



A new Data-Driven Approach for Comparative Assessment of Baseline Load Profiles Supporting the Planning of Future Charging Infrastructure

Johannes Galenzowski
johannes.galenzowski@kit.edu
Karlsruhe Institute of Technology
Karlsruhe, Germany

Simon Waczowicz
simon.waczowicz@kit.edu
Karlsruhe Institute of Technology
Karlsruhe, Germany

Veit Hagenmeyer
veit.hagenmeyer@kit.edu
Karlsruhe Institute of Technology
Karlsruhe, Germany

ABSTRACT

In order to achieve the worldwide set ambitious climate goals, the identification and characterization of flexibility in city districts can reduce grid loads and avoid grid congestion. Unlike other flexibility indicators in the literature, the present paper introduces a new flexibility indicator that uses a data-driven approach to determine flexibility from actual measured load profiles. We present this new indicator by considering flexibility in the context of planning charging infrastructure with a valley filling approach. For this use case, we introduce a data-analysis workflow to apply the presented flexibility indicator. The described data-analysis workflow is applied to data from a real-world city district.

Based on the results from the real-world data, we show that the highest peak load and the least flexible peak are not always identical. Therefore, it is not sufficient to consider only the highest peak loads to adequately describe flexibility. Furthermore, we discuss that additional flexibility can be used as another degree of freedom to optimize the charging power or the charging duration. In the presented real-world data, we show that the maximum required charging power is determined by the most inflexible peak and can be the same or smaller for all peaks with a higher flexibility. Moreover, we highlight the difference between considering buildings individually and combining them as a district.

CCS CONCEPTS

• **Hardware** → **Smart grid**; • **Mathematics of computing** → **Time series analysis**; • **Information systems** → **Data analytics**.

KEYWORDS

flexibility, city district, charging infrastructure planning, valley filling, real-world data

ACM Reference Format:

Johannes Galenzowski, Simon Waczowicz, and Veit Hagenmeyer. 2023. A new Data-Driven Approach for Comparative Assessment of Baseline Load Profiles Supporting the Planning of Future Charging Infrastructure. In *The 14th ACM International Conference on Future Energy Systems (e-Energy '23 Companion)*, June 20–23, 2023, Orlando, FL, USA. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3599733.3600245>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

e-Energy '23 Companion, June 20–23, 2023, Orlando, FL, USA

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0227-3/23/06.

<https://doi.org/10.1145/3599733.3600245>

1 INTRODUCTION

Governments around the world set ambitious goals on carbon reduction to combat climate change. For example, the German federal government aims at reducing carbon emissions by 55 % until the year 2030 compared to 1990 [25]. Worldwide, the residential sector contributes 20 %, the service sector 14 % and the transport sector 35 % to the primary energy consumption [6]. Since those sectors all come together in city districts, city districts play a significant role in the transformation of the energy sector.

In view of advancing energy transition, the exploitation of flexibility potential in city districts becomes crucial, in order to reduce grid loads and avoid grid congestions [35]. To identify and quantify the flexibility potential, suitable methodology and flexibility indicators are needed. A variety of papers like [13, 18, 21, 22, 24] and [4, 5, 16] address the description and quantification of flexibility. Furthermore, recently published meta analyses like [20], [19] or [10] give a holistic overview of all kinds of different methods and indicators describing flexibility. Those indicators, for example, consider the spread of key figures, e.g. emissions or costs, between low and high load periods (see flexibility factor in [10]), the self-sufficiency and self-consumption indicator [10] or the capacity of controllable loads, in order to reduce energy consumption in case of a demand response measure [10].

As those papers present a high number of different indicators, it can be concluded that no single indicator fits all applications and research questions. Papers like [3] show that flexibility is always in some form linked to physical devices, that can either reduce or shift their demand in time. As governments accelerate the transition towards a sustainable transport system, electric vehicles (EVs) and charging infrastructure (CI) become increasingly significant devices for flexibility provision. For example, according to the 'Klimaschutzprogramm 2030' [25] of the German federal state, one million charging points shall be installed until the year 2030 from which, less than 7 % was reached in September 2022 [7].

With those high numbers of CI still in the planning phase, there is an urgent need for flexibility indicators, that do not only describe the operation but also help in the planning phase of new CI. However, most flexibility indicators are centered around the building heating, ventilation, and air conditioning (HVAC) sphere and are not applicable on CI. Further, these flexibility indicators are not primarily based on measurement data and require detailed modelling [10].

