02443
Stochastic Simulation

DTU Compute
Department of Applied Mathematics and Computer Science

Final Project Report: Simulation of Solar Energy

Brian Tang (s192389), Reda Lahlou (s192431), Vlad Tampiza (s200384)

June 29, 2020

Contents

List of Figures

Abstract

In the past decades, the electric power demand per household has considerably increased. The electric power production has respectively grown to meet the massive demands. However, the traditional methods of electricity production are known to have an irreversible negative impact on the earth's atmosphere. In a context in which the power demand is assumed to continue to increase in the future, the most viable solution to the existing problem becomes green energy systems. This project thus reviews the efficiency of a solar energy system of mainly two components: the solar panels and the storage batteries. Besides, the project specifically addresses different types of households such as one, two or three-person households. The main product of this report is therefore simulations and statistics regarding the efficiency of a green energy system depending on variables such as the gender, age or season, as well as system characteristics such as battery storage levels of solar panels output.

The survey data [2] regarding the power consumption of a sample population has been provided by the teacher and the teaching assistants of the Stochastic Simulation course. Besides, the data regarding the power production using a solar panel system has been extracted from open source data catalogs of solar radiation levels [8]. The irradiation level data have been observed and registered in London, United Kingdom in 2005.

1 Introduction

Green energy systems have existed for a long time. However, the high costs, steady maintenance and other factors have impeded the implementation of such systems to a sensible high extent in the modern world. In order to provide a thorough analysis of their impact and promote their consideration as alternatives, countless simulations and optimization methods are discussed, run and verified through statistical analysis.

This project aims to present the efficiency of solar systems implemented in the life of various households, from individuals living alone to families with children. Moreover, the project also addresses different strategies of usage of the solar panels in different seasons. All the results presented in this project contain an uncertainty measurement based on several simulations so that the real value of the estimates lie withing a shown confidence interval.

1.1 Project Description

The first step in this project was data preparation, in which the data was filtered, resampled and finally split according to various categories and theoretical considerations. The time-discretization and thus resolution of the simulators was set as a **constant 1 hour sample** due to the project focus falling on the variety of influencing variables instead of activity accuracy. A sample activity list from the survey data can be seen in figure 1.

	individual	t_start	whatdoing	year	month	hour	serial	num_adult	num_child	sex	age
0	11070419-1	2014-04-18 22:00:00+00:00	0	2014	4	22	11070419	3	2	1	42
1	11070419-1	2014-04-18 23:00:00+00:00	0	2014	4	23	11070419	3	2	1	42
2	11070419-1	2014-04-19 00:00:00+00:00	0	2014	4	0	11070419	3	2	1	42
3	11070419-1	2014-04-19 01:00:00+00:00	0	2014	4	1	11070419	3	2	1	42
4	11070419-1	2014-04-19 02:00:00+00:00	0	2014	4	2	11070419	3	2	1	42

Figure 1: Sample of activity list

A detailed theoretical explanation is provided in sections 2 and 3.

The next step consisted in finding solar radiation data for a specific location and calculating the maximum power production based on the physical properties of the Solar Panels used in the simulation. Subsequently, the activity simulators were created with the help of Inhomogeneous Markov Chains (IMC) [4]. The data preparation and simulators section ends with discussing the validity of the proposed activity simulators.

A control system based on the desired battery output and storage level is built and provided as a solution to the extra or missing energy differences between the solar panel production and actual consumption. More specifically, the control system aims to handle the energy realization of the system. The control system follows a static strategy in which the excess or lacking energy is either imported or exported.

Finally, the activity simulator is run together with the control system and thus final statistical estimators are provided for the efficiency of the solar energy systems. The project ends with conclusions, plots showing the energy realization of the system, recommendations and optimization ideas for the future.

2 Data Preparation

2.1 Exploratory data analysis

A first investigation in the 4 datasets provided shows there is a great diversity of households. There are men or women living alone and also households up to 8 people. Besides, participant's age is evenly distributed below 60 as highlighted by figure 2. Therefore, there is an interesting opportunity to define categories within the datasets and compare their activities.

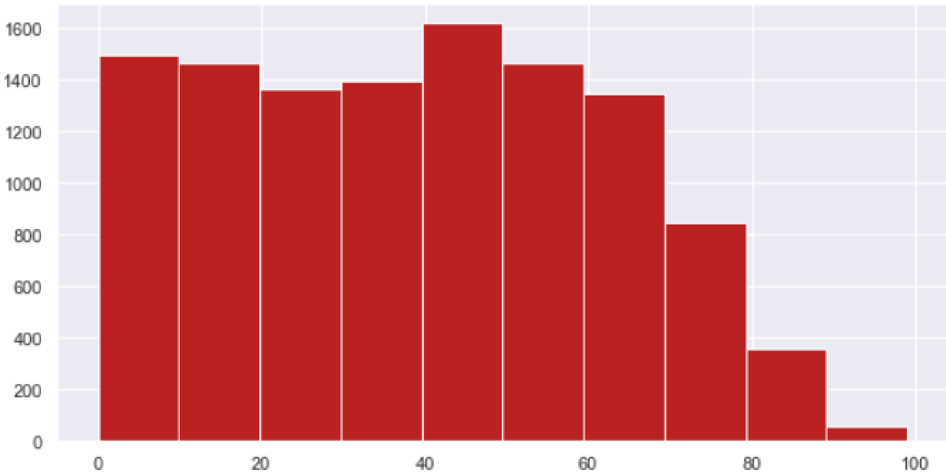


Figure 2: Age of individuals

A second insight drawing from the data is that the frequency of each activity differs a lot (figure 3). The most common one is by far activity 0 which corresponds to *"personal care"*.

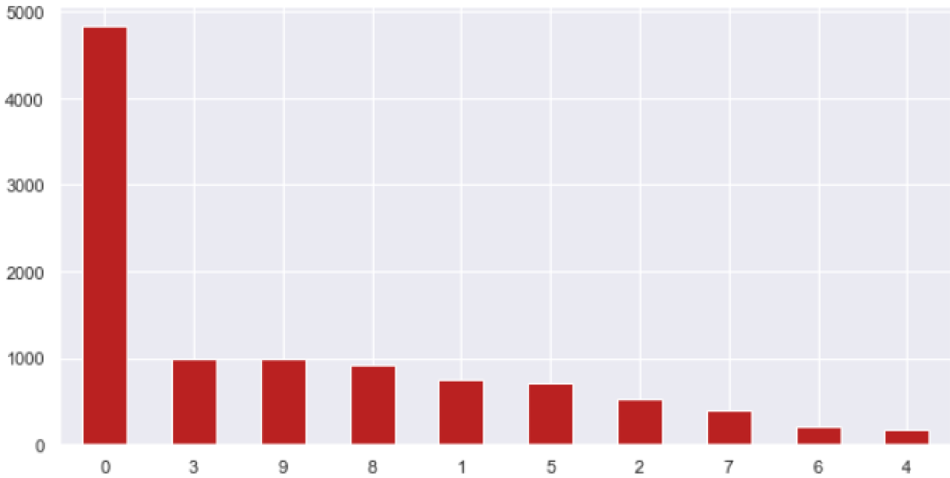


Figure 3: Most common activities

Focusing on a single individual's activities (figure 4) confirms the fact that activity 0 is predominant. It mostly occurs during the night. Other activities are scattered during the day, often not lasting more than one hour.

