multi-criteria analysis, in which relevance of the available zones is ranked according to predefined weighted decision criteria [22]. Criterion selection and ranking is mostly based on literature consensus and opinion from local expertise [23]. As these studies are not focused on energy demand management, the methodology is based on static data, i.e. resource temporal variation is not considered, and energy demand satisfaction is not taken into account. It does not aim at providing the exact location and size of the sites whose temporal power generation satisfies the hourly demand according to predetermined scenarios. Instead, it generates a cartography of the areas where RE source development appears technically and economically suitable with respect to selected criteria. We describe below some of the latest studies.

Al Garni et al. [24] have achieved the first GIS-MCDM study in Saudi Arabia for selecting the best sites for large-scale solar PV plants. They first derive potential areas from a restriction overlay analysis, and then achieve AHP to weigh the seven criteria they selected, that is 1) yearly solar irradiation and air temperature, and 2) slope, land aspects, distance to urban area, distance to roads distance to power lines. Accordingly, they generate a land suitability index (LSI) to assess the relevance of the potential sites through 5 classes, from least to most suitable.

Applied to the country of Nigeria, Ayodele et al. [25] analyze wind farm selection by combining GIS and multi-criteria type-2 fuzzy analytic hierarchy process (IT2-FAHP). The use of fuzzy sets in addition to AHP is achieved to represent expert's linguistic judgement in order to address uncertainty in decision making. Yearly averaged wind speed at 10 m hub height was retrieved from 28 ground stations. Most significant selected criteria are respectively wind speed (50 % weight), proximity to grid lines (24 %) and slope (17 %). Suitability is eventually classified according to a score from 0 (not suitable) to 4 (extremely suitable). A suitability map is finally computed based on the previous approach and a GIS-based constraint map where unsuitable areas have been removed.

Yang et al. [21] propose an innovative GIS-based study over China, where they bridge the gap in the scientific knowledge on distribution and size of large scale PV potential under geographical and technical constraints in the country. From geographical constraints, the authors derive suitable areas for utility-scale solar PV plants and then use land conversion factors under different geographical and technical conditions to estimate PV technical potential at high resolution. They also assess the life-cycle CO₂ reduction potential from the use of large-scale solar PV in China.

Compared to other studies, Also, Jung et al. [26] provide topographic and shading effects criteria into the selection of suitable sites on highway fill slopes from digital numerical maps. The authors aggregate solar radiation over the year, estimate shaded areas from contour-based topographic data, and rasterize the highway network so that high solar energy grid cells within fill slopes are finally clustered into suitable sites. They apply their method near Gongju in South Korea to rank 10 potential installation sites.

In summary, studies at urban scale are not implementable at broader scale as it is, not only regarding the predefined geometric constraints of the problem (rooftop areas) but also according to the new spatiotemporal scenarios and topics the utility scale brings out, such as climate and weather variations throughout the territory, intermittence (no storage), terrain issues (slope, land use, etc.) or even the size of the study area. On the other hand, widely used bottom-up engineering models allow techno-economic feasibility of energy system configurations to be assessed, and aim at providing optimal RE capacity amounts per region (system) or sub-regions (sub-systems), with respect to given objectives (for instance a 100 % renewable-based power system). Their main objective is to guide energy policy road maps at global scale and at long time horizon, hence they do not link temporal scenarios with the geographical constraints and costs of a given territory. Finally, at the national scale, GIS-based MCDM approaches address the issue of giving information to the decision makers about suitable locations

for RE source development. As their objective is not related to the energy demand side, these studies are based on static resource data and do not aim at determining optimal scenarios of geographical sites for future RE systems. Finally it is also important to note that in all these studies (local scale, bottom-up, GIS MCDM), handling of the data uncertainty in future costs, demand or resource potentials is generally not considered.

Essentially, to date we are not aware of computational approaches that tackle the spatiotemporal optimization problem consisting of identifying the best parcels that increase solar energy penetration into the network at minimal cost, while satisfying a region or country's specific constraints (terrain, resource, infrastructures, etc) and related costs. Hence we propose an integrated computational approach that handles this problem in terms of spatial placement and a resource planning problem. It addresses both land site selection and sizing from geographical constraints, and resource planning to satisfy the demand at minimal cost from an hourly temporal pattern viewpoint. This approach stands for a novel framework based on GIS spatiotemporal data and constraint processing, connected to a robust optimization (RO) knapsack model to plan renewable energy scenarios. Since our problem looks at future strategic planning most data at hand are uncertain (costs, demand, radiation potential), a robust approach is proposed to mitigate uncertainty of anticipated forecast values. The combination GIS-RO generates an excellent decision support system that allows for the planning of parcel scenarios (locations and areas) that will best increase the RE power into the network at minimal cost, according to the decision maker risk adversity. A GIS can handle very large volumes of data, including remote sensing images for solar radiation indicators, land use maps, and various networks maps (electrical, roads, water). The application of global and multi-layers geographical constraints and various control parameters allows for an effective deterministic pruning of the region, to determine suitable candidate parcels, and their relevant properties without impairing the optimization problem.

Summary of contributions. Our contributions can be summarized as follows: 1) The design and implementation of the GREECE prototype, *Geographical REnewable Energy Candidate Extraction* that tackles the extraction of potential PV sites over a regional or national land, by computing sets of polygons (potential RE areas) from different constraint layers (protected areas, urban zones, etc.) and control parameters (buffers, distance and surface thresholds); it also characterizes the region's spatiotemporal features and associated costs by providing for each candidate parcel *de-spatialized* attributes (shortest distance to the grid, maximum area, shape, average slope, land use, etc.) and temporal resource patterns; 2) the specification and implementation of a robust optimization system and Pareto frontier, that derives from these candidates different scenarios of optimal RE sites, satisfying hourly energy demand and overall photovoltaic (PV) grid penetration constraints, geographical network connection and installation costs, such that the RE production is maximized at minimal cost. Given the uncertainties present in all the forecast data (radiation measures, energy demand, RE costs per KM², etc.) we propose a robust optimization model that offers a reliable modeling of uncertain data from real data sets, and enables different solution scenarios based on the decision maker's attitude to uncertainty.

To illustrate our approach, we apply our model to French Guiana. According to the 2015 *Energy Transition Act*, France's energy policy has included the following targets for overseas regions: 50 % of renewable in final energy consumption in 2020, and full energy self-sufficiency by 2030 [27]. In French Guiana, it is hence planned to double or triple the installed capacity of biomass, wind, solar PV and small hydro between 2018 and 2023 [27]. French Guiana typically combines significant 1) geographical constraints with most of the landscape made up of primary forest we must protect [28], and 2) power constraints with low capacity for integration of intermittent RE sources due to a small isolated electrical network [29]. Reaching department's renewable share targets cannot therefore be decoupled from land management planning. In this paper we focus particularly on solar PV, as it shall share 75 % of the intermittent capacity and one third of the whole renewable power to be installed by 2023 [27]. Eventually, by using our integrated model, we show how and where large-scale solar PV should be exploited in French Guiana to ensure balance between landscape preservation, grid stability, and cost effectiveness.

This paper is structured as follows: we define the general problem and our approach to solve it in section 2; we then explain in detail the two parts of our integrated model in section 3 (GIS) and section 4 (optimization); the case study (French Guiana) for which we have applied this framework is presented in section 5; finally, we discuss the corresponding results and present the further developments in section 6.

2. General problem and proposed approach

This paper considers the following strategic energy planning problem: How to determine *location and size of the best sites* for solar PV plant installation 1) at utility scale (regional or national), such that 2) the energy supply into the grid satisfies the regulation for intermittent sources, and 3) the forecast hourly electricity demand is satisfied 4) at minimal cost, for a future

horizon 2030. Meanwhile, this problem has 5) no predefined geometric boundaries for site selection (such as rooftop areas, administrative boundaries, etc.), and should allow for 6) risk adversity and safest investment scenarios to be known to the decision maker.

A key element of our contribution is that we take into account the interdependence between the temporal and spatial variability of the solar radiation, thus we characterize the intermittence in a three dimensional scale (temporal radiation patterns distributed at a regional scale). If we ignore the temporal dimension, the problem is reduced to the selection of suitable sites according to average and global penetration load into the grid network. The variability of radiation in time and space does however impact the hourly penetration over cost ratio (proximity to the grid, site dimensions) as well as the maximum PV penetration per hour (as an intermittent RE source). Hence the proposed framework tackles these essential and interdependent dimensions of the problem through a two-stage approach: one GIS based model, and a robust optimization model.

One common challenge when coping with spatiotemporal analysis of RE sources is the need for high computational capacities [9]. We have handled this issue by favoring a procedural approach where the GIS model reduces the search space passed on to optimization. Essentially, our approach takes benefit of the existing geographical constraints to discretize the space and eventually mitigate the set of feasible solutions. It ensures the global complexity of the problem to remain circumscribed: for the case study (see section 5.1.1), execution CPU time for the GREECE module was 460 s (candidate sites extraction) + 340 s (solar GHI extraction + time series generation), handled as a one off spatial placement preprocessing; CPU time of each optimization run was then between 23 and 47 s depending on the bound set on the constrained objective function.

The GIS-model, we call GREECE, takes as input different spatial layers of data, a set of parameters that can be tuned by the user, and generates a set of candidate sites with their respective features, including maximum associated PV capacity, attributes (slope, shape, land use) and the hourly resource time series distribution per candidate site. The solar resource is identified from remote sensing radiation maps, and varies in time t and space. Thus for each given tuple of coordinates¹ (x, y, t) within the territory corresponds one value r of the resource. The objective of GREECE is, by considering all data and geographical constraints, to prune and partition the territory into potential areas for PV sites alongside their corresponding solar radiation, as hourly time series. That is, eventually, to each element PS_i of the set of potential sites PS is associated one time series of resource values R_{it} . This leaves the issues of determining optimal sites and size to the optimization module according to hourly power demand constraints (tailored for PV and for global coverage) and the objective functions at hand.

The optimization model is a robust optimization model [30], and has been published recently in a computer science conference [31]. This means that all forecast data are represented as intervals instead of a stochastic distribution [32]. The satisfaction of the interval linear constraints involving such data (energy demand projections and costs projections), guarantees that the solutions produced are such that consistent values that hold for such constraints are kept, and only inconsistent ones are removed. This model takes as input the set of candidate sites and their attributes from GREECE output, the hourly load penetration constraints, and global demand satisfaction constraints and derives an optimal set of sites with their computed size such that the constraints are satisfied using a Pareto optimization approach. The two objectives at hand are the maximization of the production and minimization of the total cost.

An energy conversion model might also be used between both GREECE and optimization model in order to rather retrieve system's output power from resource. In the case of solar PV, it consists in converting solar radiation into AC power with respect to endogenous features (PV cell type, PV field inclination and orientation, inverter's efficiency, PV rack mounting, etc.) and exogenous factors (ground albedo, wind speed, air temperature, diffuse fraction, etc.). Instead of using simplistic conversion from resource during optimization, the use of such an in-between model ensures accurate and realistic site selection.

The models are described in the following sections. Figure 1 depicts the architecture of the proposed framework to determine the best locations, size and capacities of PV sites for long-term installations in selected areas. This framework has been implemented and evaluated for the French Guiana region. The data layers for this region were provided by world databases [33], local administration (e.g. ONF, *National Forests Office* [34] and IGN, *National Geographic Institute* [35]), France's electric utility [36] (EDF, *Électricité de France*) and previous solar radiation studies from Albarelo et al. [37] and Fillol et al. [38]. All forecast data on demand, and installation and maintenance costs are based on standard projection data. PV site deployment for French Guiana is planned for the horizon 2030, to complement the current PV facilities, as well as hydro and biomass. A hybrid PV and biomass study is underway but we focus on the PV component in this article. The

¹Theoretically energies such as wind and solar also vary with altitude z but we will not consider it in this study.

implementation shows the effectiveness of accounting fully for the different temporal patterns of solar radiation in different areas, in this strategic planning of PV sites.