Paper that cover planning of CI like [13, 22, 28, 31] all approach the planning from the perspective of expected charging behavior, considering different vehicle types and usage characteristics and

concluding the required size of CI to fulfill this demand. This assumes that the on-site conditions in a city are ideally adaptable to the identified charging demand. In reality, however, CI is part of the superordinate, already existing city infrastructure (like parking spaces or the distribution grid with other existing energy consumers). Therefore, a suitable flexibility indicator needs to take the characteristics of the existing superordinate infrastructure into account.

The relevant aspects of the superordinate infrastructure are to some degree of physical and to some of legal nature. From the hierarchical perspective, a city consists of different districts, that themselves consist of different buildings (see Figure 1). From a legal standpoint, the connection point between the building and the superordinate grid, i.e. the grid connection point (GCP) is of interest (e.g. 'Kundenanlage' in the German law [8]). The GCP can be associated with a single building or aggregate multiple buildings in other cases (see Figure 1). In the private grid, subordinate to the GCP, energy can be exchanged without the supplier being regulated as a utility company, leading to reduced costs and bureaucracy [8]. Also, the energy billing is carried out on the load profile of each GCP individually [9]. While papers like [27], [23] or [36] describe the demand for flexibility in the energy market and how to quantify flexibility of existing flexible plants, we focus on characterizing the potential for future supply through future CI. Our approach makes different GCPs comparable. This information can then be used to be matched against the flexibility needs of the market and the superordinate grids to make a planning decision.

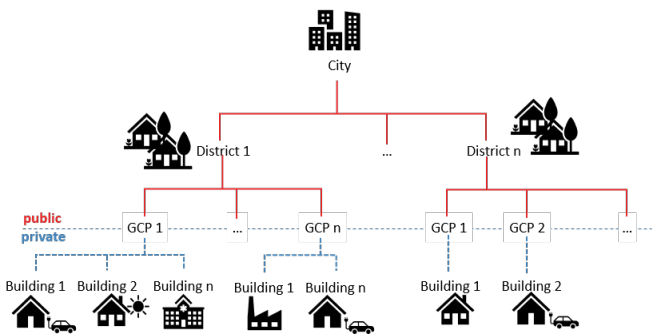


Figure 1: Hierarchy of superordinate infrastructure: Multiple buildings at one GCP (left) or one building only per GCP and coupling of buildings only through public grid (right)

Therefore, suppliers of local renewable energy and operators of the private grid in one GCPs only optimize the internal consumption inside their single GCP and do not consider the other GCPs in the district. Therefore, a suitable flexibility indicator needs to make the different GCPs comparable. Furthermore, research like [21] or [26] suggest, that aggregation of buildings to a district leads to advantages in flexibility and grid support. Since the aggregation at district level faces multiple legal problems in the current system (e.g., for Germany see [12]), the advantages of combining multiple GCPs to a district also need to be critically examined.

Besides grid congestion, cost reduction is a main motive for exploiting flexibility [10]. In countries like Germany, billing periods

can not only be on a yearly, but also on a monthly basis, according to StromNEV § 19 [9]. The charged prices are composed of an energy related share (¢/kWh) and a power related share (€/kW). A monthly billing period, in this case, means, that the power component is charged for each month individually on its highest monthly consumption peak. Therefore, a suitable flexibility indicator should also be able to characterize flexibility regarding monthly power peaks.

To avoid new and even higher peaks, charging needs to take place during off-peak times. In the context of load shaping approaches as defined in [10], this means a valley filling approach. Valley filling is widely examined in literature [1, 2, 11, 15, 17, 30, 32–34]. Those works all present different algorithm and control strategies, like decentralized coordination [32, 33], control through price signals [1, 15, 17, 33] or further grid friendly behavior, like reduced power ramps [30, 33] or the consideration of the limits of the local power transformer or a local grid [11, 15].

However, [1, 2, 11, 14, 15, 17, 30, 32–34] all target at optimizing the operation of CI during everyday production usage. Therefore, they do not focus on providing indicators that quantify the flexibility in a way to make GCPs comparable to support the planning phase of new CI.