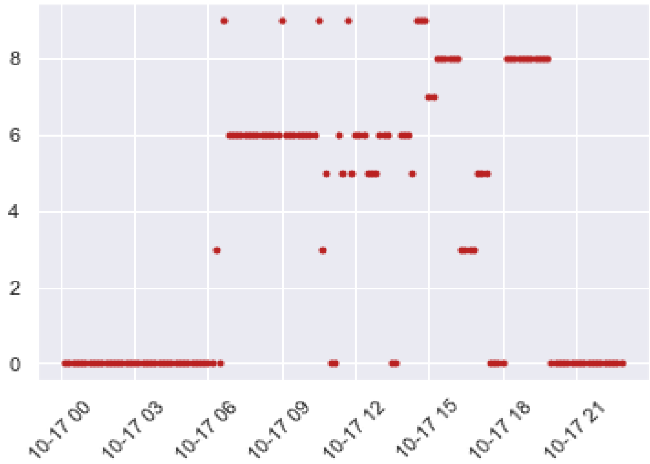


Figure 4: Activity of a single person during one day

2.2 Data time re-sampling

The main difficulty in the data preparation came from the structure of the data. They have a 10 minutes resolution. However, they are not time series given that the sequence of activities is not recorded at successive equally spaced points in time. If the activity lasts more than 10 minutes, its duration is recorded in another variable. Filling all "10 minutes time slots" leads to data that seems closer to a time series.

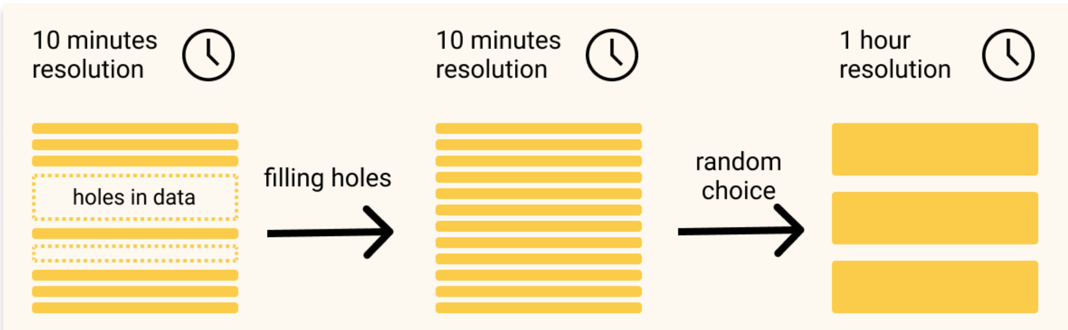


Figure 5: Defining a new time resolution process

Then, we wanted to change data resolution, going from a 10 minutes resolution to a 1-hour resolution. Working with 10 minutes intervals would require a lot of computational resources given that the aim is to simulate activities for a whole year. However, many individuals actually do multiple activities within one hour, as emphasized in figure 4. Therefore, a time resampling implies a loss of data. A random weighted choice is done to select which activity to keep. An activity which has occurred more is more likely to be selected.

2.3 Solar irradiance data

The data for solar irradiance was taken from the SoDa project [8]. They were generated for a location close to London, for the year 2005.

The global horizontal irradiance U $[kWh/m^2]$ was given for every hour of the year. The corresponding energy E $[kWh]$ was then calculated as follows:

$$E = U * A * \eta \tag{1}$$

, where A is the area of the solar panel in m^2, and η the efficiency of the panels. η has been taken equal to 15%, which corresponds to a realistic value found in the literature [7]. The tilt of the panels has been neglected.

It should be noted that 79 values out of 8760 were missing. They were approximated by taking the average irradiance during the corresponding hour in the corresponding month.

2.4 Inhomogeneous Markov Chains

To build an activity simulator, we first want to define transition matrices for the Markov chain. There are 24 matrices because we need one for each hour of the day. That is why the Markov chain is time-inhomogeneous. The size of the matrices is 10 X 10 since there are 10 activities.

The probability to jump from activity i at hour h to activity j at hour h+1 is given by the following formula:

$$P(X_{h+1} = j | X_h = i) = \frac{Count((X_{h+1} = j, X_h = i))}{Count(X_h = i)} \tag{2}$$

Then, a Markov chain is built according to the computed transition matrices and becomes the activity simulator.

3 Activity Simulation

3.1 Data categorization

The data provided through the survey contains various information on the households and their inhabitants. In this section, the variations in activity patterns between different categories of population will be assessed. It will make it possible to build IMCs which are able to describe more specifically the behavior of each category of inhabitants.

3.1.1 Gender and age influence

It could be guessed that a child would have a different activity pattern than an adult, and that a man would have a different pattern than a woman. Figure 6 shows the density of the activities of these 3 categories during the year. An individual is considered as a child if his age is strictly lower than 18.

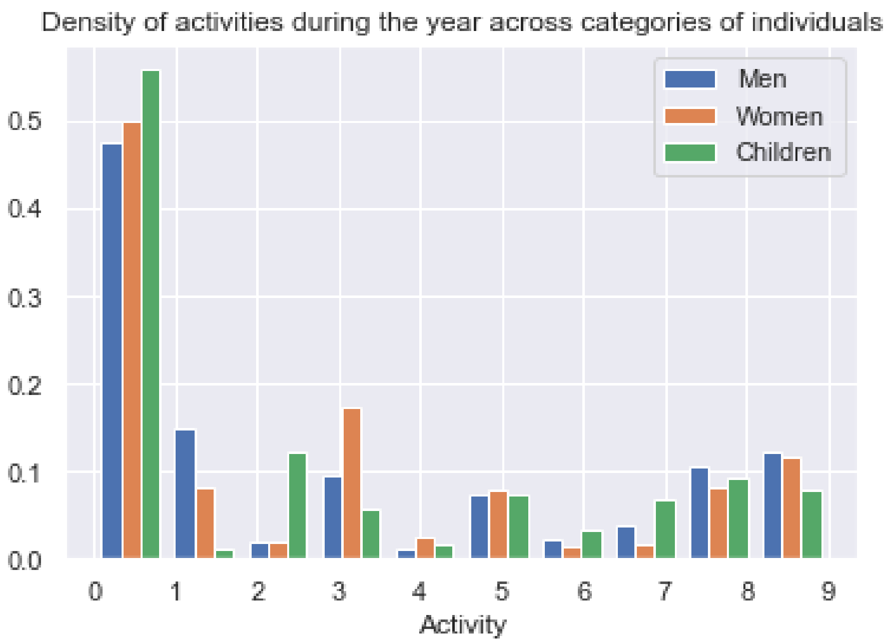


Figure 6: Density on activities depending on the individuals categories

Significant variations can be observed in the density of various activities. For example, children are over-represented in activity 2, which corresponds to *"study"*. Another example is category 3 *"household care"*, in which women are over-represented. These unbalances justify the decision to categorize the population in this way.