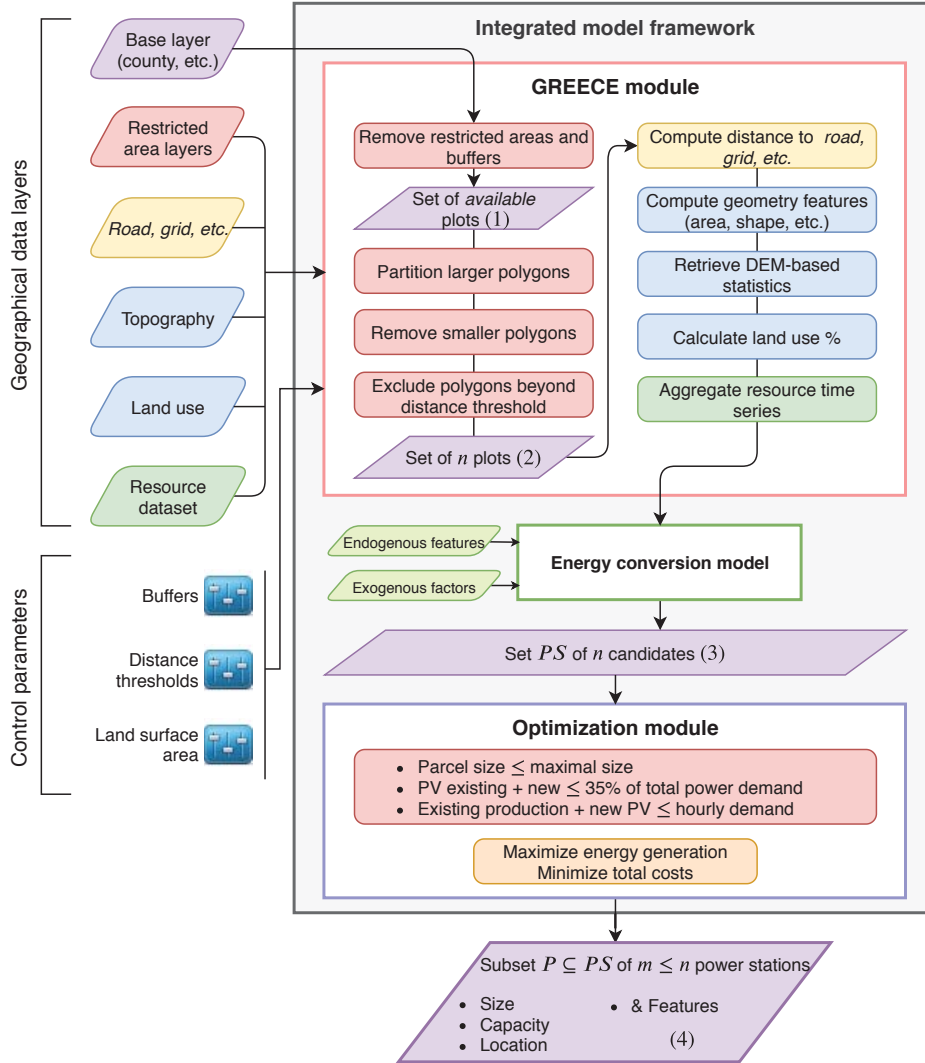


Figure 1: Schematic view of the proposed integrated GIS-optimization framework.

3. PV sites extraction and spatial data discretization problem: The GREECE model

The first stage problem computes a set of candidate geographical sites such that a set of geographical, topographic, and land usage constraints are satisfied, enriched with digitalized features to characterize each candidate site. We propose the GREECE model to achieve this goal, in a GIS environment, by handling the different spatial data layers sequentially. The computed candidate sites, or polygons, are then de-spatialized, meaning identified independently of their coordinates, and passed on with computed characteristic features of relevance to the optimization module. Typically, those features exhibit the geometry of the corresponding polygon, the land and terrain attributes, maximum energy load potential, and distances to

geographical objects of interest (such as road or grid networks). These characteristics may be used to derive cost values and as selection criterion in the optimization stage. The outcome is a set of available land plots with their respective features.

The GREECE model takes as entry 1) Geographical data layers (territory, constraints, networks, terrain, resource), and 2) Adjustable geometric parameters (buffers, distance thresholds, land surface area). Layer data are specific to a territory and have been previously obtained by any institution for the given territory (e.g. the IGN institute in France) and by remote sensing with respect to the resource. They depend on the territory's own particularities and might be selected with respect to expert opinion, planning objectives and literature. The set of range control parameters has been defined to allow some candidate extraction to be tuned according to the user own requirements. Different land management scenarios might in fact be implemented by playing with the corresponding parameter values. Indeed, depending on buffer, threshold and surface area values set by the user, the resulting set can differ, and be tailored to specific expert knowledge in a deterministic manner. This guarantees a deterministic pruning of inadequate sites, that reduces the complexity of the combinatorial optimization process while remaining consistent, i.e. not removing potentially valuable sites. GREECE is implemented in Python and has been developed using various open-source geographic and data structure libraries ([gdal](#) [39], [geopandas](#) [40], [numpy](#) [41], [rtree](#), [shapely](#), [networkx](#)).

3.1. Geographical data layers

As depicted in Figure 1, input data layers correspond to the boundaries of the study area (base layer), restricted areas, specific objects for which distance to resulting polygons must be computed (e.g. road, grid), terrain features (land use and topography) and the resource of interest (here solar radiation maps). Restricted areas (in red) stand for the land where facilities cannot or should not be established. Typically, those are urban areas, ecological zones, watercourses, military sectors, cultural heritage, etc. [24]. They may also represent zones too far from specific objects (e.g. electrical grid) [23]. Geographic elements for which distance must be computed (in yellow) can be of any kind but are generally related to connection and accessibility costs such as road and grid networks. Topography and land use (in blue) allow terrain to be characterized within each polygon. Finally, resource (in green) is a dataset of rasters or vectors defined over as many time steps as available. Typically, all these data might be retrieved from national and international geographic databases as well as specialized agencies.

3.2. Control parameters

Range control parameters allow different land management scenarios to be taken into account. Buffers surrounding geographic objects depend both on their type and on the kind of power station, e.g. wind [23], biomass [42] or solar PV [24]. Distance thresholds to given layers (e.g. road network, electrical grid) stand for the limit beyond which establishing a facility gets economically irrelevant and so corresponding plots must be excluded [23]. Depending on the values defined for buffers and thresholds, corresponding restricted areas will be more or less significant. Finally, land surface area allows different strategies for land management by setting minimum and maximum surface thresholds. Typically, resulting plots may be too small or too large with respect to the kind of power plants to be set up within. According to selected values, smaller and larger plots might respectively be excluded from the study and partitioned.

Though all control parameters affect the size of the set of available sites passed to optimization, buffers and land surface area mainly depend on existing regulations and on the size of selected applications. On the other hand, distance thresholds stand for economic relevance of the corresponding sites. Accordingly, when assigning values to those thresholds, we may either limit the set of potential sites, by considering smaller values, in order to reduce optimization time, but at the expense of accuracy (economic assumptions made *a priori*); or give more authority to optimization, by considering higher values and extending the set of potential sites, in order to enhance accuracy at the expense of computing speed.

3.3. Generating the set PS of suitable polygons

3.3.1. Pruning areas from the base and constraint layers

The GREECE model first computes the intersection between the territory *base layer* and *restricted area layers* using given buffer values. The computation relies on optimized methods to split largest polygons [43], and finally re-aggregates all contiguous results. The result is a set of available plots where facilities could be set up (step (1) in Figure 1).

3.3.2. Bounding plot sizes using tuned parameters

The model then integrates some *surface bounding strategies* as parameterized by the expert. There are two possible thresholds. Basically, when a surface area is below a minimum threshold specified by the user, the model will remove corresponding plots from the final set. Similarly, when the user stipulates a maximum area threshold, the model will split the corresponding plots according to *simple cut* or *partition* rules that can not be below the minimum threshold as well. The simple cut is similar to the method described previously for fast polygon intersection [43]. It does not consider how a parcel should be split but instead provides a set of square-like smaller parts whose area is necessarily below the maximum threshold.

The partition method is more complex but provides deeper insight into the surface management. This is the one we used. We apply graph partitioning theory through the METIS package originally developed by Karypis and Kumar [44]. A polygon is first split into smaller parts, either square-like polygons or hexagons, now representing a graph of interconnected nodes (sub-plots) weighted by their area. A k -way partitioning scheme is then performed so that the original plot is divided into k parts, each one being an aggregate of various nodes. By minimizing the number of edge cuts and forcing contiguous partitions [45], the model ensures aggregating parcels with correct geometry. As nodes are weighted by their areas (surface constraint), it guarantees that each parcel finally covers the area corresponding to the maximum threshold defined at first. The output is a set of n polygons of acceptable surface (step (2) in Figure 1). Thus, the total area of the final set of polygons is smaller or equal to the one computed in step (1).

Figure 2 illustrates this process for a 16-way graph partitioning of a 781 km² polygon into plots of 50 km², using an hexagonal honeycomb mesh. The algorithm derives 15 plots between 49 km² and 51 km² plus one (marked with an asterisk) which fills the leftover space (≈ 32 km²). The output is a finite set of polygons of acceptable surface (below the maximum threshold).

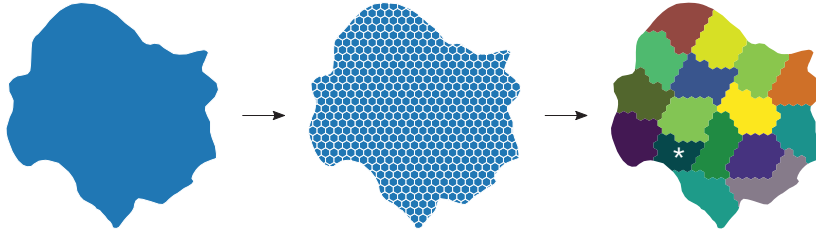


Figure 2: Surface partitioning using a honeycomb mesh

Note that, in this work, we have more particularly focused on *surface* management, but it should be noted that other strategies are possible. For instance, graph partitioning allows implementation of multiple constraints [45]. It is therefore possible to perform partitions meeting other objectives in addition to equal areas. Typically, we may want to divide the land with respect to similar terrain features or land use attributes. However, the number of constraints also limits the chance of reaching a satisfactory solution. It should therefore be used with caution while thoroughly knowing what kind of compromise we intend to meet, and not impact the optimization module with too restrictive candidate sites.

3.3.3. Removing polygons beyond distance thresholds

Based on distance computed from their centroid and thresholds specified by the user, furthest plots from specific objects such as road (accessibility costs) and electrical networks (connection costs) might be excluded.

3.3.4. Adding quantifiable features to the plots

The second part of the GREECE model enriches the set of candidate sites with additional features whose main purpose is to feed the optimization model such that all aspects involving a cost or resource potentials are handled in the optimization module. To do so we use the remaining spatial data layers. Distance from power facilities to specific geographic objects or locations gives information on related costs [46]. For instance, distance to the grid or to substations is convertible into connection costs and distance to the road into accessibility costs. GREECE computes minimal (euclidean) distance from potential candidates' centroid or boundary to any selected layer. In general, we shall prefer the centroid method, as it makes cost comparison more fair: every distance will be retrieved from one unique reference point.

For each polygon, the model will derive geometry (area, shape) as well as topography statistics, based on Digital Elevation Models (DEM), and land use share. Those characteristics might in fact decide of the *usability* of a parcel. As pointed out by Al Garni and Awasthi, terrain slope is for instance one critical criterion when regarding the installation of any power facility [47]. In the case of wind and solar, available polygon surface also limits how much nominal capacity might eventually be set up [24]. In the same way, final cost for establishing a power station might also be affected by land type and land use.

Regarding terrain, GREECE extracts the DEM-based distribution of the main descriptors (slope, aspect, elevation) and then computes corresponding mean and standard deviation (std). Typically, the std indicates how hilly a land plot is, while the mean slope and aspect tell us about its main inclination and orientation. Eventually, all these features (shape, land use and terrain) might be converted into the corresponding installation costs or used as exclusion thresholds in the optimization module.

Finally, GREECE computes how much energy (maximum) can be derived within each polygon. We will see in section 5 that in the case of solar PV in French Guiana, we have derived the resource from a dataset of satellite-based solar radiation maps. Typically, resource is aggregated within a given polygon and for each available time step by averaging all raster cell values it intersects with. In the case of very large parcels or rough spatial resolution, the result might also be weighted according to corresponding cell surfaces. If temporal resolution is too low, the model may use stochastic methods such as [48] or [49] for generating synthetic downscale data from average global records. The result is a time series of resource values for the polygon of interest (e.g. global horizontal irradiation – GHI – in the case of solar energy). Furthermore, energy conversion models can easily be implemented at the end of the chain in order to retrieve the system’s final output energy potential (e.g. PV output rated energy from GHI and other parameters).

3.4. Output

As a GIS-based model, GREECE generates 2 outputs: one geographical layer and its corresponding dataset of attributes. Only the dataset is passed on to optimization, while the geographical layer ensures optimal candidate sites to be eventually located. To each candidate PS_i in output set of potential sites PS (step (3) in Figure 1) corresponds static (area, shape, land use, etc.) and dynamic (time series of resource values R_{it}) digitalized attributes. Table 1 depicts an example of dataset generated by the GREECE model.

Table 1: Example of feature attributes per polygon generated by GREECE.