Since, as [10] points out, valley filling aims at increasing the consumption in off-peak hours and those are characterized by a baseline load profile¹ the assumed baseline load profile is of high relevance. The baseline load profile in [33] is based on data of a Californian system operator, [34] uses measured data from a transformer in Shenzhen and [11] utilizes the load profile of the IEEE 14-bus system. However, only single load profiles with a limited time range are used by [11, 33, 34] and a comparison of multiple load profiles over a whole year is missing. In other paper like [1, 2, 15, 32] it is not even clear, what the baseline load profile is based on. A study and comparison of several measured real-world baseline load profiles and their difference or suitability for valley filling is missing. To conclude, a data-driven approach for quantifying CI flexibility of measured baseline load profiles of existing buildings or districts regarding the suitability for valley filling is needed.

When considering a data-driven approach, it should be based on features used for flexibility description, like proposed in [18]. One main feature is the time for which a flexibility source is available [18]. For example, a charging session of an EV lasts from 20 min to 6 h depending on the location type [13]. The limitation on available charging power is highest around peaks of the baseline load profile. Therefore, the flexibility indicator must be able to reflect the limited availability in the surroundings of a load peak.

To sum up, a suitable flexibility indicator needs to consider characteristics of charging infrastructure like limited availability of vehicles and limited charging power. Furthermore, it must support the planning phase of new CI. In addition, the flexibility indicator should make different GCPs comparable. And lastly, it needs to consider monthly and yearly consumption peaks and be extractable from existing baseline load profiles without further assumptions needed for vehicle or user behavior. Instead, it should only quantify

¹Baseline meaning the measured load profile under current consumption and behavioral conditions without the newly planned CI.

the suitability of a GCP for the future installation of grid friendly CI (that does not produce additional peaks besides the existing ones) through a data-analysis of baseline load profiles.

The remainder of the paper is organized as follows: Section 2 presents our methodology to calculate the flexibility indicator. In section 3, we present a real-world district that contains four different GCPs. In section 4, we discuss the application of our methodology on the real-world data. Additionally, subsection 4.2 presents the comparison of the different GCPs, subsection 4.3 examines the correlation of flexibility and peak height, and subsection 4.4 discusses flexibility on a district level compared to each GCP individually. Section 5 closes the present article, giving relevant conclusions and an outlook for future work.

2 METHODOLOGY

In section 2 we present the data-driven approach, fulfilling the requirements laid out in section 1. The full workflow of the data-analysis process is depicted in Figure 2. Since the approach is based on analyzing a baseline load profile regarding power peaks, a first step is to acquire and prepare the baseline load profiles of each GCP (subsection 2.1) and to identify the different types of peaks (subsection 2.2). In the next step, the flexibility indicator is defined and calculated for each GCP (subsection 2.3). Afterward, a filtering step is included to select only the relevant peaks of each GCP for further analysis and comparison (subsection 2.4). Furthermore, the power limitation aspect of CI is considered in subsection 2.5. The obtained results are then used to compare the different GCPs in order to answer the further research questions posed in section 1.

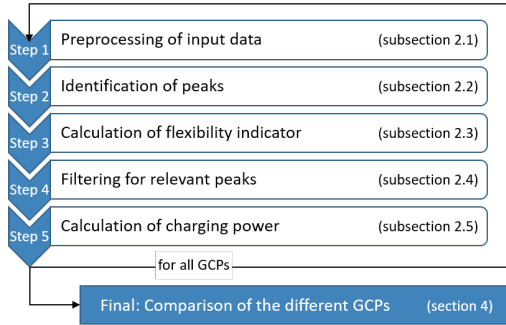


Figure 2: Overview of data-analysis workflow for all GCPs individually with a final comparison of the different GCPs

2.1 Preprocessing of input data

When considering preprocessing, features of interest are: time zone, temporal resolution, value resolution and time frame of available data. The resolution is relevant for both the temporal and the value axis, since measured values are usually both time discrete and value discrete. The temporal resolution (Δt) is assumed as 15 min equidistant. The value resolution (Δp) depends on the quality of the power measurement and can greatly influence the results. For example, at Δp , it might appear that multiple 15 min power values have the same magnitude, resulting in multiple peaks that appear to be equally important but exhibit potentially quite different load

profile shapes before and after the peak. Because billing of peaks is carried out on a time frame of a month or a year, the available data needs to cover at least one year per GCP. Further research is needed to investigate, whether the relevant key figures could also be extrapolated on a data basis of a smaller time frame.