3.1.2 Seasonal influence

The dependency of the activity patterns on the weather is a crucial aspect of the model built in this project. Indeed, electricity consumption is usually higher during the winter, which is also a season with less sun than the others. Building an IMC by using directly all the data available for the year would prevent from observing this peak of consumption, which plays a central role in the dimensioning of a power system.

First, the variation in activities has been assessed on plot 7, which shows the density of the activities depending on the season. It should be noted that the seasons were approximated as follows: winter goes from January to March, spring for April to June, summer from July to September and fall from October to December.

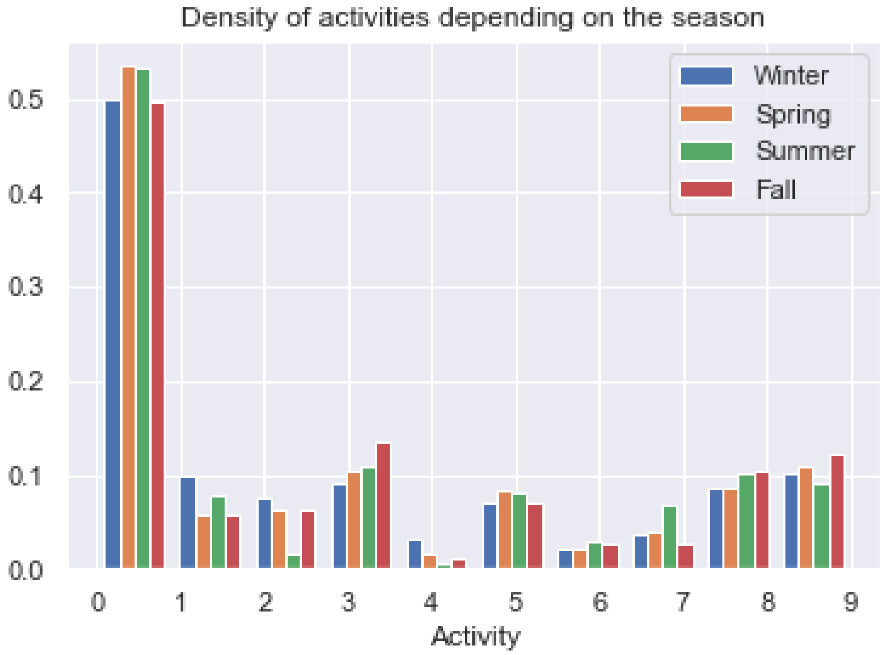


Figure 7: Density on activities depending on the season

Again, significant variations can be observed in the density of activities - although to a lesser extent than between individuals categories. For example, activity 2 *"study"* is under-represented during the summer - which is probably due to the summer holidays.

3.2 Building the activity simulator for a typical household

Four different IMCs were built for each individual category, generating activities according to the observations in the 4 distinct seasons. The figure 8 describes how these 4 activity lists were finally merged to generate an activity list for the whole year.

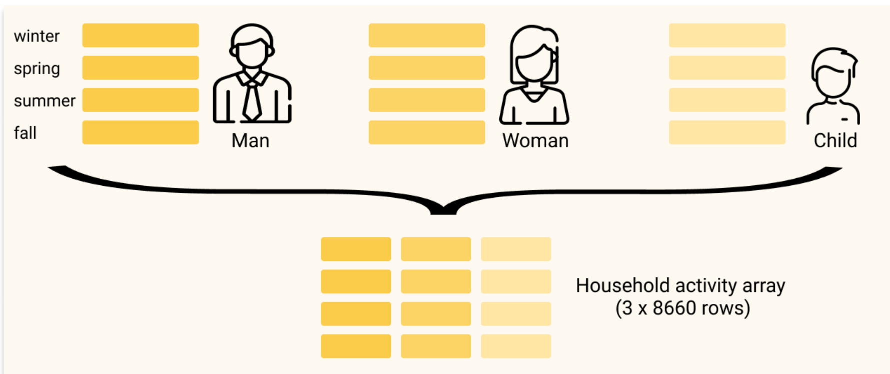


Figure 8: Categorization to compute one household activities

3.3 Validation of data models

In order to assess for the proper functioning of the 3 IMCs built in this project, a training/testing approach has been developed. For each individual category, the initial dataset was divided into 2 parts: a training set, composed of 80% of the data, and a testing set, composed of 20% of the data. The IMC is built using only the training test, and it is employed to generate activities for the period encompassed in the the testing set. The results are then compared to the actual activities observed in the testing set. The testing set was built by selecting random individuals one by one, until the observations they are linked to correspond approximately to 20% of the data. The training set then corresponds to the remaining observations. Figure 9 shows an occurrence of this validation process conducted on the woman IMC.

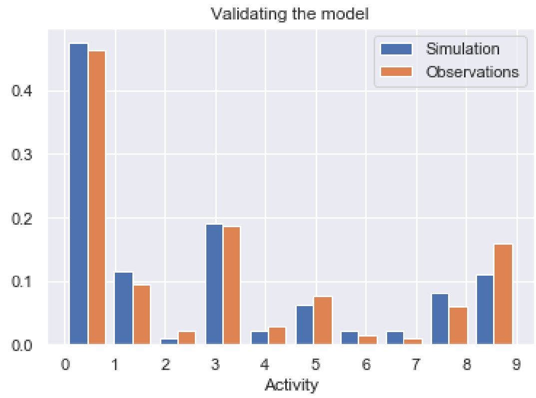
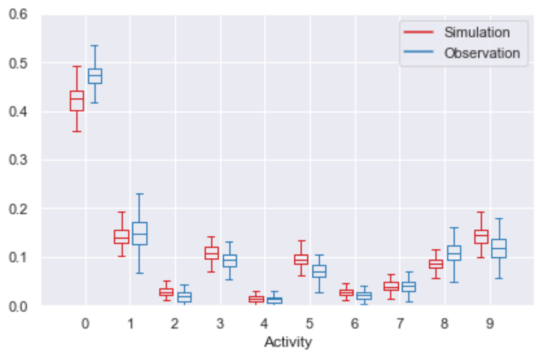
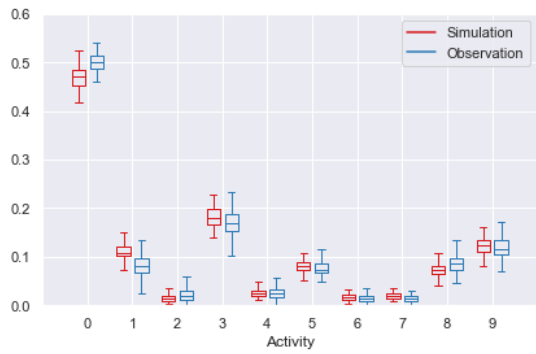