Polygon id	Area (ha)	Shape	Distance to the grid (km)	Land use share			Elevation (m)		Slope (%)		Aspect(°)		Resource time series
				Grass	Savanna	Farming	mean	std	mean	std	mean	std	
Polygon 1	45	0.71	2.3	0.2	0.5	0.3	11.1	1.6	2.9	1.3	190.1	91.1	R_{1t}
Polygon 2	30	0.42	3.6	0.8	0.2	0	33.8	8.2	3.7	1.9	125.4	93.3	R_{2t}
...	
Polygon n	35	0.55	5.8	0.0	0.0	1.0	7.2	2.1	1.6	0.8	193.3	75.9	R_{nt}

¹ Aspect of 180° means south-facing orientation.

4. Optimal planning and sizing of PV parks (OPSPV)

This section describes the optimization problem, its modelling and our approach to solve it. Let PS be the set of candidate PV sites obtained from the GREECE model, i.e. areas satisfying all the geographical and expert requirements and enriched with de-spatialized and discretized features. The core outputs of the OPSPV problem are 1) a subset of sites to be installed to satisfy the energy demand, 2) the dimensions of those sites and associated costs. The complexity of this problem lies in the ubiquitous data uncertainty at hand (forecast energy demand, costs, and solar radiation) and the flexible PV site dimensions (thus production capacity) to be determined, and whose corresponding installation costs follow a non linear function relative to the area. In this section we formalize this problem, and describe our model and approach.

4.1. Problem specification

This optimization problem is permeated with uncertainty, that resides mainly in the fact that strategic planning is based on forecast data relative to the projected growth of energy demand and various costs, as well as the projected radiation values

based on existing measurements. To model data uncertainty we consider a robust model, with the aim of proposing reliable solutions that can be implemented in real life.

Robust optimization seeks a certain measure of robustness against uncertainty that can be represented as a deterministic variability of the value of the data and the solution(s) [50]. It originates from the use of worst case analysis to treat uncertainty, and has found applications in fields such as finance, logistics, engineering, portfolio management, etc. [51]. From a modelling point of view we talk about robust optimization or reliability based optimization. The modelling approach encloses the uncertain data projected within robust bounds [52]. For instance if the network energy demand in January, 2016 is of 85.26 MWh at 9am, its projected value for 2030 would be 123.46 MWh, but remains clearly uncertain. A robust model would enclose the projected value within robust bounds such that a projected value a becomes:

$$a \longrightarrow \tilde{a} = [a - \hat{a}, a + \hat{a}]$$

where \hat{a} denotes the degree of uncertainty to adjust the robustness of the data with regard to the knowledge at hand and projection in terms of degree of conservatism. In terms of model we consider an interval linear model for which we seek robust optimal solutions. Basically a constraint of the form $aX \leq b$ where a and b are uncertain data parameters, will be modelled as $[\underline{a}, \bar{a}] \times X = [\underline{b}, \bar{b}]$. Extensive practical, theoretical and algorithmic work in this field allow such interval linear models to find their equivalent linear program in terms of best case scenario, worst case and convex region of possibilities, with the overall goal of minimizing costs:

$$\min \sum_i c_i X_i$$

The convex region is defined by the following equivalent set of linear constraints in terms of solution set:

$$S_{X_i} = \left\{ X_i \in R^+, \sum_i [\underline{a}_{ih}, \bar{a}_{ih}] \times X_i = [\underline{b}_h, \bar{b}_h], \forall h \right\} = \left\{ X_i \in R^+, \sum_i \underline{a}_{ih} \times X_i \leq \bar{b}_h \ \& \ \sum_i \bar{a}_{ih} \times X_i \geq \underline{b}_h, \forall h \right\}$$

The best case scenario corresponds to the highest possible resources (upper bound of a), and lowest required limit (lower bound of b), thus for one interval linear constraint we have:

$$\sum_i [\underline{a}_{ih}, \bar{a}_{ih}] \times X_i \leq [\underline{b}_h, \bar{b}_h] \implies \sum_i \bar{a}_{ih} \times X_i \leq \underline{b}_h$$

The worst case scenario corresponds to the lowest possible resources (lower bound of a), and highest limit to reach (upper bound of b), thus the constraint:

$$\sum_i [\underline{a}_{ih}, \bar{a}_{ih}] \times X_i \leq [\underline{b}_h, \bar{b}_h] \implies \sum_i \underline{a}_{ih} \times X_i \leq \bar{b}_h$$

In our strategic planning problem, the model will consider the different scenarios that relate to the risk adversity of the decision maker.

The OPSPV, energy strategy planning problem can be defined as a two-dimensional fractional knapsack problem with additional constraints (choice of the size of the PV candidate site with bounded size). Indeed a minimal and maximal size of a PV park has been set by the expert and geographical constraints in the GREECE model, and the installation cost is not fixed but relative to the park size according to a non-linear function. The ratio (€/kW) is high for small parks and gets smaller as the size grows. In this analogy, the knapsack corresponds to the maximum forecast energy to be provided by intermittent RE (in kW/hr), and the items are the sites with their potential supply (kW/hr) plus their associated costs (installation relative to the size (€/kW), and connection to the grid and substation (€/m)). The objective is to maximize hourly demand satisfaction while minimizing global costs. In addition the problem is permeated with data uncertainty which adds another level of complexity.

4.2. Problem modeling

Variables. Given that two aspects are associated with each candidate site, their costs and their production (size), we consider two sets of variables that need to be linked to each other. Boolean variables relate to the selection of a site to which various fixed costs are attached, and the area variables ranging in an interval of real values relate to the energy production and the variable installation cost of a new PV site according to its determined size.

Given:

Units

Hour (per year) $h \in H = \{1, \dots, 8760\}$ Site index $i \in N = \{1, \dots, n\}$

Energy demand and production

Current hourly production from intermittent energy sources (kWh) $Eint_h$ Current hourly production from non-intermittent energy sources (kWh) Ep_h Estimated global (forecasted) hourly power demand (kWh) $Dem_h = [\underline{Dem}_h, \overline{Dem}_h]$ Nominal power per unit area (kW/m²) $Pnom$

Candidates PV sites

Minimum and maximum *area* for each candidate parcel (m²) $Smin, Smax_i$ Potential Site $PS_{i \in N}$ Estimated *hourly production* per PV unit (kWh/m²) $Ppv_{h,i} = [\underline{Ppv}_{h,i}, \overline{Ppv}_{h,i}]$ Minimal distance from the grid to centroid of a candidate PV site (m) Dg_i

Costs

Transmission line unit cost (€/m) $Clan$ Substation unit power cost (€/kW) $Csta$ **Find:**The set P of sites to build a PV plant on,The surface SA_i to consider for each candidate site i that is selected**Cost functions:**Sum of all costs of PV installation (Minimize) $C = \sum_i (Ccap_i + Cop_i + Ccon_i + Csta_i)$ Capital cost of implementation of new PV power plant (€) $Ccap_i$ Annual operational cost per new PV plant (€) Cop_i Connection costs for each new PV plant, transmission lines (€) $Ccon_i$ Capital cost for new substation (€) $Csta_i$ Total added PV energy production (Maximize) $\sum_h PV_h$ Unsatisfied forecast demand (Minimize) $\sum_h |Dem_h - PV_h|$ **Such that the following constraints hold:**

PV existing and added production must be less than 35% of the total energy demand per hour

PV site size can't exceed maximal given size

PV newly added production plus existing production can't exceed the total hourly demand

Figure 3: OPSPV problem specification

 $\forall i \in N = \{1, \dots, n\}, \mathbf{B}_i \in \{0, 1\}$ 1 if parcel is selected, 0 otherwise $\forall i \in N = \{1, \dots, n\}, \mathbf{SA}_i \in [0.00 \dots Smax_i]$ Area of a solar PV plant

The first set of constraints states that the amount of intermittent energy resource into the network should be less than 35 % of the total forecast energy demand per hour (upper bound) [27, 53]. The time stamp is the hour, so there is one constraint per hour of the year.

$$\forall_h \sum_i \mathbf{SA}_i \times [\underline{Ppv}_{h,i}, \overline{Ppv}_{h,i}] + Eint_h \leq 0.35 \times [\underline{Dem}_h, \overline{Dem}_h]$$

Knapsack constraints. The second set of constraints relates to the satisfaction of the forecast energy demand, using existing resources augmented with new PV production. It seeks to determine the capacity of new PV plants to satisfy the anticipated demand. The time stamp is the hour. Two scenarios are considered: 2) best case scenario (highest forecast GHI and lowest forecast demand) and 3) worst case scenario (lowest forecast GHI and highest forecast demand). This allows to study the impact of the decision maker risk adversity in planning the creation of new PV plants.

$$\begin{aligned}
\text{Best case scenario:} \quad & \forall_h, \quad \sum_i \mathbf{SA}_i \times \overline{Ppv_{h,i}} + Eint_h + Ep_h \leq \underline{Dem}_h \\
\text{Worst case scenario:} \quad & \forall_h, \quad \sum_i \mathbf{SA}_i \times \underline{Ppv_{h,i}} + Eint_h + Ep_h \leq \overline{Dem}_h
\end{aligned}$$

The third set of constraints establishes a link between the size of the PV parks and their selection as successful candidates or not. This relationship is needed to connect the energy production and the various costs. If a plant size is not null then the site is selected, and conversely if a site is not selected its size is forced to be null.

$$\forall_i, \quad \mathbf{SA}_i \leq Smax_i \times \mathbf{B}_i, \quad \mathbf{B}_i \times Smin \leq \mathbf{SA}_i$$

Objectives and cost functions. Overall the strategic energy planning over the 2030 horizon has two main cost functions: 1) to maximize the total hourly energy demand production over the year through new PV energy production, 2) to minimize the total costs related to PV installation, connection to the grid, etc. Since these functions are in different units they will not be combined into a single weighted function that will not be meaningful, but instead solved by seeking the Pareto frontier, i.e. optimizing each function while constraining the other one, and reciprocally.

Maximize production. The first one is to maximize the production of PV (where $Ppv_{h,i}$ can take the best or worst case value depending on the scenario at hand):

$$\sum_i \sum_h \mathbf{SA}_i \times Ppv_{h,i}$$

Minimize costs: Modelling non-linear functions. This problem takes into account four cost functions that relate to each PV plant as defined in Fig. 3. Capital and operational costs are generally approximated as linear functions such as in [54] or in [55]. However, this approach is unrealistic as both the Cap_i and Cop_i costs are non-linear and depend on the size of the number thus the number of PV panels [56]. Basically the less the number of panels the highest the relative cost per panel. Accounting for the non-linearity of the costs is important when we allow for a variability in the number of panels on a site. We have thus considered an innovative approach using a piece-wise linear function such that a_i is the coefficient of the slope, and y_i the value of the coordinate where the new slope begins. It is illustrated below for Cap_i , and Cop_i follows a similar specification with different constants. Values have been set from [56]:

$$Cap_i = \begin{cases} a_1 \times Pnom \times \mathbf{SA}_i + y_1 & \text{if } 0_{MW} \leq \mathbf{SA}_i \times Pnom \leq 1_{MW} \\ a_2 \times Pnom \times \mathbf{SA}_i + y_2 & \text{if } 1_{MW} \leq \mathbf{SA}_i \times Pnom \leq 10_{MW} \\ a_3 \times Pnom \times \mathbf{SA}_i + y_3 & \text{if } 10_{MW} \leq \mathbf{SA}_i \times Pnom \end{cases}$$

On the other hand, the unit connection cost of a PV park, and the capital cost of a new station for a park, are both linear functions that depend respectively on the selection of a park with its distance to the grid, and the unit cost of a substation proportional to the size of the park. We have the following functions:

$$Ccon_i = Clan \times Dg_i \times \mathbf{B}_i, \quad Csta_i = Csta \times Pnom \times \mathbf{SA}_i$$

5. Case study: Optimizing PV parks implantation and production in French Guiana

In this work, we have applied our model framework to strategic planning of new solar PV power facilities in French Guiana, according to characteristics and policy targets of the region for the horizon 2030. This overseas department is located in the northern coast of South America, and shares border with Brazil (south-east), Suriname (west) and Atlantic ocean (north-east). It is depicted in Figure 4.

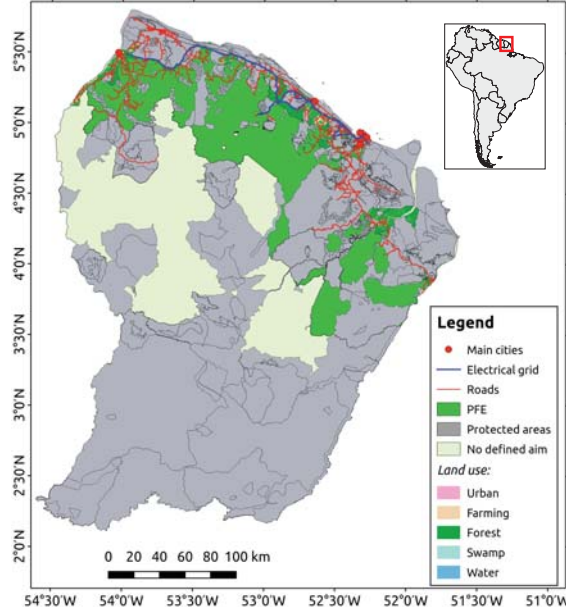


Figure 4: Map of French Guiana with corresponding geographical features (constraint layers, networks, etc.).