2.2 Identification of peaks

Firstly, all potentially relevant peak data points need to be identified. These are power maxima at the level:

- p_{gy} : global year (calendar feature)
- p_{gm} : global month (calendar feature)
- p_l : local (in given number of neighboring data points)

The set of local maxima includes the monthly maxima which include the yearly maxima:

$$p_{gy} \subset p_{gm} \subset p_l \quad (1)$$

The yearly maxima (p_{gy}) and monthly maxima (p_{gm}) are defined as the highest 15 min power value in a given yearly or monthly time frame. If the power value resolution is low in relation to the absolute power level, more than one data point may have the same global power level in the given time frame. In this case, both data points are classified as yearly/monthly peaks.

The local maxima are identified by finding the highest (or equal) power value in a given surrounding of each data point (*order*) for the entire load profile. With a 15 min temporal resolution, a single day has 96 data points that result in an order of 48. This makes it possible to identify approximately one local maximum per day without being limited through calendar features (e.g., a maximum right around midnight, that would be counted twice if identified by the day as a calendar feature).

2.3 Calculation of flexibility indicator

As discussed in section 1, the flexibility is characterized by the baseline load profile during a certain time frame t_f around any peaks specified in subsection 2.2. The main function of the time frame t_f is to have a comparable metric for each load profile. The size of the time frame influences the expressiveness of the flexibility indicator. Regarding CI, the longest expected charging duration is 6 h (see section 1). Therefore, it can be assumed that t_f of 6 h is the best suite as a starting point for this paper. A t_f of 6 h leads to an evaluated range from -3 h to 3 h around each peak time (t_l). With a $\Delta t = 15$ min resolution, that leads to 12 values before and after t_l , resulting in: $t_f = \{t_l + k \cdot \Delta t | k \in [-12, 12] \cap \mathbb{Z}\}$.

The flexibility indicator is calculated for every local peak. In a first step, the flexible energy around each peak ($e_{l,flex}$) is calculated as:

$$e_{l,flex} = \sum_{i=t_l-12}^{t_l+12} (p_g - p_i) \cdot \Delta t \quad (2)$$

Figure 3 shows the baseline load profile (gray) and for the relevant time frame t_f in the surrounding of t_l also $e_{l,flex}$ in red.

The global peak power p_g can either be the yearly maximum p_{gy} or the monthly maximum, p_{gm} depending on the billing system for the individual GCP. To compare also the suitability of the two different billing methods for the different GCPs, both, the monthly and the yearly billing system are evaluated in this paper. The value of $e_{l,flex}$ is calculated for every local power maximum (p_l) and

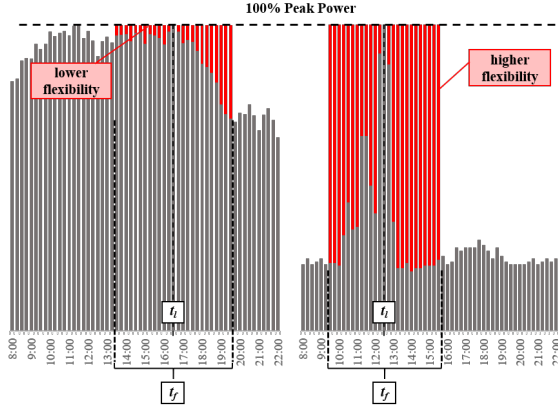


Figure 3: Examples for one peak with low flexibility (left) and one peak with high flexibility (right)

therefore, for each timestamp, t_l of a local power maximum. We define the inflexible energy $e_{l,fix}$ by:

$$e_{l,fix} = \sum_{i=t_l-12}^{t_l+12} p_i \cdot \Delta t \quad (3)$$

The flexibility indicator f is then defined as the share of flexible energy $e_{l,flex}$ of the overall energy ($e_{l,flex} + e_{l,fix}$) in a given time frame t_f . It ranges from 0.96 as an ideal peak with the highest flexibility to zero as the worst peak with the lowest flexibility²:

$$f = \frac{e_{l,flex}}{(e_{l,flex} + e_{l,fix})} \in [0, 0.96] \quad (4)$$

As mentioned, both billing methods (p_{gy} and p_{gm}) are evaluated, leading to f_y (yearly billing period) and f_m (monthly billing period).