Figure 9: Density of activities in the simulated and observed datasets

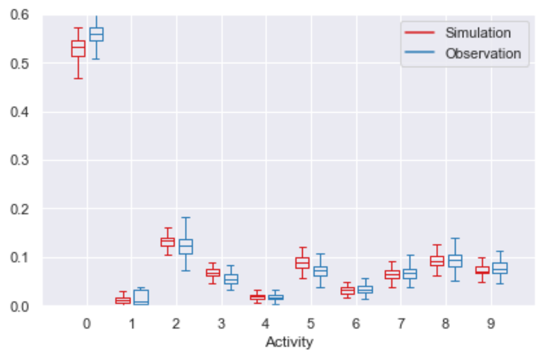
The density of the activities appears to be quite similar between the simulated and the observed activity lists. However, it has been observed that this result is highly dependent on the training and testing sets, and so subject to random variations. In order to take this phenomenon into account 100 simulations were conducted for each of the 3 IMCs. The corresponding boxplots were drawn and can be seen on figure 10.



(a) IMC validation for men



(b) IMC validation for women



(c) IMC validation for children

Figure 10: Density of activities in the simulated and observed datasets

Overall, the density of activities generated by the IMCs appears to follow the observed density, which validates their proper functioning.

However, it can be noticed that the IMCs seem to generate too few activities 0 - especially for men. The reason may be that the testing and training sets are built by merging the activities of different individuals. Thus, each time 2 individuals are merged, a wrong activity jump appears between the last activity of the first individual and the last activity of the second individual. Another explanation could be that the categories do not represent homogeneous populations. Some other categorizations could be studied, following for example the type of the household or the employment situation of the individual.

4 Control System

The Control System is used to manage the energy input and output of the solar-battery system. More specifically, it saves extra energy produced during sunny days, and provides steady electricity input when the consumed power is greater than the produced power. Given that the control system strategy does not change with time, the strategy can be defined as static. A more detailed picture showing how the energy is handled after transferring the solar radiation into electric power is shown in the figure below.

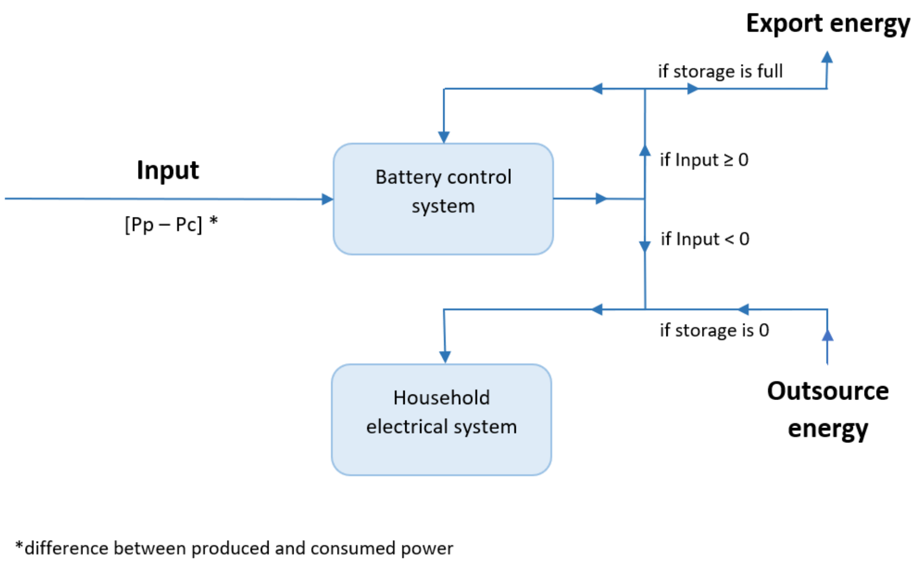


Figure 11: Control System Strategy representation

4.1 Initial control considerations and limitations

The most critical aspect of the control system is to be able to store as much as possible energy and, at the same time, use this energy as efficiently as possible. Therefore, after researching the market, a TESLA [6] battery was chosen to represent an applicable battery for a household solar energy system. The maximum Storage Level of the battery is given as **MaxStorage = 13.5** [**kW·h**] and the maximum Output as **MaxOutput = 5** [**kW**], meaning that the battery cannot discharge more than 5kWh in one hour of continuous activity. It is generally assumed that a multiple battery system strategy will not differ from a singular battery system, thus, for an increased number of batteries the maximum storage and output levels will increase linearly as $N \cdot MaxStorage$, N being the number of batteries used.

As seen in Figure 11 the extra energy is curtailed outside or exported, and the lacking energy is combined with imported energy to meet the customer demand. This simple strategy implies several key-disadvantages:

- The extra energy is lost whenever batteries are fully charged
- Energy is imported only at the point at which the battery system is fully discharged.
- The system is time-invariant, meaning that seasonal influences cannot be taken into account and the energy input is treated the same regardless of future meteorological conditions

Possible solutions to the presented disadvantages are shown in the section 4.3

4.2 Data filtering through the control system

At this stage, the activity list is transformed into power consumption with the help of the A2P matrix. It is important to emphasize the use of the "bottom-up" method [5] to express the power consumption of several household members at the same time. In this method, the overall power consumption at time t is equal to the sum of individual power consumption contributions of each of the household members. After creating the consumption and production matrices, the data is run through the battery system simulator and the final output thus becomes the imported and exported energy ratios, mean storage and output levels of the batteries and other relevant parameters presented in section 5. The following picture shows the average battery output for different scenarios for a single man household.