5.1. Data acquisition and processing

As described in Fig. 3, this problem uses solar and land management resources, various network resources (electrical grid, roads), as well as economical data. While the resource data are specific to the territory of French Guiana, the economical data that contribute to defining the cost function are drawn from published generic documents. In this section we describe in more detail how we extracted the solar and land management resources, and recall the data used for the various costs and the definition of the cost function.

5.1.1. Geographical constraints and land management

French Guiana is typically a region with great land constraints. More than 96 % of the 83,534 km² of the territory is in fact occupied by forest [57]. As pressure of human activities on this ecosystem has been increasing over the last 30 years, various protective measures have been taken accordingly [58]. Figure 4 shows the union of all protected areas (ecological and landscape protection, registered sites, Amazonian Park, etc.), retrieved from the world database on protected areas (WPA) [33], ONF [59] and IGN [35], alongside the electrical network connecting main cities between them. Power transmission lines are limited to the north, in a narrow band along the shore, and essentially surrounded by forest [34]. South of this band lies the *Permanent Forest Estate* (PFE) of Guiana which is made for forestry purposes only [34]. As no distance higher than 20 km is generally taken into account for grid connection [46], it is clear at first glance how carefully potential sites should be targeted.

Table 2 presents the different assumptions we have made about restricted areas according to 3 different scenarios. Restriction layers (protected areas, land use, networks, etc.) have been fetched from different world [33] and national geographic databases (e.g. ONF for land use [34] and forests [59], and IGN [35]). Buffer values have been retrieved from the literature and adapted to our problem; corresponding references are included in the table. When no value was available from the literature, we have taken the one closer to the corresponding land use properties (e.g. "orchard/rice farming" buffer has been defined from the value corresponding to "agriculture", and "dune/sand" from "wetland"). We have considered 3 land management scenarios, from *restrictive* to *permissive*, generating 3 different sets of exploitable plots. The scenarios differ in the buffer values considered as well as in the accounting or not for the restriction layers (a straight line in Table 2 means that the corresponding restriction layer is not considered). For example grassland is never restricted, and agricultural lands may not be restricted (permissive scenario), or be restricted and take different buffer values accordingly (*restrictive* and *intermediate* scenarios).

Table 2: Restricted areas used in this study according to land management scenarios.

Layer	Buffer (m)			Reference(s)
	Restrictive	Intermediate	Permissive	
Protected areas	1000	750	500	[60]
PFE/Forest/ Mangrove	1000	500	200	[23]
Urban areas	1000	500	200	[24, 60]
Flood savanna	500	200	100	[61]
Water bodies	500	200	100	[61, 62]
Shoreline	500	200	100	[62]
Electrical grid	500	200	100	[14, 24, 61]
Road network	500	200	30	[14, 24, 61]
Wetland	200	100	0	[61]
Dune/sand	200	100	0	
Orchard/rice farming	100	0	0	
Agriculture	100	0	—	[23]
Arable land	0	—	—	
Rocky/dry savanna	0	—	—	[23]
Grassland	—	—	—	

Table 3: Distance threshold.

Layer	Threshold (m)
Electrical grid	20,000
Road network	20,000

In the case of the maximum distance to road and electrical networks, a common value is 10 km [23]. We have preferred to double it in order to give priority to the optimization module for assessing the cost of the furthest zones (Table 3).

Table 4: Min/max surface area of resulting land parcels.

Criterion	Plot area (ha)	Reference
Minimum	1.5	[63]
Maximum	50	[53]

Table 4 shows the values set for minimum and maximum plot area, adapted from the literature. Regarding upper value, investment into power generation shall be achieved by tranche of 20 MW in the region [53]. From Table 5, it would correspond to a solar PV plant of about 36 ha. Therefore, we have set the maximum land surface to 50 ha, leaving some extra space for flexibility purposes and potential future improvements. The land has eventually been partitioned with respect to this threshold and the method depicted in section 3.3.

Scenario outputs corroborate what we previously noted about the significant constraints the territory is subject to. While the *permissive* strategy resulted in 133 candidates covering about 30 km², *intermediate* and *restrictive* gave either few (9) or no candidates at all. In the general case, each resulting set would feed the optimization and produce as many distinct optimal solutions left at the discretion of the decision and policy makers. Here, the specific constraints of French Guiana have resulted

in the *permissive* scenario as the only relevant land configuration, that is the only scenario for which a significant number of potential sites is generated. Hence we will focus on this scenario in the following analysis, whose geographical distribution of the corresponding land parcels is depicted in Figure 5.

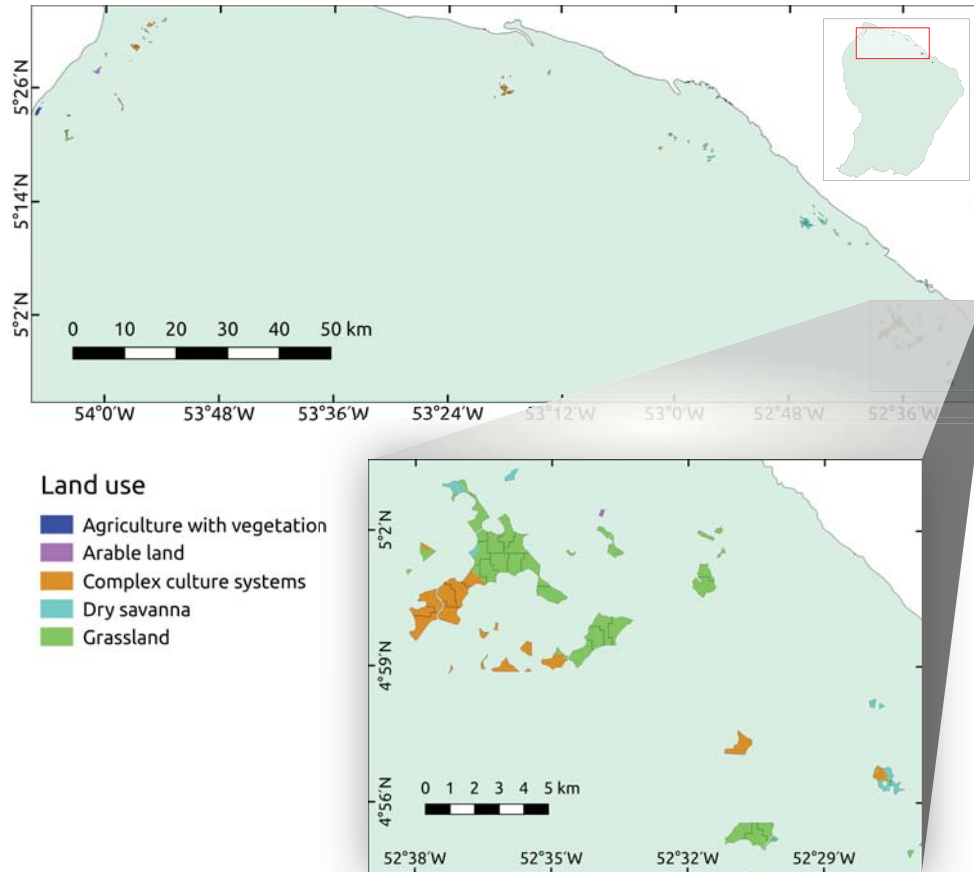


Figure 5: Set of land parcels generated by the GREECE model for French Guiana when considering the *permissive* scenario.

5.1.2. Solar resource and PV power generation

Solar PV has been given a significant share into French Guiana’s energy planning policy for the next 10 years: it stands for about 75 % of the intermittent power to be installed by 2023 [27]. Our analysis has hence focused on optimizing the distribution of those facilities throughout the territory. To be as close as possible to reality, most of the energy system models currently provide analysis at *hourly* time scale [9]. In this study, we have therefore followed the same pattern for simulation and optimization of the given PV facilities. Energy demand hourly data have been retrieved from France’s electric utility EDF open data server [36].

We describe below how we generated hourly solar potential values within each of the polygons computed by the GREECE model. Note that when latitude and longitude are required in GHI and PV sub-models described thereafter, we always refer to polygon’s centroid coordinates.

5.1.3. Solar potential

Climate of French Guiana is equatorial and divided into two main seasons: dry season, from August to December, and rainy season, from December to June [57]. It affects yearly variability of solar radiation accordingly. To assess the resulting

solar potential within each land parcel, we have first relied on the monthly satellite-based GHI maps of the region. Those maps were generated at 0.03° resolution by applying an optimized version of the Heliosat-2 method to GOES-13 satellite images [37]. As depicted previously (see section 3.3), we have finally aggregated monthly irradiation within each polygon by averaging all corresponding raster cell values. Finally, in order to obtain resource potential at hourly scale, we have scaled down the previous GHI monthly time series by using an updated version [64] of the synthetic generation model developed by Aguiar *et al.* [49, 65]. We have beforehand delimited cluster zones for each month with respect to GHI amount, that is gathering of raster cells presenting close solar irradiation values. Polygons belonging to the same cluster have been given the same pattern in daily clear-sky index sequence over the year. It actually prevents neighboring polygons from having too different solar radiation patterns over time. Hourly GHI time series have then been derived from those daily sequences. Figure 6 shows an example of monthly *cluster zones* used in generating hourly GHI for each candidate site based on February solar GHI map.

It is important to note that we have generated daily and hourly patterns only because of the original monthly temporal resolution of solar GHI map products over the region. Ideally, one would instead directly use satellite-based radiation data at daily or hourly scale. Our model can take full advantage of spatiotemporal remote sensing data at hourly scale.

5.1.4. PV system's output energy

As shown in Figure 1, it is possible to integrate a conversion model between GREECE and the optimization module, in order to obtain energy data from the previously aggregated GHI resource. Regarding this study, we have thus integrated solar PV conversion into our model framework, by adapting the `pvl` Python library from the Sandia National Laboratory (SNL) [66]. This model makes use of the hourly GHI data described beforehand as well as both endogenous and exogenous factors.

Endogenous features correspond to system's characteristics and are depicted in Table 5. As power plant size is regarded as a variable in optimization, we have considered as a facility any gathering of several basic PV *units*. One PV unit is composed of various strings of modules connected to an inverter, and occupies a specific amount of space (land usage). Size, nominal capacity and electricity generation of a solar PV plant are eventually summed up over all units it is made of. Common values, such as for PV-to-inverter ratio [67], land usage [68] or PV array tilt and azimuth [69], have been used.

Table 5: Characteristics of one PV unit used in energy simulation.

Nominal power (P^*)	6 kW
Number of modules	20
Inverter power	5 kW
PV-to-inverter ratio	1.2
Array tilt	Site latitude
Array azimuth	180°
Total module area	32.0 m^2
Land usage (A_{mod})	106.7 m^2
$P_{nom} = P^*/A_{mod}$	0.056 kW/m^2

Exogenous factors typically stand for weather and climatic parameters, that is solar radiation, ambient temperature T_{air} , wind speed V or albedo A . In first approximation, the last three have been taken as constant and equal to typical values ($T_{air} = 25^\circ\text{C}$, $V = 1 \text{ m/s}$, $A = 0.2$). In order to assess irradiance in module's plane, it was also necessary to split GHI into its beam (DNI) and diffuse (DHI) components. To do so, we have first converted irradiation into irradiance (average solar power over the hour). Then, we have retrieved DNI and DHI by using a diffuse fraction model. In the case of no available data, the one developed by Erbs *et al.* [70] is generally recommended [71]. Using the conversion model, we finally get 8760 values of the output average AC power from one PV unit (in Wh/h) for every polygon.