2.4 Filtering for relevant peaks

The peak with the lowest flexibility f_{min} limits and characterizes the whole GCP. In addition, the second and third most inflexible peaks characterize the GCP. However, below a certain threshold, the peaks are too low to be representative of the characteristics of the GCP. This threshold is defined using normalized local peaks $p_{n,l} = p_l/p_{gy}$ (yearly) and $p_{n,l} = p_l/p_{gm}$ (monthly) and the characteristic flexibility f_{min} ($\min(f_y)$ or $\min(f_m)$) correspondingly as:

$$\text{if } (p_{n,l} + f_{min} > 1) \text{ then } \rightarrow \text{relevant} \quad (5)$$

This assumption filters out all peaks, allowing for permanent charging with average charging power without the need for load shifting. f_{min} also depends on the billing period as described in subsection 2.3. The rating is therefore also carried out on a yearly and a monthly basis.

²For the t_f used in this paper, the best case are 24 Δt with 100% $e_{l,flex}$ and the peak Δt with 100% $e_{l,fix}$ leading to $f = 24/25 = 0.96$.

2.5 Calculation of charging power

Subsection 2.4 discusses the filtering for relevant peaks. The peaks that required some sort of load shifting were identified as the relevant peaks. However, as the other relevant peaks all show a higher flexibility than f_{min} , they exhibit a certain surplus of flexibility and therefore a further degree of freedom. Figure 4 shows two different ways to use the surplus flexibility: for charging power optimization or for charging time optimization. Due to costs, it is of interest to minimize the charging power needed to exploit the maximum flexibility. Consequently, the minimum required charging power is calculated as a further indicator to characterize the GCP.

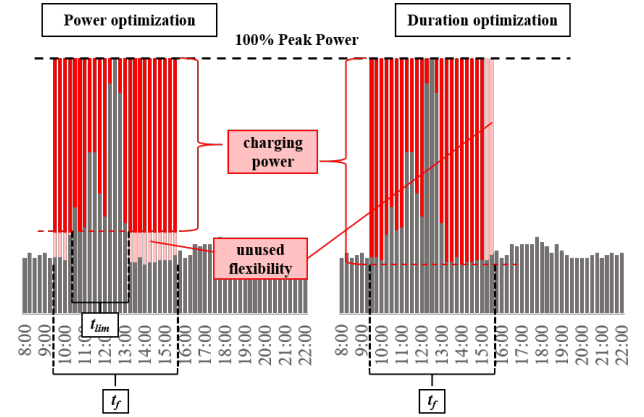


Figure 4: Example for utilization of flexibility with charging limit: Utilization of flexibility for power optimization (left) and for charging duration optimization (right)

The following paragraph describes the algorithm to calculate the minimal required charging power (see Algorithm 1). The reference is the energy that can be charged at the most inflexible peak ($e_{min} = e_{l,flex}(f_{min})$). At least this amount should be chargeable at all other peaks as well. Therefore, during the calculation, the charging power p_{cha} is gradually increased, until the integral over all time steps ($t \in t_f$) matches the energy chargeable at the most inflexible peak. For timestamps, where a charging power limitation is needed ($t \in t_{lim}$), the maximum power reserve until the peak power is used for charging. We use the normalized (on the year peak) baseline load profile $p_n[t]$. The charging power during t_{lim} is then given as $(1 - p_n[t])$. For the remaining timestamps ($t \in t_{unlim}$), charging with the maximum charging power p_{cha} is assumed. The algorithm tries to find a data point with the next higher charging power $p_{cha,tmp}$ until the required energy e_{min} is met. Since Δt and t_f are fixed values for all peaks, and $\Delta e = p \cdot \Delta t$, the term Δt can be set to one and the energy can be determined by $e = \sum \Delta e = \sum p$ as a sum of the powers of the individual time intervals. The charged energy e_{cha} consists of a part during limited e_{lim} and a part during unlimited e_{unlim} charging. The charging power during the unlimited time per definition needs to be higher or the same as the power during the limited times. Therefore, if the energy requirement is fulfilled, the maximum charging power can be calculated as the energy charged during unlimited charging ($e_{unlim} = e_{min} - e_{lim}$) divided by the number of time steps of unlimited charging ($\text{len}(t_{unlim})$).