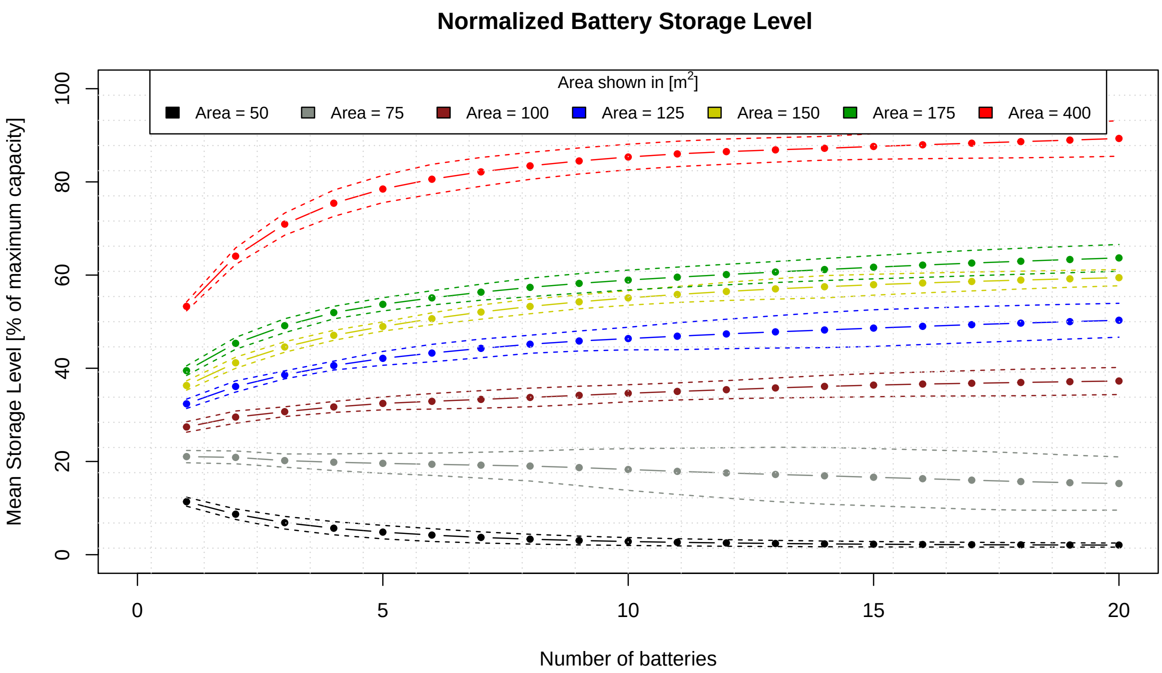


Figure 12: Average Battery Storage levels

The figure shows that for small energy production, the use of one battery is enough to store as much as possible. If one uses more batteries to store this amount of energy, the batteries will discharge extremely fast while charging them will be feasibly impossible and thus result in a mean battery level of 0 %.

On the other hand, for large energy production, the storage is more efficient as the number of batteries increases. That is, the more batteries are available to be charged, the more energy will be stored and therefore used. However, after a specific limit, the increase in the number of batteries becomes irrelevant. Regardless of the energy production volume, it is always desirable to have an average storage level greater than 50 % in order to be able to deliver maximum output at any given instant. The upcoming figure presents the Normalized Battery Output level for different scenarios.

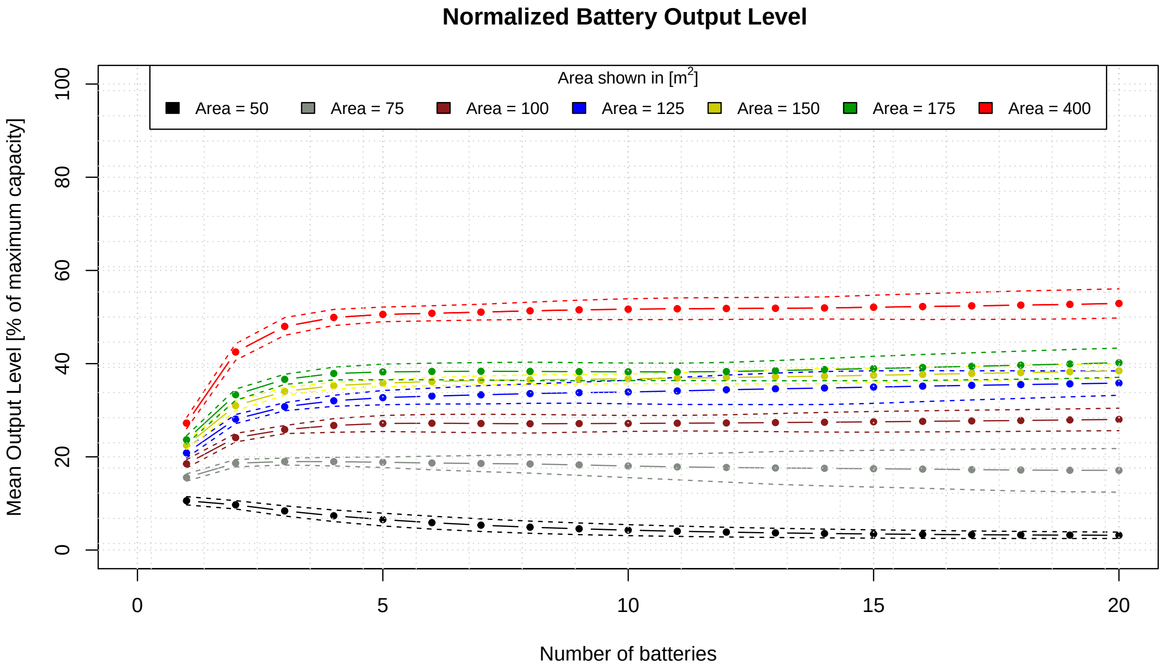


Figure 13: Average Battery Output levels

The figure clearly shows how an increased number of batteries increases the output level of the system. However, as mentioned earlier, it is not necessary to increase the number of batteries for small energy production as it would result in, generally, fully discharged batteries throughout the whole year since the input energy is not large enough to charge them accordingly. If, however, the production levels are high, it is vital to use more than just 1 battery in order to obtain an efficient system. The figure shows how the mean output level tends to be close to half of the maximum potential. This means that for such a scenario, the output would vary in between 0 and maximum of its potential i.e. the best case scenario for a green system of 100 % as the number of times the system discharges is equivalent to the number of times it charges.

4.3 Proposed optimizations to control system

It has been previously argued that a static control system might result in inefficient energy distribution and usage. This section presents several possible solutions to the mentioned disadvantages.

Sell the extra energy instead of curtail.

Although the green system efficiency will not be affected by this action, the overall system efficiency will increase according to the amount of sold energy. That is, the resources spent on imported energy can be recovered by the resources gained after selling the extra energy. Thus, the physical efficiency of the system remains the same while the overall efficiency could be increased so that the system will be self-sufficient.

Create a more flexible charging/discharging strategy.

Arguably, the system efficiency could be improved if the strategy does not allow the batteries to be fully discharged or charged before curtailing or importing energy, resulting in energy-saving strategies for the batteries.

Implement a dynamical(time-dependent) strategy.

The following suggestion is by far the most urgent implementation if one looks to improve the green system efficiency. The current strategy does not allow the system to save critical amounts of energy before consecutive days with no sunshine, and at the same time, does not take into account days with extra sunshine in which a high amount of energy is available. Thus, by introducing a strategy in which a great number of batteries work during winter, and respectively, fewer batteries are used during summer, the system could greatly benefit.

5 Results and discussion

The model built was used to simulate a whole year of electricity consumption and generation for different configurations of households. Five main statistics are calculated:

- **The system ratio:** it corresponds to the fraction of the electricity consumed that is not imported from the grid (it either comes directly from the solar panels or from the batteries)
- **The import ratio:** it corresponds to the fraction of the electricity consumed that is imported from the grid
- **The export ratio:** it corresponds to the fraction of the electricity produced that is not consumed by the household (it is either curtailed or exported to the grid)
- **The mean storage level:** it corresponds to the average storage level of the batteries, given as a fraction of the total storage capacity available
- **The mean output level:** it corresponds to the average energy output of the batteries, given as a fraction of the maximum energy output

For each simulation, each of the 5 parameters $\hat{x}$ was estimated 10 times. A 95% confidence interval was built as follows:

$$I = [E(\hat{x}) - t_{0.975,9} \cdot \sigma(\hat{x}), E(\hat{x}) + t_{0.975,9} \cdot \sigma(\hat{x})] \tag{3}$$

, where $E(\hat{x})$ and $\sigma(\hat{x})$ are respectively the expected value and the standard deviation of the parameter, and $t_{0.975,9}$ is the 97.5% quantile of the t-distribution with 9 degrees of freedom (10 simulations run minus 1 parameter estimated).