5.1.5. Costs for PV technology

The optimization model will seek a compromise solution between the maximization of the PV production and the minimization of the various costs. These have been defined in the general model formally. In this section we present the actual

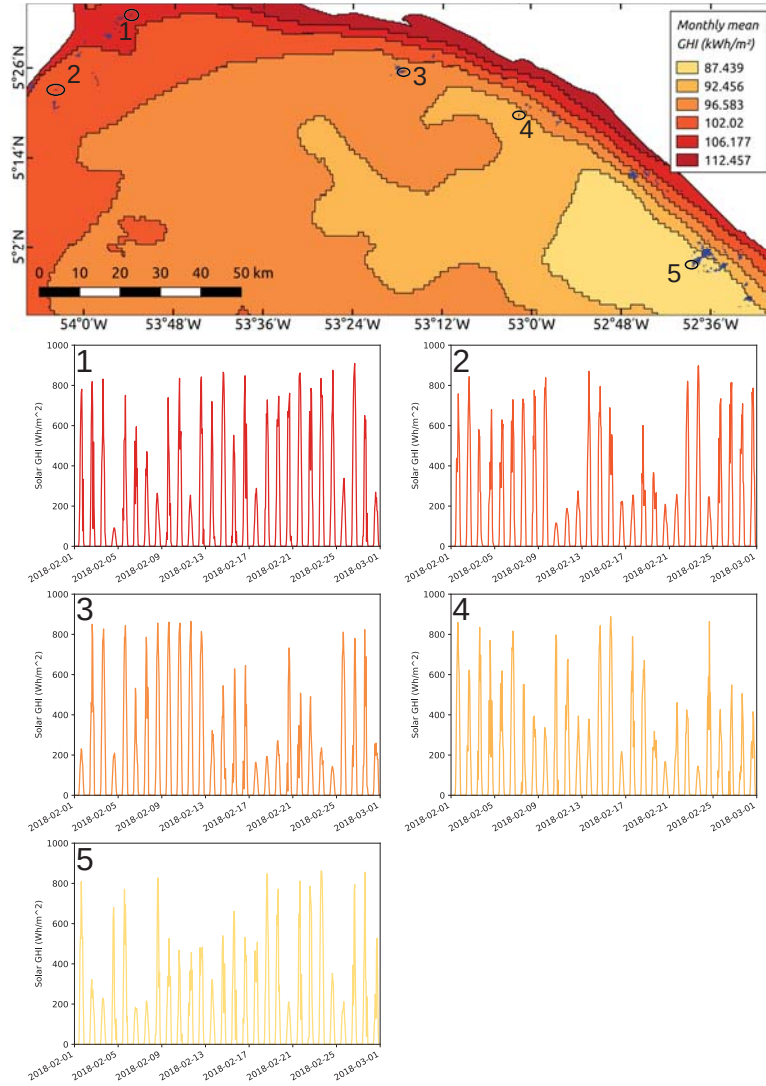


Figure 6: Example of monthly solar potential cluster zones (February) derived from solar GHI raster maps elaborated by Albarelo et al. [37] and Fillol et al. [38], and corresponding generated temporal patterns for candidate PV sites based on the *permissive scenario* (see section 5.1.1).

values used in our case study. Some are unitary fixed values and some depend on the power associated to a given park and its creation.

Grid connection and substation costs. The fixed unit costs are given in Table 6 for the connection to the grid, through power lines and creation of substations.

Table 6: PV plant fixed unit costs.

Cost type	Value	Reference
Grid-connection cost $Clan$	1000 €/m	EDF
New substation cost $Csta$	100 €/kW	[72]

Capital and operational costs. Regarding the four cost functions, two are linear functions that relate to the above values, and two the capital cost and operational cost for a new PV park are non-linear functions for which the coefficients of the piece-wise linear model we considered are defined in Table 7 [56].

Table 7: PV plant capital and annual operation cost piece-wise parameters.

Capital cost of a new PV power plant (C_{cap_i})		Annual operations cost (C_{op_i})	
Slope a_n	y_n value	Slope b_n	y_n value
$a_1 = 2701$	$y_1 = 0$	$b_1 = 19$	$y_1 = 0$
$a_2 = 1800$	$y_2 = 901,000$	$b_2 = 16$	$y_2 = 3000$
$a_3 = 1200$	$y_3 = 6,901,000$	$b_3 = 10$	$y_3 = 63,000$

5.2. Scenarios

One of the core elements in our strategic planning approach is to evaluate the impact of risk adversity of the decision maker. To do so, we consider the approach made of best and worst case demand and production scenarios, so that we get a range of robust solutions. As detailed in section 4.2, the best case corresponds to the lowest forecast demand in 2030 with the highest forecast solar GHI; the worst case consider the highest forecast demand with the lowest forecast solar resource.

5.2.1. Global energy production and demand: current and forecast

In our case study we have considered the existing production and consumption data from 2016 as the baseline to draw forecast data for the horizon 2030 [36]. The distribution for 2016 was: hydropower (48%), thermal (45%), solar PV (6%), and biomass (1%) [53]. In order to establish forecast data for the potential PV production on each candidate site, we have considered 3 different data sets: 1) a static projection from current GHI data, 2) a 10 % stochastic increase (mitigation) where each site temporal pattern is enhanced (decreased) by a uniform random noise $U(0, 0.1)$ at each hour, 3) a 20 % stochastic increase (mitigation) where patterns are enhanced (decreased) by a uniform random noise $U(0, 0.2)$, in the best case scenario (worst case scenario).

Regarding our strategic planning problem of PV penetration in the Guiana’s electricity network, we made the assumption that the current production is stable and maintained, and we considered best and worst case scenarios in the increasing power demand based on EDF forecasting study in Guiana [53]. The best scenario assumes a 2% annual growth and the worst scenario a 5% growth from 2016 [73]. The 2016 hourly demand data was based on the hourly production data, since demand data are not available, but conversely the production was sufficient in 2016. Fig. 7 illustrates the daily hour variations over 2030 and projection of the demand for the best and worst case scenarios.

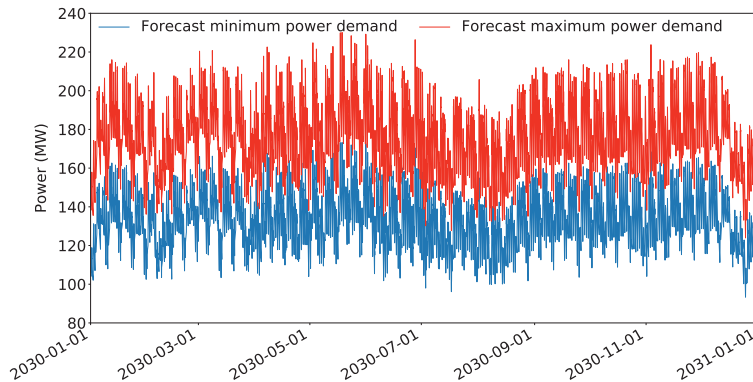


Figure 7: 2030 forecast power demand over daily hours.

6. Results and analysis

We here present the results of our integrated model for the permissive land management scenario (see Table 2). GREECE was implemented in Python using libraries recalled in the paper, and led to the extraction of 133 candidate parcels with their relevant features. Those de-spatialized data items were then implemented into the OPSPV module. Different numerical experiments were run for both best and worst scenarios, so that the total PV energy production would be maximized over the year 2030 and the total costs minimized. As it is a two-objective optimization, we have derived Pareto charts of compromising solutions for each scenario, by computing maximum amounts of generated energy for various cost thresholds. Typically, the computation time of one optimization run was between 23 and 47 s.

We have focused the following analysis on 4 specific points: 1) the contribution of the knapsack approach to the accuracy of the optimal solutions (section 6.1), 2) the impact of the 35 % constraint on the number of solar PV plants, their installed capacity and the final amount of produced energy (section 6.2), 3) the relation between solar GHI temporal patterns, location and size of the candidate sites, and the corresponding solutions in the Pareto set (section 6.3), 4) several aspects of relevance to the decision maker and network manager (section 6.4).

6.1. Fractional knapsack optimization

The fractional knapsack approach allows the size of a power plant within a given site to be optimized. In order to fully apprehend the gain in accuracy of this approach, we have derived two Pareto charts for one of the scenarios (worst case), either with or without considering it. In the second case, the size of a solar PV plant is assumed to be equal to the size of the corresponding parcel (*full land use*). Both resulting Pareto fronts are depicted in Fig. 8. We can observe that the fractional result acts as the Pareto asymptotic limit for both cases, while the full land use gives a staircase-like Pareto front hitting the previous limit when the size of the PV plants is close to the size of the selected sites.

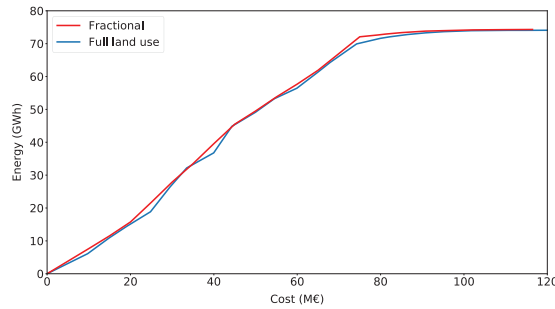


Figure 8: Worst case scenario Pareto solutions based on fractional (red) and full land use method (blue).

In the case of French Guiana, the difference between solar GHI patterns is not too sharpened, the permissive land management scenario results in a broad range of parcel's sizes and the sites are not so far from one another. Hence the gain in accuracy might be more limited than in the case of larger regions (e.g. countries) at lower or higher latitudes and with very heterogeneous climatic conditions. Even though, in this very specific case, the difference between both Pareto limits ranges up to about 3 GWh or 7.5 % of the energy generated in the full land use approach (around 40 M€).

6.2. Impact of the 35 % constraint

The maximum penetration limit of 35 % is related to the power generation from intermittent resources, and it is still considered a necessary constraint to be respected for the islanded or isolated power network in order to maintain its stability [74]. It is important for the investors to keep in contact with the network manager because if the power produced by new PV plants is greater than this limit, the last connected PV plant to the grid will be switched off. In French Guiana, this limit is set at 30 % of the power demand at any time [53] though, in the coming years, it is expected to increase to 35 % [27] when more dispatchable power facilities shall be installed (e.g. biomass, solar PV with storage, etc.).

The impact of this penetration limit (here, 35 % of the total hourly power demand) is depicted in Fig. 9 in the worst case scenario with no cost threshold (C_∞). It is especially relevant for that kind of scenario as the forecast power demand is at its maximum (5 % growth): the share of existing generation is lower, resulting in more new PV power being injected into the

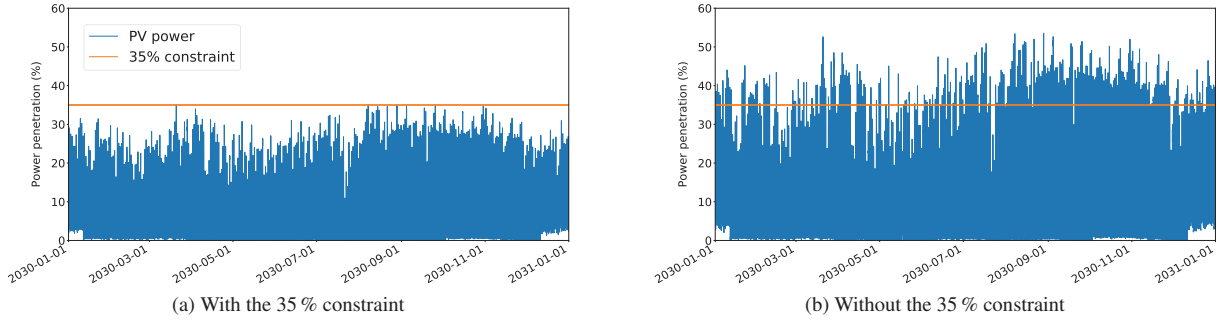


Figure 9: Effect of the 35 % constraint on final PV power generation

network and so more chances to get above the 35 % upper bound. It is actually observed in Fig. 9b, where 23 % of the daily hours (1043 hours) exceed the limit when the constraint is not applied. Table 8 shows that the corresponding installed capacity and generated energy almost double when the constraint is removed.

Table 8: Effect of the 35 % constraint on final PV generation

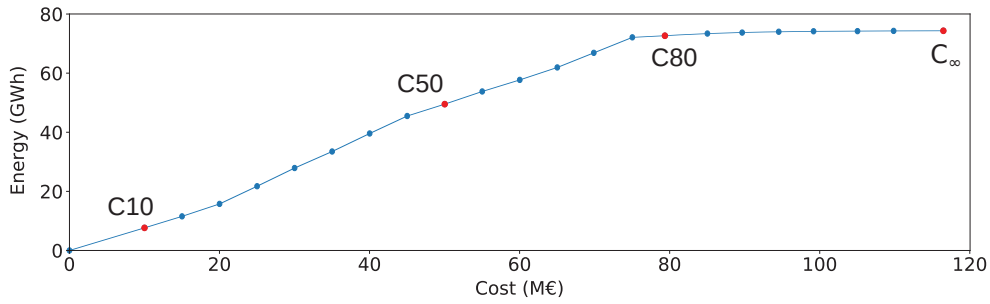
Scenario	Total costs (M€)	Total energy (GWh/year)	Nb of parks	Total installed power (MW)
Figure 9a	116.44	74.32	11	45.65
Figure 9b	195.92	135.28	16	83.10

It is clear from this analysis that this threshold might be a significant drag on the growth of RE share in a given network. Various actions could be undertaken in order to heighten this limit, such as using modern energy storage technologies to store the extra energy, developing smart grids, improving the forecasting of intermittent RE production, performing power demand management or supplying power from dispatchable renewable resources. Those actions would keep the balance between power generation and demand in the region at any time, and reduce the risks of electrical failures [7].