Algorithm 1 Identifying maximum required charging power

```

 $e_{min} \leftarrow \text{len}(t_f) \cdot \min(f_y)$ 
while  $e_{cha} < e_{min}$  do
   $p_{unlim} \leftarrow p_n$  where  $p_n < (1 - p_{cha})$ 
   $p_{cha,tmp} \leftarrow (1 - \max(p_{unlim}))$ 
  for  $t \in t_f$  do
    if  $t \in t_{lim}$  then
       $e_{lim} \leftarrow e_{lim} + (1 - p_n[t])$ 
    else if  $t \in t_{unlim}$  then
       $e_{unlim} \leftarrow e_{unlim} + (1 - p_{cha,tmp})$ 
    end if
  end for
   $e_{cha,tmp} \leftarrow e_{lim} + e_{unlim}$ 
  if  $e_{cha,tmp} < e_{min}$  then
     $p_{cha} \leftarrow p_{cha,tmp}$ 
     $e_{cha} \leftarrow e_{cha,tmp}$ 
  else if  $e_{cha,tmp} \geq e_{min}$  then
     $p_{cha} = (e_{min} - e_{lim}) / \text{len}(t_{unlim})$ 
    break
  end if
end while

```

3 REAL-WORLD DISTRICT AND DATA BASE

In section 3, we present the real-world district in the city of Karlsruhe, Germany, and present the data base.

3.1 Real-world district

The district is of a mixed residential-commercial type. Regarding charging profiles, charging at work is present as well as the charging of car-sharing vehicles and charging by residents. Since it is an urban district, the potential for further renewable energy generation is limited. The district consists of four GCPs with a multitude of buildings subordinate to each GCP (see Figure 5; compare Figure 1). The GCPs are coupled through the public 20 kV grid. Regarding the construction year, the district is quite heterogeneous. It reaches from buildings constructed less than 5 years ago to ones constructed more than 120 years ago.



Figure 5: Buildings of the four GCPs in the real-world district in Karlsruhe, Germany [29]

As mentioned in section 1, the GCP is a relevant billing boundary. Energy supplier within the GCPs are not regulated as utility companies. Furthermore, within the GCPs in this specific district, local generation and consumption is further subsidized³. The advantages of reduced regulation and subsidies are omitted, when utilizing the public grid for energy exchange. Therefore, the advantages of considering the whole district versus the single GCPs need to be critically examined in subsection 4.4.

3.2 Data base

The real-world load profiles used in the present paper were measured by the utility company for billing purposes. The temporal resolution (compare subsection 2.1) of the load profiles is equidistant 15 min and the data is available for the whole year of 2021. All load profiles were converted in UTC with the start timestamps of each period and the power consumption for each GCP. For plots of the load profiles, see Appendix A.

As mentioned in subsection 2.1 the value resolution Δp significantly influences the number of identified local peaks p_l . The p_l values were obtained following the method described in subsection 2.2 with an order of 48. The Δp of the real-world load profiles ranges from 0.001 % to 2.128 % given in percent of the peak power p_{qy} as shown in Table 1. GCPs with a lower value resolution (higher Δp) present a significantly higher number of p_l (compared Table 1). For example, for GCP 2, 1372 p_l were identified instead of the expected average of 365 per year (resp. 366 in 2021 due to leap year). The number of additionally identified p_l can be characterized by the occurrence of adjacent maxima or several maxima in one calendar day. Of GCP 2's total of 1372 p_l , 670 have directly adjacent p_l with the same value. In 1013 cases of the 1372 p_l , more than one p_l per calendar day was identified due to an indistinguishable peak level. However, since they do represent real world power maxima and it is not possible to extract a more precise peak value, all data points at the peak level were treated as equally relevant in this work. In future work, other methods like filtering according to flexibility indicator or peak slope could be examined.

Table 1: Lower value resolution Δp leads to a lower number of distinguishable peak power levels and multiple values, either directly neighboring or within the same day exhibiting an identical peak level and therefore leads to more peak values (number of identified p_l) than the ideal 365-366 per year

	GCP 1	GCP 2	GCP 3	GCP 4
$(\Delta p / p_{qy})$	0.001 %	2.128 %	0.049 %	0.901 %
number of identified local peaks p_l	375	1372	430	501
number of peaks with immediate neighbor at same power level	22	670	74	110
number of additional peaks per calendar day	22	1013	117	153

³'Mietstromzuschlag' according to EEG 2021 §21 3.2 and reduced grid fees according to StromNEV §18 1.1.

4 RESULTS AND DISCUSSION

In section 4, we discuss the results obtained from applying the methodology presented in section 2 to the data base presented in section 3.