5.1 Households with single men, women and children

Three households were simulated, each one composed of a single individual: a man, a woman and a child. Figure 15 shows the evolution of the system ratio depending on the number of batteries for different sizes of solar panels.

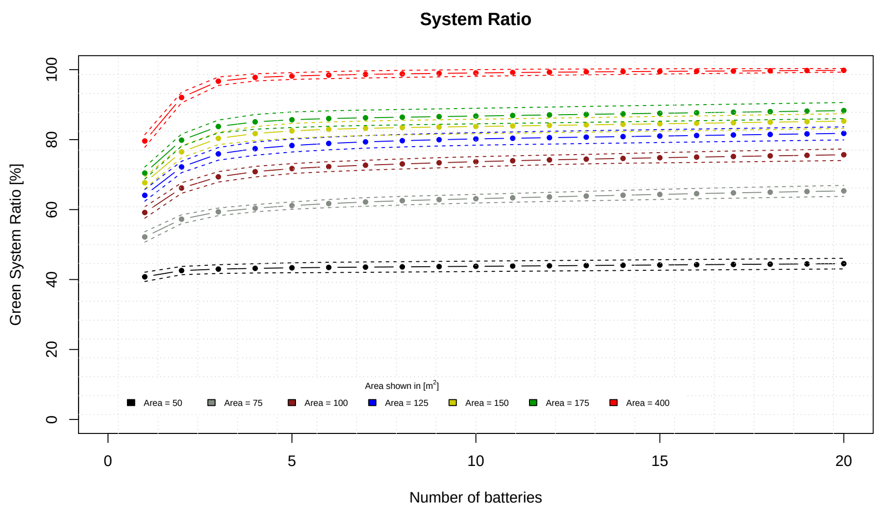
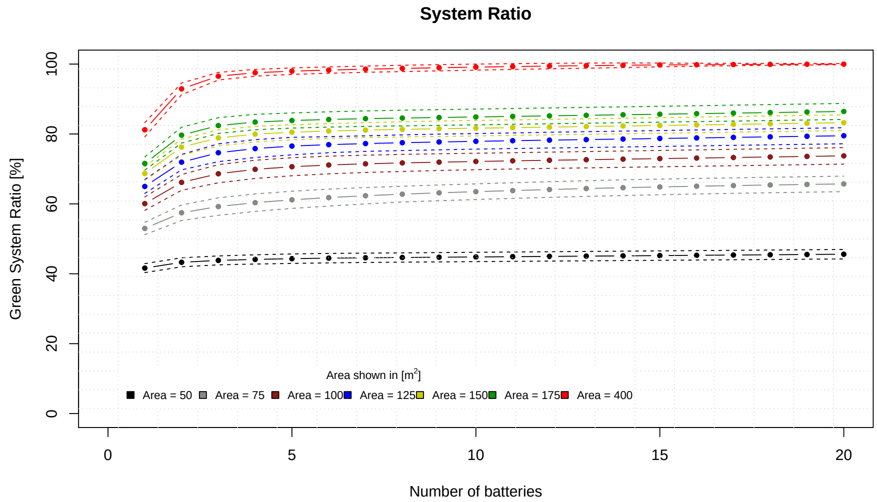
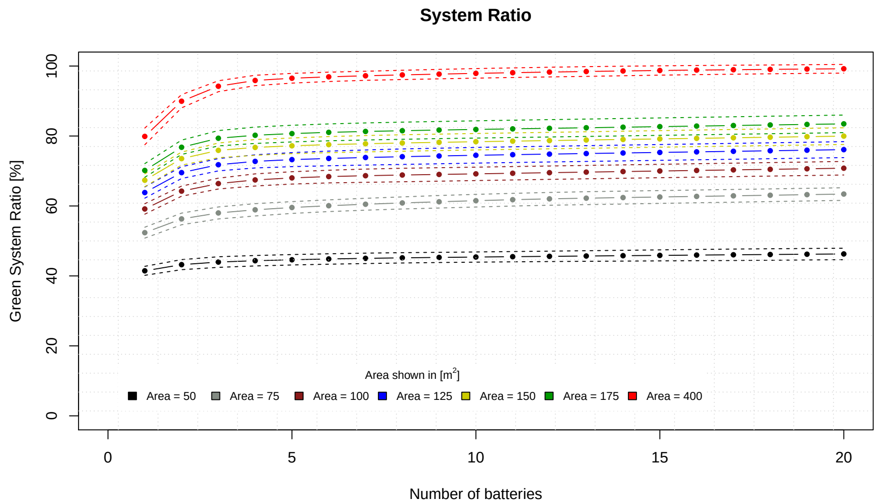


Figure 14: Single man household



(a) Single woman household



(b) Single child

Figure 15: System ratio depending on the number of the batteries for different households

It can be observed that adding new batteries increases the system ratio when the number of batteries is low. But after a certain number of batteries, the ratio reaches a plateau. The magnitude of this phenomenon depends on the size on the solar panels. The higher the size, the more quickly the ratio increases with the number of batteries (higher derivative), and the more batteries is needed to reach a plateau.

The difference between the different categories is also visible through these graphs. For example, for an area of $150m^2$, a system ratio of 80% can be reached with 3 batteries for a man, with 4 batteries for a woman, and more than 10 batteries for a child. Such an area represents a realistic value for the available rooftop space. This reflects the difference in the activity patterns between the different categories.

5.2 Household with a couple

Figure 16 shows the evolution of the system ratio depending on the number of batteries for different sizes of solar panels for a couple (man+woman).