6.3. spatiotemporal solar GHI patterns and corresponding Pareto solutions

Each compromising solution provided by the Pareto chart for a given scenario is related to both location and size of the selected candidate sites. When combining the available candidates from GREECE, the OPSPV module seeks to match the power demand as well as to balance the power generation over time due to the 35 % upper limit. We illustrate this interdependence between location of the solar PV plants and corresponding temporal solar GHI patterns by looking at how PV sites are selected when both the amount of generated energy and the total cost grow along the Pareto. As a result, Fig. 10 depicts the geographical location of the selected PV sites alongside GHI cluster zones (see Fig. 6 for the corresponding temporal patterns) according to different cost thresholds on the Pareto chart. Table 9 also gives the corresponding: number of parks, total installed power and installed capacity per solar PV plant.

As expected, as we move to the right of the Pareto set in Fig. 10 (increasing produced energy and cost), the more the number of PV parks and the higher the global installed power. Selected PV sites also remain close to the power grid as the connection costs rise with the corresponding distance. The OPSPV looks for filling up the *free energy volume*, that is the available volume between existing power production and either the power demand or the 35 % upper bound, which is initially empty. Larger parks can therefore be built, as power generation over time does not reach any limit and cheapest solutions will necessarily be selected, i.e. solutions with fewer plants. Then, the more the energy generated, and so the fewer the energy volume left, the more the OPSPV looks for sites both smaller and located in different solar GHI cluster zones. The latter ensures balancing the intermittent global output power in order to remain below any upper threshold (power demand, 35 % limit) over time. The combination of different solar GHI patterns results in *geographical dispersion* of the solar PV plants when increasing their share into the energy mix. It is especially significant when no cost constraint is applied (C_∞)



Number of parks, total installed capacity, number of GHI clusters

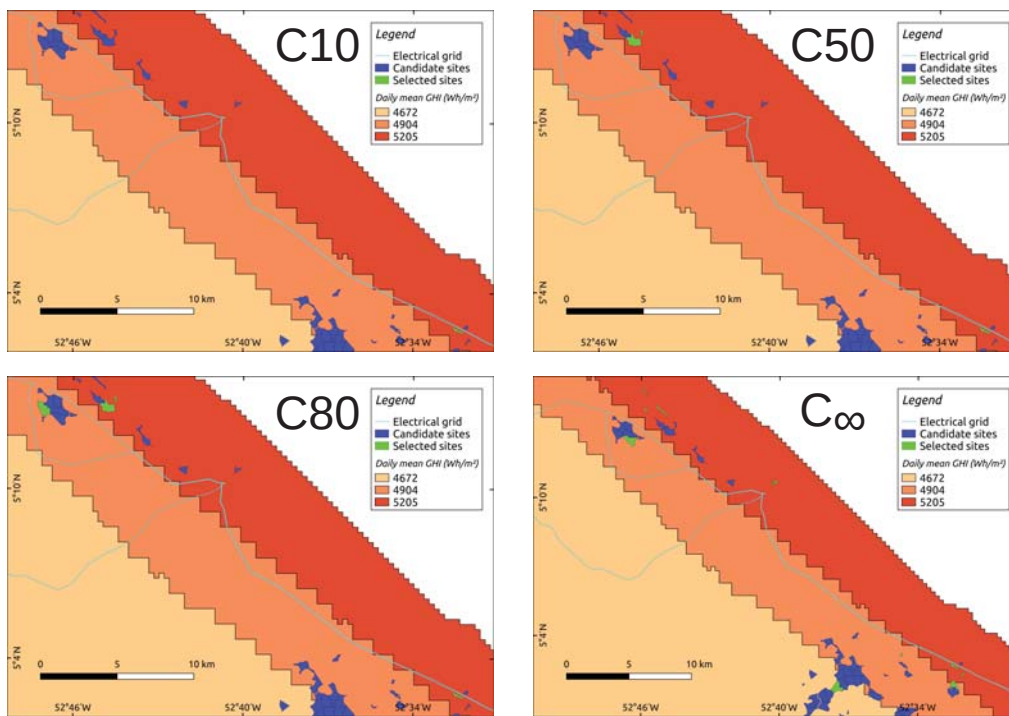


Figure 10: Evolution of PV site geographical location with Pareto cost.

and the optimization only seeks to maximize the global output energy. In order to fill up the energy volume as much as possible, the OPSPV heads for combining the temporal patterns between them. Accordingly, though the generated energy barely grows between C80 and C ∞ , the number of selected sites rises dramatically, as well as their geographical dispersion (i.e. the number of GHI clusters, see Fig. 10), while the size of the corresponding power plants naturally decreases (see Table 9). As the number of PV stations increases but total installed capacity remains unchanged, total costs start growing faster than the produced energy and the Pareto line thus reaches its plateau (around C75).

Those results emphasize the capacity of our integrated approach to perform *geographical aggregation* with contextual real data, i.e. balancing the intermittency of RE sources by spreading out the corresponding installations throughout a given territory as shown by Haurant [75] or Jewell et al. [76].

Table 9: Characteristics of the PV parks depicted in Fig. 10.

Cost level	Generated energy (GWh/year)	Nb of parks	Total installed power (MW)	Set of installed capacities (MW)
C10	7.65	1	4.66	{4.66}
C50	49.80	2	30.50	{1.16, 29.34}
C80	72.64	3	44.98	{5.36, 10.28, 29.34}
C ∞	74.31	11	45.65	{1.31, 1.34, 2.19, 2.22, 2.22, 2.47, 2.89, 4.38, 5.36, 6.69, 14.58}

6.4. Towards helping decision making and network managing

In this section, we analyze specific aspects of relevance to both the decision maker and network manager (power plant energy investors in French Guiana and EDF) that are made possible with the combination of geographical and temporal contextual information and robust optimization: 1) the study of the risk adversity comparing best and worst case scenarios (section 6.4.1), 2) the identification of robust planning investment scenarios where the optimal plants show to be identical regardless of the degree of spatiotemporal uncertainty on the power resource (section 6.4.2).

6.4.1. Risk adversity: best vs worst case scenario

By comparing best and worst case scenarios, one may assess the robustness of the different energy planning investment scenarios available to the decision maker, in spite of uncertainty on both forecast power demand and resource. To evaluate the range of robust solutions, and their associated risk, we consider the corresponding Pareto charts shown in Fig. 11. Those charts have been derived from the principle exposed in section 4.2 and by applying the 20 % stochastic increase (best case) and mitigation (worst case) on the resource described in section 5.2.1.

Essentially, Fig. 11 gives information about the volume of solar PV plants one may install within the region without threatening the power grid in the long-term, and thus allows for risk adversity to be estimated. The robust solution corresponds to the convex region, where power generation from any selected solar PV plant can fill up the *free energy volume* regardless of the scenario, but at different speeds (i.e. the slope of each Pareto curve). The best case stands for the smallest volume filled at the highest pace, so that as long as the worst case Pareto is below the best case, the investment remains safe. The corresponding PV sites can be explored in the limits of their maximum capacity estimated by the robust optimization. In contrast, once the best case reaches its plateau, and so Pareto lines cross each other (around 45 M€), the leftover energy volume in that scenario gets too low: though any park can still be selected in the worst case (larger remaining volume), only smaller ones can make it in the best case. At this point, the more the energy generated in the worst case scenario, the more it exceeds the energy generation limit in the best case: the risk grows as much as the gap between both lines.

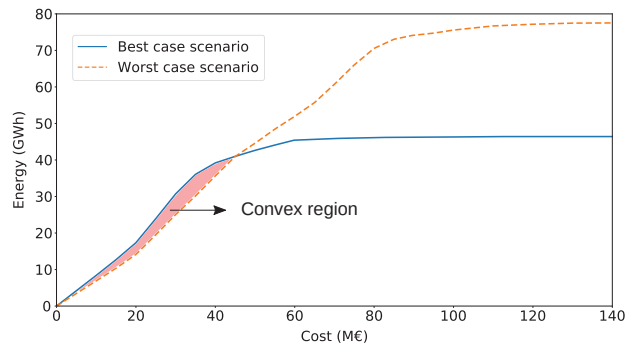


Figure 11: Best vs. worst case scenario.

6.4.2. Candidate site's location robustness with uncertainty

For one given scenario, it is interesting to assess how uncertainty affects the location of the selected PV sites. A more robust solution would be one where either candidate sites or corresponding temporal patterns remain the same for different

confidence levels. To illustrate this statement, we have applied different degrees of spatiotemporal uncertainty to the resource in the worst case scenario. Pareto charts of Fig. 12a respectively depict: i) estimated GHI, energy potential (in blue), ii) its mitigation by a random uniform noise per hour and per parcel, between 0 and 10 % (in green) and iii) between 0 and 20 % (in red). In that case, the safety zone lies below a cost of 70 M€ (C70), meaning that the selected PV sites are the same for all resource time series projections. Those sites are sorted by ascending parcel size area from the smallest (P1) to the largest (P4) and are visible in Fig. 12b. The investment costs grow naturally when the optimal PV plant size grows within its corresponding parcel area (fractional knapsack optimization).

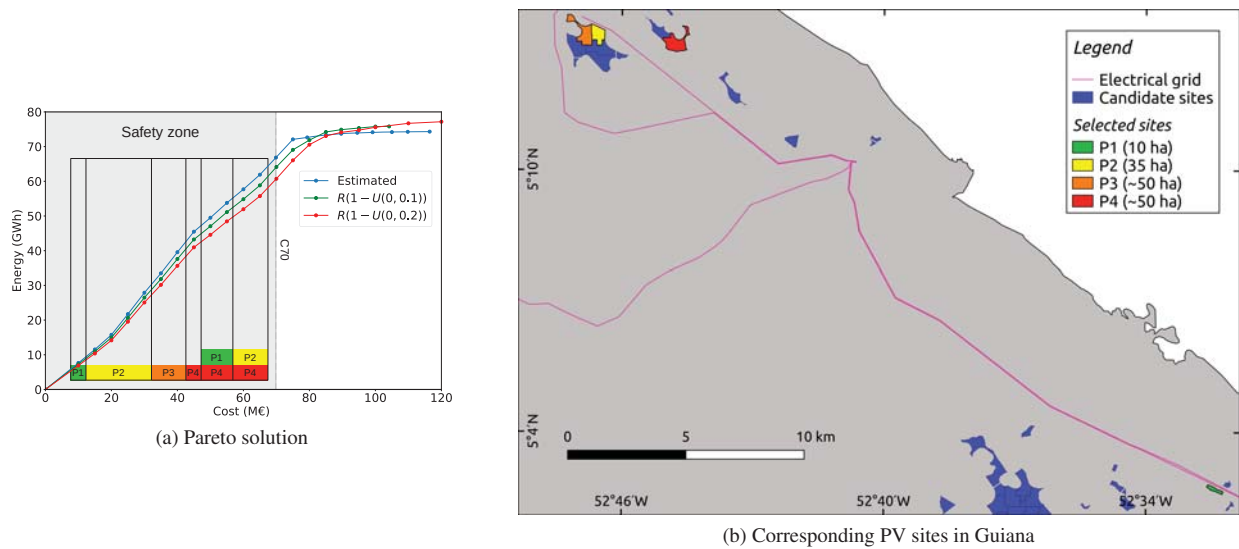


Figure 12: Robustness of the worst-case scenario solution when resource is permeated with different uncertainty levels.