4.1 Application of data-analysis workflow

Since the data-analysis workflow has already been described in detail in section 2, only the specifics of the results are discussed in the following. Besides the quality of the data base, which was described in subsection 3.2, the results of the filtering process and the determined flexibility indicators are the main factors to be examined.

Peaks are in their nature outliers and extreme values. In average, only 1.93 % of all peaks of the real-world validation dataset are relevant for flexibility characterization on a yearly basis. Significantly more data points, 29 % of all peaks, characterize the monthly flexibility (compare Table 2). Future work also needs to examine the sensitivity of the percentage of relevant peaks toward t_f (that was chosen as a 6 h time frame in this work, see subsection 2.3). The other results of the analysis workflow are presented in the

Table 2: In contrast to p_l , the absolute number of relevant peaks on a yearly basis (f_y) or a monthly basis (f_m) is not significantly influenced by Δp for the data basis used in this work. The additional peaks are irrelevant for the characterization of the GCP (see subsection 2.4)

	GCP 1	GCP 2	GCP 3	GCP 4
number of identified p_l	375	1372	430	501
relevant on yearly basis	2	15	9	20
relevant on monthly basis	112	128	221	129
irrelevant	263	1244	209	372

following directly in the form of a comparison of the four GCPs.

4.2 Comparing the data-analysis results for the four GCPs

The monthly and yearly flexibility show the same characteristics of the GCPs for the given dataset. Meaning, the higher the flexibility on a yearly basis, the higher it is also on a monthly basis (see Table 3 and Figure 6).

Table 3: Characteristic flexibility f_{min} and corresponding normalized power p_n

	GCP 1	GCP 2	GCP 3	GCP 4
$f_{min,y}$	0.14	0.07	0.14	0.13
$p_{n,y}$ at $f_{min,y}$	1.000	0.957	1.000	0.964
mean $f_{min,m}$	0.44	0.08	0.23	0.16
mean $p_{n,m}$ at $f_{min,m}$	0.959	0.998	0.893	0.953

GCP 1 and GCP 3 exhibit the highest and GCP 2 the lowest flexibility. This indicates that information on the flexibility of different

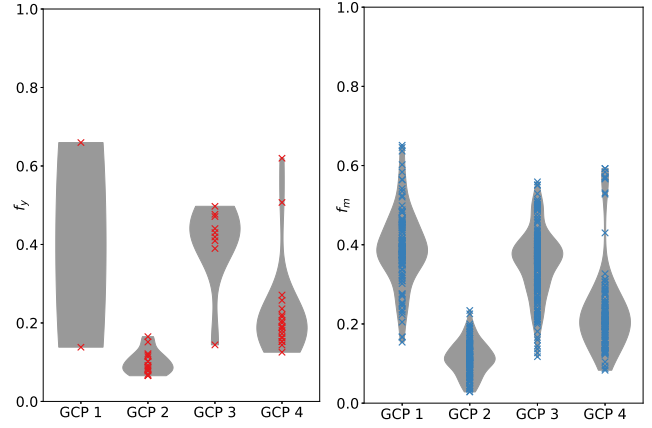


Figure 6: Comparison of f_y and f_m show similar tendencies regarding the minimum exhibited flexibility and the flexibility distribution for all GCPs

GCPs could be extracted even of less than a year of data. Future research should further examine this relation and take an in-depth look at monthly billing systems and conduct a more detailed analysis of single months.

On a yearly basis, one single peak data point may exhibit a significantly lower flexibility than all other relevant peaks. In the presented data, this was the case for GCP 1 and GCP 3 (see Figure 6). In those cases, allowing underfulfillment for a single peak could significantly increase the flexibility. The CI could be planned according to the next higher flexibility by only under fulfilling a single 6 h time frame. The presented approach explicitly does not have the aim to describe operating strategies or control algorithms and should only support the planning of new CI. However, including analysis of the sensitivity towards underfulfillment could lead to further understanding and better planning in future works.

As shown in Table 4 and Figure 7, to achieve the flexibility displayed in Table 3 at some time steps, the charging power needs to be higher than the mean value to compensate the missing energy of reduced charging periods around the peaks. At all four GCPs the maximum charging power is reached at the most inflexible peak. In average, more than the double (2.2 times) of the mean charging power is needed to achieve the before mentioned flexibility. While in average a power of 11.83 % of the peak consumption in CI can be installed, some GCPs (GCP1 and GCP 3) allow significantly higher charging power than the others.