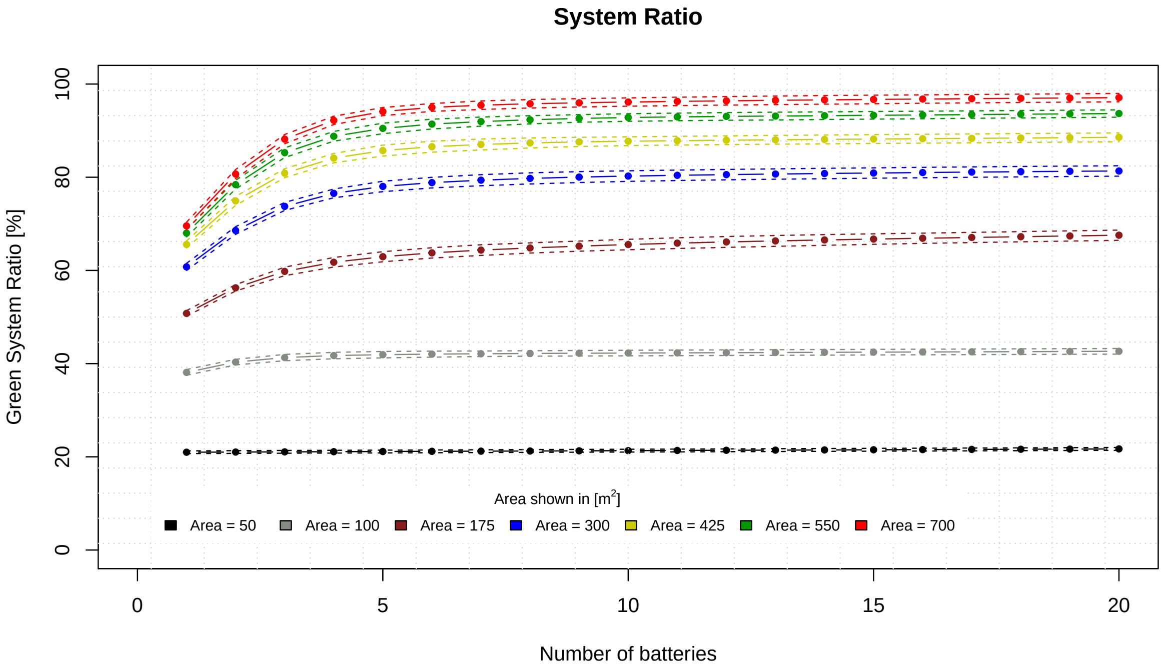


Figure 16: System ratio depending on the number of the batteries for a couple

Compared to a household with a single person, it can be observed that the size of the panels needed to be able to reach 100% ratio has increased considerably (700 m^2 against 400m^2). Moreover, for a low size of panels (50 m^2), the benefits of adding more than 1 battery is almost nonexistent.

5.3 Household with a family

Figures 18 and 19 show the evolution of all 5 indicators defined in the beginning of this section depending on the number of batteries for different sizes of solar panels for a family (man+woman+2 children).

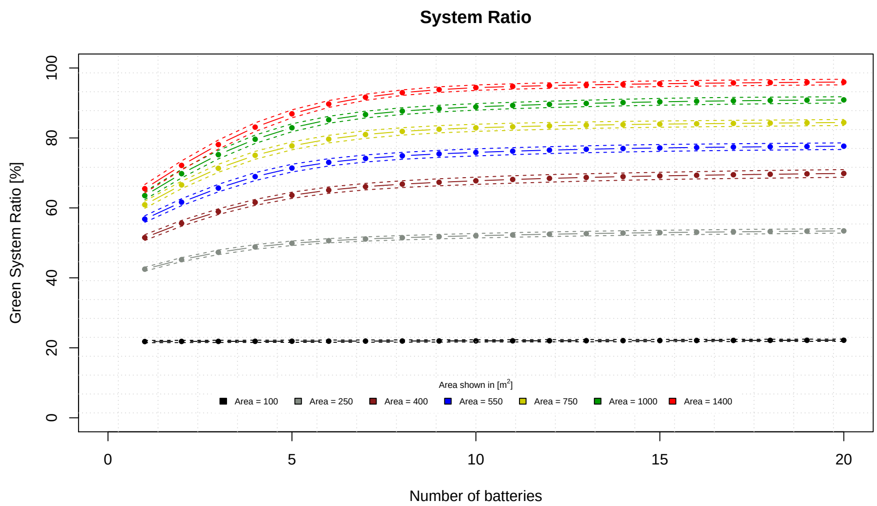
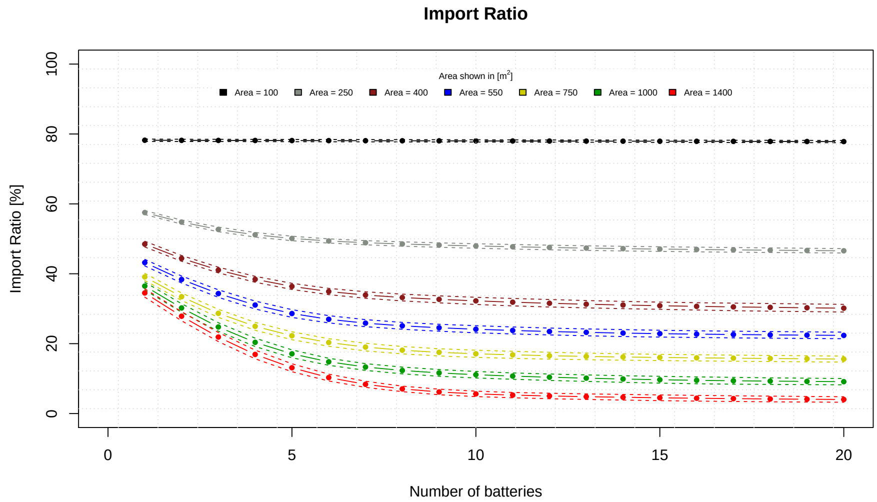
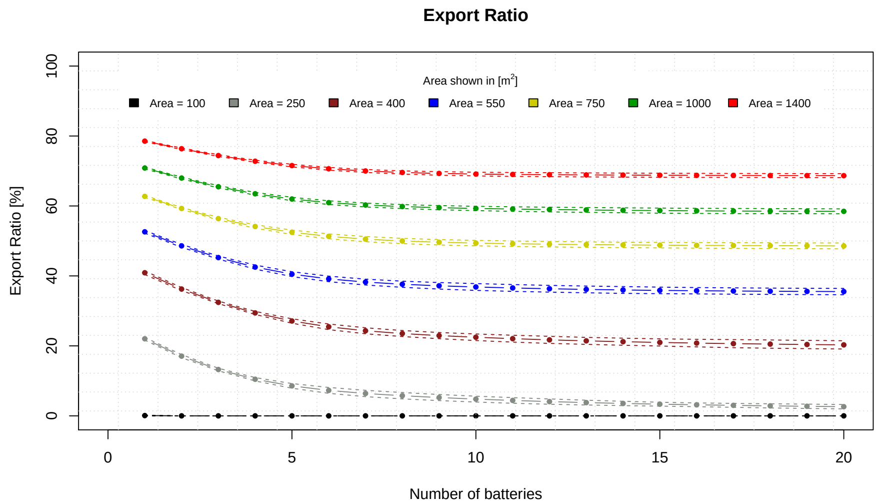