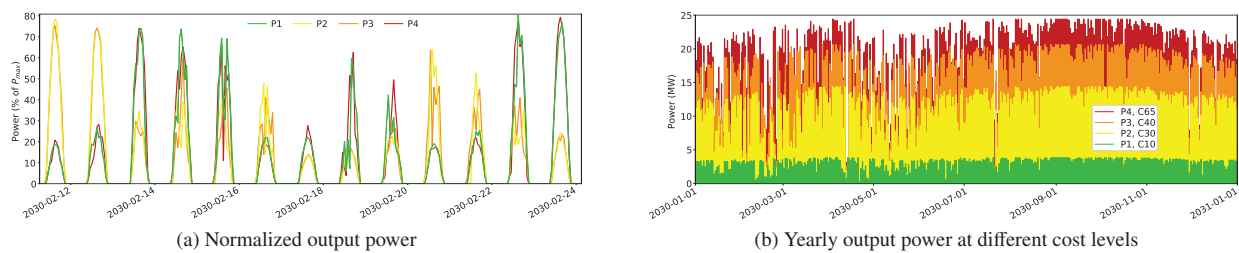


Figure 13: Output power from solar PV plants of Fig. 12

Fig. 13 depicts the corresponding output power from given solar PV plants selected by the robust optimization, whose placement is visible in Fig. 12b, in the case of estimated solar GHI. Fig. 13a depicts the resulting spatial variation of the solar GHI patterns (derived by GREECE) on the generated power from the optimal solar PV plants (location and size). P1 and P4 sites (P2 and P3) have similar patterns for they are located in the same solar potential cluster zone (see Fig. 6). Essentially, it shows how our integrated approach manages real site spatial arrangement so that the global output power from the optimal solar PV plants is balanced through time: impact of their RE intermittency on future network power management is limited. Finally, Fig. 13b shows how the robust optimization makes power generation over time gradually filling up the free energy volume by selecting different sites and power plants' capacity.

7. Conclusion

In this article we have dealt with both geographical and temporal constraints for planning scenarios of intermittent renewable power at broad scale. We have developed a model at the interface between both GIS and holistic approaches. In addition to GIS-based studies combined with multi-criteria analysis, we take into consideration the temporal variation of both the resource and the power demand for identifying suitable sites. Instead of suitability maps, we provide the exact sites, i.e. location and size, where power plants should be built so that the share of intermittent renewable energy sources is maximized within a given region. We also go beyond usual bottom-up energy models by adding geographical constraints, and true possibilities granted by the territory, to the optimization problem. To that extent, our integrated approach is primarily a decision making tool for local policy makers and network managers. It is primarily intended to provide real contextual solutions within present and future renewable energy policies at the regional and national scale. We have contributed and shown how our integrated model allows for retrieving real spatial distribution of power plants resulting in mitigation of global renewable energy intermittency at the hour. A key contribution of our work was to bring forward the importance of seeking trade-offs between land management and energy transition goals, especially in territories such as French Guiana where ecosystems must be protected. In that case, only the permissive scenario allows for solar PV plants to be established. Decision-makers should therefore be eventually advised by environmental experts so that global sustainability is achieved hand in hand with the local environment. Future work includes the generalization of this model to cost-effectiveness of energy storage, to biomass that raises a complex temporal issue of the renewability of the resource, and to power flow modeling in order to test the validity of each solution in terms of grid voltage, as it has already been done at lower scale [10].

References

- [1] E. Hache, A. Palle, *Renewable energy source integration into power networks, research trends and policy implications: A bibliometric and research actors survey analysis*, *Energy Policy* 124 (2019) 23–35. doi:10.1016/j.enpol.2018.09.036.
URL <https://doi.org/10.1016/j.enpol.2018.09.036>
- [2] L. Gacitua, P. Gallegos, R. Henriquez-Auba, Lorca, M. Negrete-Pincetic, D. Olivares, A. Valenzuela, G. Wenzel, *A comprehensive review on expansion planning: Models and tools for energy policy analysis*, *Renewable and Sustainable Energy Reviews* 98 (2018) 346–360. doi:10.1016/j.rser.2018.08.043.
URL <https://doi.org/10.1016/j.rser.2018.08.043>
- [3] R. Lacal Arantegui, A. Jäger-Waldau, *Photovoltaics and wind status in the European Union after the Paris Agreement*, *Renewable and Sustainable Energy Reviews* 81 (2018) 2460–2471. doi:10.1016/j.rser.2017.06.052.
URL <https://doi.org/10.1016/j.rser.2017.06.052>
- [4] P. Das, J. Mathur, R. Bhakar, A. Kanudia, *Implications of short-term renewable energy resource intermittency in long-term power system planning*, *Energy Strategy Reviews* 22 (2018) 1–15. doi:10.1016/j.esr.2018.06.005.
URL <https://doi.org/10.1016/j.esr.2018.06.005>
- [5] R. Li, Y. Jin, *A wind speed interval prediction system based on multi-objective optimization for machine learning method*, *Applied Energy* 228 (2018) 2207 – 2220. doi:https://doi.org/10.1016/j.apenergy.2018.07.032.
URL <http://www.sciencedirect.com/science/article/pii/S0306261918310584>
- [6] P. Jiang, R. Li, H. Li, *Multi-objective algorithm for the design of prediction intervals for wind power forecasting model*, *Applied Mathematical Modelling* 67 (2019) 101 – 122. doi:https://doi.org/10.1016/j.apm.2018.10.019.
URL <http://www.sciencedirect.com/science/article/pii/S0307904X18305110>
- [7] X. Li, M. Paster, J. Stubbins, *The dynamics of electricity grid operation with increasing renewables and the path toward maximum renewable deployment*, *Renewable and Sustainable Energy Reviews* 47 (2015) 1007–1015. doi:10.1016/j.rser.2015.03.039.
URL <http://www.sciencedirect.com/science/article/pii/S1364032115001926>

- [8] L. Barbosa, D. Bogdanov, P. Vainikka, C. Breyer, Hydro, wind and solar power as a base for a 100% Renewable Energy supply for South and Central America, *PLoS ONE* 12 (3) (2017) 1–28. doi:10.1371/journal.pone.0173820. URL <http://dx.plos.org/10.1371/journal.pone.0173820>
- [9] L. Ramirez Camargo, G. Stoeglehner, Spatiotemporal modelling for integrated spatial and energy planning, *Energy, Sustainability and Society* 8 (1) (2018) 32. doi:10.1186/s13705-018-0174-z. URL <https://energysustainsoc.springeropen.com/articles/10.1186/s13705-018-0174-z>
- [10] S. Kucuksari, A. M. Khaleghi, M. Hamidi, Y. Zhang, F. Szidarovszky, G. Bayraksan, Y. J. Son, An Integrated GIS, optimization and simulation framework for optimal PV size and location in campus area environments, *Applied Energy* 113 (2014) 1601–1613. doi:10.1016/j.apenergy.2013.09.002. URL <http://dx.doi.org/10.1016/j.apenergy.2013.09.002>
- [11] G. Mavromatidis, K. Orehounig, J. Carmeliet, Evaluation of photovoltaic integration potential in a village, *Solar Energy* 121 (2015) 152–168. doi:10.1016/j.solener.2015.03.044.
- [12] A. Alhamwi, W. Medjroubi, T. Vogt, C. Agert, Development of a GIS-based platform for the allocation and optimisation of distributed storage in urban energy systems, *Applied Energy* 251 (May) (2019) 113360. doi:10.1016/j.apenergy.2019.113360. URL <https://doi.org/10.1016/j.apenergy.2019.113360>
- [13] P. Huang, Z. Ma, L. Xiao, Y. Sun, Geographic Information System-assisted optimal design of renewable powered electric vehicle charging stations in high-density cities, *Applied Energy* 255 (2) (2019) 113855. doi:10.1016/j.apenergy.2019.113855. URL <https://doi.org/10.1016/j.apenergy.2019.113855>
- [14] M. L. Sabo, N. Mariun, H. Hizam, M. A. Mohd Radzi, A. Zakaria, Spatial matching of large-scale grid-connected photovoltaic power generation with utility demand in Peninsular Malaysia, *Applied Energy* 191 (2017) 663–688. doi:10.1016/j.apenergy.2017.01.087. URL <http://dx.doi.org/10.1016/j.apenergy.2017.01.087>
- [15] S. Pfenninger, A. Hawkes, J. Keirstead, Energy systems modeling for twenty-first century energy challenges, *Renewable and Sustainable Energy Reviews* 33 (2014) 74–86. doi:10.1016/j.rser.2014.02.003. URL <http://dx.doi.org/10.1016/j.rser.2014.02.003>
- [16] C. Breyer, D. Bogdanov, A. Aghahosseini, A. Gulagi, M. Child, A. S. Oyewo, J. Farfan, K. Sadovskaia, P. Vainikka, Solar photovoltaics demand for the global energy transition in the power sector, *Progress in Photovoltaics: Research and Applications* (2017) 505–523doi:10.1002/pip.2950. URL <http://doi.wiley.com/10.1002/pip.2950>
- [17] W. Zappa, M. Junginger, M. van den Broek, Is a 100% renewable European power system feasible by 2050?, *Applied Energy* 233-234 (January 2018) (2019) 1027–1050. doi:10.1016/j.apenergy.2018.08.109. URL <https://doi.org/10.1016/j.apenergy.2018.08.109>
- [18] T. Luz, P. Moura, 100 % Renewable energy planning with complementarity and flexibility based on a multi-objective assessment, *Applied Energy* 255 (2019) 113819. doi:10.1016/j.apenergy.2019.113819. URL <https://doi.org/10.1016/j.apenergy.2019.113819>
- [19] M. Giamalaki, T. Tsoutsos, Sustainable siting of solar power installations in Mediterranean using a GIS/AHP approach, *Renewable Energy* 141 (2019) 64–75. doi:10.1016/j.renene.2019.03.100. URL <https://doi.org/10.1016/j.renene.2019.03.100>
- [20] A. Z. Dhunny, J. R. S. Doorga, Z. Allam, M. R. Lollchund, R. Boojhawon, Identification of optimal wind , solar and hybrid wind-solar farming sites using fuzzy logic modelling, *Energy* 188 (2019) 116056. doi:10.1016/j.energy.2019.116056. URL <https://doi.org/10.1016/j.energy.2019.116056>

- [21] Q. Yang, T. Huang, S. Wang, J. Li, S. Dai, S. Wright, Y. Wang, H. Peng, A GIS-based high spatial resolution assessment of large-scale PV generation potential in China, *Applied Energy* 247 (2019) 254–269. doi:10.1016/j.apenergy.2019.04.005.
URL <https://doi.org/10.1016/j.apenergy.2019.04.005>
- [22] L. Castro-Santos, G. P. Garcia, T. Simões, A. Estanqueiro, Planning of the installation of offshore renewable energies: A GIS approach of the Portuguese roadmap, *Renewable Energy* 132 (2019) 1251–1262. doi:10.1016/j.renene.2018.09.031.
URL <https://linkinghub.elsevier.com/retrieve/pii/S0960148118310942>
- [23] S. Ali, J. Taweekun, K. Techato, J. Waewsak, S. Gyawali, GIS based site suitability assessment for wind and solar farms in Songkhla, Thailand, *Renewable Energy* 132 (2019) 1360–1372. doi:<https://doi.org/10.1016/j.renene.2018.09.035>.
URL <http://www.sciencedirect.com/science/article/pii/S096014811831098X>
- [24] H. Z. Al Garni, A. Awasthi, Solar PV power plant site selection using a GIS-AHP based approach with application in Saudi Arabia, *Applied Energy* 206 (2017) 1225–1240. doi:10.1016/j.apenergy.2017.10.024.
URL <https://doi.org/10.1016/j.apenergy.2017.10.024>
- [25] T. R. Ayodele, A. S. Ogunjuyigbe, O. Odigie, J. L. Munda, A multi-criteria GIS based model for wind farm site selection using interval type-2 fuzzy analytic hierarchy process: The case study of Nigeria, *Applied Energy* 228 (2018) 1853–1869. doi:10.1016/j.apenergy.2018.07.051.
URL <https://doi.org/10.1016/j.apenergy.2018.07.051>
- [26] J. Jung, S. U. Han, B. Kim, Digital numerical map-oriented estimation of solar energy potential for site selection of photovoltaic solar panels on national highway slopes, *Applied Energy* 242 (February) (2019) 57–68. doi:10.1016/j.apenergy.2019.03.101.
URL <https://doi.org/10.1016/j.apenergy.2019.03.101>
- [27] CTG, Programmation pluriannuelle de l'énergie (PPE) 2016-2018 et 2019-2023 de la Guyane, Tech. rep., Collectivité Territoriale de Guyane (2017).
- [28] S. L. Franklin, R. S. Pindyck, Tropical forests, tipping points, and the social cost of deforestation, *Ecological Economics* 153 (2018) 161–171. doi:10.1016/j.ecolecon.2018.06.003.
URL <https://doi.org/10.1016/j.ecolecon.2018.06.003>
- [29] A. Etxegarai, P. Eguia, E. Torres, A. Iturregi, V. Valverde, Review of grid connection requirements for generation assets in weak power grids, *Renewable and Sustainable Energy Reviews* 41 (2015) 1501–1514. arXiv:arXiv:1011.1669v3, doi:10.1016/j.rser.2014.09.030.
URL <http://www.sciencedirect.com/science/article/pii/S136403211400803X>
- [30] B. L. Gorissen, İhsan Yanıkoğlu, D. den Hertog, A practical guide to robust optimization, *Omega* 53 (2015) 124 – 137. doi:<https://doi.org/10.1016/j.omega.2014.12.006>.
URL <http://www.sciencedirect.com/science/article/pii/S0305048314001698>
- [31] N. Al-Kurdi, B. Pillot, C. Gervet, L. Linguet, Towards robust scenarios of spatio-temporal renewable energy planning: A GIS-RO approach, in: T. Schiex, S. de Givry (Eds.), *International Conference on the Principles and Practice of Constraint Programming, CP'19*, Springer Verlag, 2019, pp. 729–747.
- [32] İhsan Yanıkoğlu, B. L. Gorissen, D. den Hertog, A survey of adjustable robust optimization, *European Journal of Operational Research* (2018) 1 – 15doi:<https://doi.org/10.1016/j.ejor.2018.08.031>.
URL <http://www.sciencedirect.com/science/article/pii/S0377221718307264>
- [33] D. Juffe-Bignoli, H. Bingham, B. MacSharry, M. Deguignet, A. Milam, N. Kingston, *World database on protected areas – User manual 1.4*, United Nations Environment Programme - World Conservation Monitoring Centre, 219 Huntingdon Road, Cambridge, UK (2016).

- [34] ONF Guyane, Occupation du sol en 2015 sur la bande littorale de la Guyane et son évolution entre 2005 et 2015, Direction Régionale ONF Guyane (2017).
- [35] IGN, **BD TOPO® Version 2.2 - Descriptif de contenu**, Institut Géographique National, Paris, France (2018).
URL http://professionnels.ign.fr/doc/DC_BDTopo_2-2.pdf
- [36] EDF, **Open Data EDF Guyane**, accessed: 2018-12-10 (2018).
URL <https://opendata-guyane.edf.fr/pages/home/>
- [37] T. Albarelo, I. Marie-Joseph, A. Primerose, F. Seyler, L. Wald, L. Linguet, **Optimizing the heliosat-II method for surface solar irradiation estimation with GOES images**, Canadian Journal of Remote Sensing 41 (2) (2015) 86–100. doi:10.1080/07038992.2015.1040876.
URL <https://doi.org/10.1080/07038992.2015.1040876>
- [38] E. Fillol, T. Albarelo, A. Primerose, L. Wald, L. Linguet, **Spatiotemporal indicators of solar energy potential in the Guiana Shield using GOES images**, Renewable Energy 111 (2017) 11–25. doi:10.1016/j.renene.2017.03.081.
URL <http://dx.doi.org/10.1016/j.renene.2017.03.081>
- [39] GDAL/OGR contributors, **GDAL/OGR Geospatial Data Abstraction software Library**, Open Source Geospatial Foundation (2018).
URL <http://gdal.org>
- [40] W. McKinney, **Data structures for statistical computing in python**, in: Proceedings of the 9th Python in Science Conference, 2010, pp. 51–56.
URL <https://conference.scipy.org/proceedings/scipy2010/pdfs/mckinney.pdf>
- [41] S. van der Walt, S. C. Colbert, G. Varoquaux, **The numpy array: A structure for efficient numerical computation**, Computing in Science Engineering 13 (2) (2011) 22–30. doi:10.1109/MCSE.2011.37.
- [42] H. Woo, M. Acuna, M. Moroni, M. S. Taskhiri, P. Turner, **Optimizing the location of biomass energy facilities by integrating Multi-Criteria Analysis (MCA) and Geographical Information Systems (GIS)**, Forests 9 (10) (2018) 1–15. doi:10.3390/f9100585.
URL <http://www.mdpi.com/1999-4907/9/10/585>
- [43] J. Arnott, **Splitting large polygons for faster intersections**, accessed: 2019-10-03 (2016).
URL <https://snorfalorpagus.net/blog/2016/03/13/splitting-large-polygons-for-faster-intersections/>
- [44] G. Karypis, V. Kumar, **A fast and highly quality multilevel scheme for partitioning irregular graphs**, SIAM Journal on Scientific Computing 20 (1) (1999) 359–392.
- [45] G. Karypis, **METIS: A software package for partitioning unstructured graphs, partitioning meshes, and computing fill-reducing orderings of sparse matrices**, Department of Computer Science & Engineering, University of Minnesota, Minneapolis, MN 55455 (2013).
- [46] T. R. Teixeira, C. A. A. S. Ribeiro, A. R. dos Santos, G. E. Marcatti, A. S. Lorenzon, N. L. M. de Castro, G. F. Domingues, H. G. Leite, S. J. M. da Costa de Menezes, P. H. S. Mota, L. A. de Almeida Telles, R. da Silva Vieira, **Forest biomass power plant installation scenarios**, Biomass and Bioenergy 108 (2018) 35 – 47. doi:<https://doi.org/10.1016/j.biombioe.2017.10.006>.
URL <http://www.sciencedirect.com/science/article/pii/S0961953417303227>
- [47] H. Z. Al Garni, A. Awasthi, **Chapter 2 - solar pv power plants site selection: A review**, in: I. Yahyaoui (Ed.), Advances in Renewable Energies and Power Technologies, Elsevier, 2018, pp. 57 – 75. doi:<https://doi.org/10.1016/B978-0-12-812959-3.00002-2>.
URL <http://www.sciencedirect.com/science/article/pii/B9780128129593000022>
- [48] V. A. Graham, K. G. T. Hollands, **A method to generate synthetic hourly solar radiation globally**, Solar Energy 44 (6) (1990) 333–341. doi:10.1016/0038-092X(90)90137-2.

- [49] R. Aguiar, M. Collares-Pereira, TAG: A time-dependent, autoregressive, Gaussian model for generating synthetic hourly radiation, *Solar Energy* 49 (3) (1992) 167–174. doi:10.1016/0038-092X(92)90068-L.
- [50] A. Ben-Tal, A. Nemirovski, Robust solutions of uncertain linear programs, *Oper. Res. L* (25) (1999) 1–13.
- [51] V. Gabrel, C. Murat, A. Thiele, Recent advances in robust optimization: An overview, *European Journal of Operational Research* 235 (3) (2014) 471–483. doi:10.1016/j.ejor.2013.09.036.
URL <http://dx.doi.org/10.1016/j.ejor.2013.09.036>
- [52] J. W. Chinneck, K. Ramadan, Linear programming with interval coefficients, *J. Operational Research Society* (51 (2)) (2000) 209 – 220.
- [53] EDF, Systèmes énergétiques insulaires Guyane - bilan prévisionnel de l'équilibre offre / demande d'électricité, Tech. rep., EDF - Direction des Systèmes Énergétiques Insulaires, Paris, France (2017).
- [54] D. Bogdanov, C. Breyer, North-East Asian Super Grid for 100% renewable energy supply: Optimal mix of energy technologies for electricity, gas and heat supply options, *Energy Conversion and Management* 112 (2016) 176–190. doi:10.1016/j.enconman.2016.01.019.
URL <http://dx.doi.org/10.1016/j.enconman.2016.01.019>
- [55] A. Heydari, A. Askarzadeh, Optimization of a biomass-based photovoltaic power plant for an off-grid application subject to loss of power supply probability concept, *Applied Energy* 165 (2016) 601–611. doi:10.1016/j.apenergy.2015.12.095.
- [56] NREL, Distributed generation renewable energy estimate of costs, <https://www.nrel.gov/analysis/tech-lcoe-re-cost-est.html> (2016).
- [57] I. Fayad, N. Baghdadi, V. Gond, J. S. Bailly, N. Barbier, M. E. Hajj, F. Fabre, Coupling potential of ICESat/GLAS and SRTM for the discrimination of forest landscape types in French Guiana, *International Geoscience and Remote Sensing Symposium (IGARSS)* 33 (2014) 2046–2049. doi:10.1109/IGARSS.2014.6946866.
URL <http://dx.doi.org/10.1016/j.jag.2014.04.005>
- [58] A. Calmont, Forests in French Guiana, between development and ecosystem protection, *VertigO - la revue électronique en sciences de l'environnement* [online] special issue 14 (2012). doi:10.4000/vertigo.12402.
URL <http://journals.openedition.org/vertigo/12402>
- [59] ONF Guyane, Programme régional de mise en valeur forestière pour la production de bois d'œuvre - période 2015-2019, Tech. rep., Direction Régionale ONF Guyane (2015).
- [60] J. J. Watson, M. D. Hudson, Regional Scale wind farm and solar farm suitability assessment using GIS-assisted multi-criteria evaluation, *Landscape and Urban Planning* 138 (2015) 20–31. doi:10.1016/j.landurbplan.2015.02.001.
URL <http://dx.doi.org/10.1016/j.landurbplan.2015.02.001>
- [61] A. Sultana, A. Kumar, Optimal siting and size of bioenergy facilities using geographic information system, *Applied Energy* 94 (2012) 192–201. doi:10.1016/j.apenergy.2012.01.052.
URL <http://dx.doi.org/10.1016/j.apenergy.2012.01.052>
- [62] S. H. Siyal, U. Mörtberg, D. Mentis, M. Welsch, I. Babelon, M. Howells, Wind energy assessment considering geographic and environmental restrictions in Sweden: A GIS-based approach, *Energy* 83 (2015) 447–461. doi:10.1016/j.energy.2015.02.044.
- [63] Q. Wang, M. M'ikiugu, I. Kinoshita, Q. Wang, M. M. M'ikiugu, I. Kinoshita, A GIS-Based Approach in Support of Spatial Planning for Renewable Energy: A Case Study of Fukushima, Japan, *Sustainability* 6 (4) (2014) 2087–2117. doi:10.3390/su6042087.
URL <http://www.mdpi.com/2071-1050/6/4/2087>

- [64] J. Remund, S. Müller, S. Kunz, B. Huguenin-Landl, C. Studer, R. Cattin, [Meteonorm Handbook part II: Theory, Global Meteorological Database Version 7 Software and Data for Engineers, Planers and Education](#) (2018). doi:10.1111/j.1600-0668.2004.00281.x.
URL <http://www.meteonorm.com>
- [65] R. J. Aguiar, M. Collares-Pereira, J. P. Conde, Simple procedure for generating sequences of daily radiation values using a library of Markov transition matrices, *Solar Energy* 40 (3) (1988) 269–279. doi:10.1016/0038-092X(88)90049-7.
- [66] W. F. Holmgren, C. W. Hansen, M. A. Mikofski, Pylib Python: a Python package for modeling solar energy systems, *Journal of Open Source Software* 3 (29) (2018) 884. doi:10.21105/joss.00884.
- [67] X. Camps, G. Velasco, J. de la Hoz, H. Martín, [Contribution to the PV-to-inverter sizing ratio determination using a custom flexible experimental setup](#), *Applied Energy* 149 (2015) 35–45. doi:10.1016/j.apenergy.2015.03.050.
URL <http://dx.doi.org/10.1016/j.apenergy.2015.03.050>
- [68] P. Gilman, SAM photovoltaic model technical reference, Tech. rep., National Renewable Energy Laboratory, Colorado, USA (2015).
- [69] E. G. Evseev, A. I. Kudish, [The assessment of different models to predict the global solar radiation on a surface tilted to the south](#), *Solar Energy* 83 (3) (2009) 377–388. doi:10.1016/j.solener.2008.08.010.
URL <http://dx.doi.org/10.1016/j.solener.2008.08.010>
- [70] D. G. Erbs, S. A. Klein, J. A. Duffie, Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation, *Solar Energy* 28 (4) (1982) 293–302. doi:10.1016/0038-092X(82)90302-4.
- [71] H. K. Elminir, Experimental and theoretical investigation of diffuse solar radiation: Data and models quality tested for Egyptian sites, *Energy* 32 (1) (2007) 73–82. doi:10.1016/j.energy.2006.01.020.
- [72] D. L. Wall, Parametric estimating for early electric substation construction cost, Master’s thesis, The University of Texas at Austin, Texas, USA (2009).
- [73] EDF, Systèmes énergétiques insulaires Guyane - bilan prévisionnel de l’équilibre offre / demande d’électricité, Tech. rep., EDF - Direction des Systèmes Énergétiques Insulaires, Paris, France (2015).
- [74] G. Notton, L. Stoyanov, M. Ezzat, V. Lararov, S. Diaf, C. Cristofari, [Integration limit of renewable energy systems in small electrical grid](#), *Energy Procedia* 6 (2011) 651 – 665. doi:<https://doi.org/10.1016/j.egypro.2011.05.075>.
URL <http://www.sciencedirect.com/science/article/pii/S187661021101486X>
- [75] P. Haurant, De la sélection multicritère de parcs photovoltaïques à la cartographie et l’étude de l’intermittence de la ressource, Ph.D. thesis, Université de Corse (2012).
- [76] W. T. Jewell, T. D. Unruh, Limits on cloud-induced fluctuation in photovoltaic generation, *IEEE Transactions on Energy Conversion* 5 (1) (1990) 8–14. doi:10.1109/60.50805.