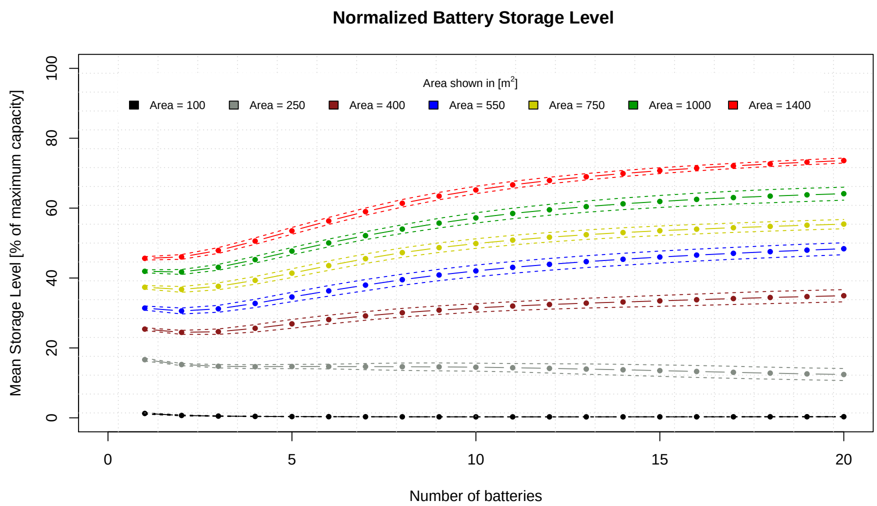
Figure 17: System ratio



(a) Import ratio



(b) Export ratio



(c) Mean battery storage level

Figure 18: Import ratio, export ratio and mean storage level depending on the number of the batteries for a family

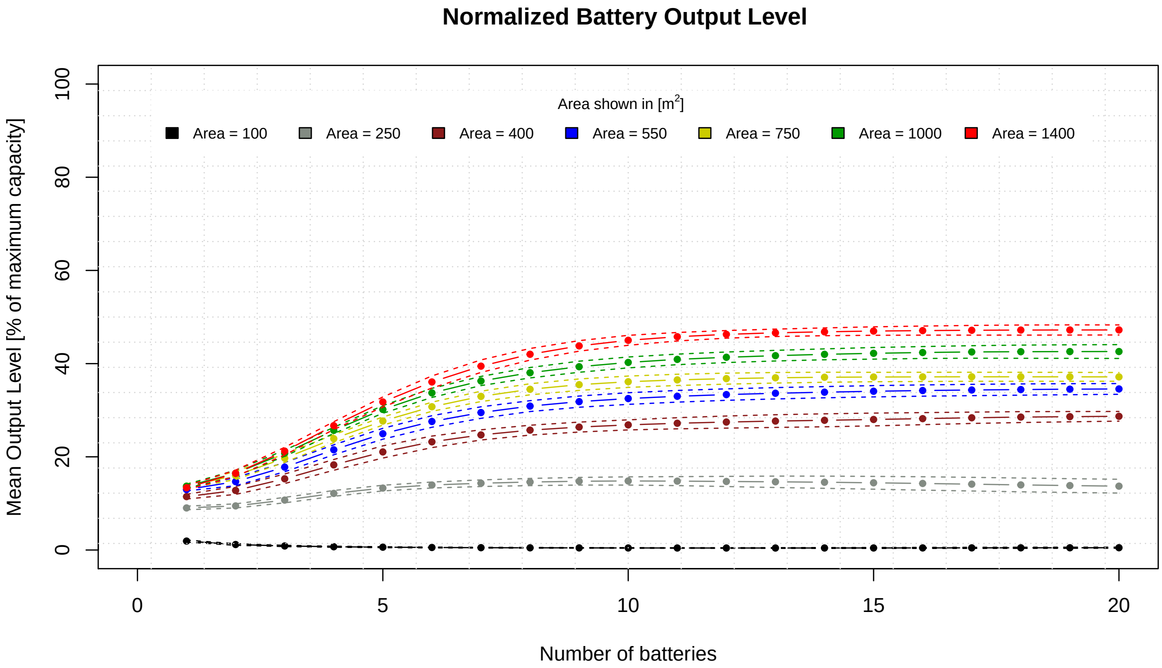


Figure 19: Mean battery output depending on the number of the batteries for a family

For a size of 100 m^2 of solar panels, the system ratio remains almost constant at a value just above 20%, while the export ratio is equal to 0%. With such a small panel size, the batteries are almost useless, since there is hardly any extra energy to store at any time. The 20% represents the energy produced by the solar panels directly consumed. This is also shown through the battery storage and output levels, which actually decrease in the beginning, to reach almost 0%. In this case, increasing the number of batteries increases the storage capacity and the maximum output, but this extra-utility is not even used, hence the decrease.

For higher sizes of panels, the following scheme is observed. The system ratio increases with the number of batteries, until it reaches a plateau. Similarly, the export ratio decreases until it reached a plateau - which is logical, as more batteries means more possibilities to store energy rather than exporting it. As for the battery indicators, they also increase until they reach a plateau, but not in the same way. They are "S-shaped", meaning that the highest derivative is not found in the beginning of the curve.

Finally, it should be noticed that running 100% on solar energy without any import from the grid does not seem feasible for a family living in a standard house (maximum roof size between 200 m^2 and 600 m^2). However, this conclusion depends on the perspective from which one looks at the system. In particular, if the surplus of energy which cannot be stored (also called "exported energy") is no longer seen as curtailed, it can be imagined that this energy is sold through the grid to another consumer, meaning that it is no longer considered lost. Thus, a family could be seen as running on 100% solar energy if the amount of energy it exports is higher than the amount of energy it imports. Following this perspective, it is possible to run on 100% solar energy with a reasonable size of panels (approximately 300 m^2).

References

[1] [Ross, Sheldon M. 2012. Simulation. Academic Press.
https://books.google.dk/books/about/Simulation.html?id=1Dwsyyty3P8C&redir_esc=y]

[2] [Gershuny, J., Sullivan, O. (2017). United Kingdom Time Use Survey, 2014-2015. UK Data Service. SN: 8128
http://doi.org/10.5255/UKDA-SN-8128-1]

[3] [UK-TUS-15 data sets
https://drive.google.com/drive/folders/1I0LBnw22jDLQIAqmA9ICLRLRXCpX8Wx1?usp=sharing]

[4] [Iversen, Emil Banning, Jan K Møller, Juan Miguel Morales, and Henrik Madsen. 2016. "Inhomogeneous Markov Models for Describing Driving Patterns."
https://ieeexplore.ieee.org/document/7414468]

[5] [Widén, Joakim, and Ewa Wäckelgård. 2010. "A High-Resolution Stochastic Model of Domestic Activity Patterns and Electricity Demand." Applied Energy 87 (6): 1880–92.
https://ideas.repec.org/a/eee/appene/v87y2010i6p1880-1892.html]

[6] [Tesla PowerWall 2 Battery for Energy Storage - Datasheet 2020
https://www.tesla.com/powerwall]

[7] [Photovoltaics Report, 2020, © Fraunhofer ISE
https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/Photovoltaics-Report.pdf]

[8] [SoDa, © 2014-2020 MINES ParisTech / Transvalor
http://www.soda-pro.com/web-services/radiation/helioclim-3-archives-for-free]