Table 4: Charging power mean and max values

values as % of p_{gy}	GCP 1	GCP 2	GCP 3	GCP 4
worst case max p_c	95.6 %	40.4 %	60.5 %	80.2 %
optimal max p_c	37.3 %	8.5 %	33.3 %	28.8 %
mean p_c	13.8 %	6.6 %	14.4 %	12.5 %
CI oversizing reg. mean	2.7	1.3	2.3	2.3

Figure 7 shows the required charging power in relation to the available flexibility. The displayed 'worst case' means the maximum charging power in t_f is used, meaning charging at a total power consumption of p_{gy} for any time, charging is active and only stopping, if the energy requirement is met or the power needs to be reduced due to a peak. In the 'optimal case', the charging power is at most the highest charging power from the most inflexible peak.

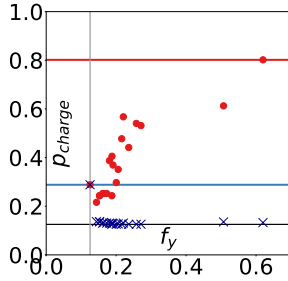


Figure 7: The most inflexible peak determines the maximum charging power, as shown here for GCP 4 (for all GCPs see Appendix C). In the unoptimized case (red o), higher charging powers occur at all peaks exhibiting greater flexibility. The excess flexibility can be used to reduce the charging power (blue x) down to the mean charging power (black line).

The charging power can in average be chosen more than 60% lower when exploiting the surplus flexibility than in the unoptimized case. Still, the installed charging power needs to be about 2.2 times higher than the average charging power to take advantage of the flexibility potential. The mean proposed charging power lies at 11.83% of the peak power for the sum of the GCPs (see Table 4). The oversizing of power is characteristic of all plants with flexibility. Otherwise, flexibility provision would not be possible at all. A big advantage of CI is that oversizing is possible at low cost, since in the case discussed it is not necessary to reinforce the supply line to the GCP or the transformer, since new peaks are prevented. Only the plug, and therefore the charging point and the car, have to be able to handle correspondingly higher power.

Since the charging power can be reduced significantly below the charging power from the most inflexible peak, not even all flexibility is used for charging power optimization. This still leaves a margin for safety or other purposes.

4.3 Correlation of flexibility and peak power

Among other things, it was noted in subsection 4.2 that the peak with the highest power peak is not necessarily identical with the most inflexible peak (for example see Figure 8. This finding is examined in more detail in the following.

As shown in Table 5 and Figure 9, peak power and f are strongly negatively correlated. This correlation is statistically significant (p-value lower than 5%) for the unfiltered values on a yearly and monthly basis and the filtered values on a monthly basis, see Table 5). However, this is not true for the yearly basis due to some short but high peaks. Because of their shortness, those high peaks also exhibit a high flexibility indicator. And because of the reduced number of relevant data points for the yearly basis, those outliers

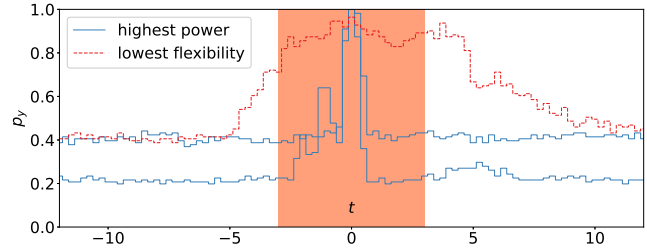


Figure 8: For GCP 4 the peak with the highest power is not the one with the lowest flexibility (all GCPs, see Appendix C)

have a relevant impact. Therefore, a key finding is, that the highest peak cannot be assumed to also be the most inflexible one.

Table 5: On a monthly basis, the peak height and flexibility show a statistically significant negative correlation for all GCPs. On a yearly basis, GCP 1 and GCP 2 show no statistically significant correlation, GCP 3 shows a statistically significant negative and GCP 4 a statistically significant positive correlation

		GCP 1	GCP 2	GCP 3	GCP 4
f_y (filtered)	PCC	-1.00	0.15	-0.92	0.64
	p-value	1	0.59	0.00038	0.0026
f_m (filtered)	PCC	-0.55	-0.82	-0.55	-0.82
	p-value	3.5e-10	6.6e-32	3.5e-19	1.6e-32

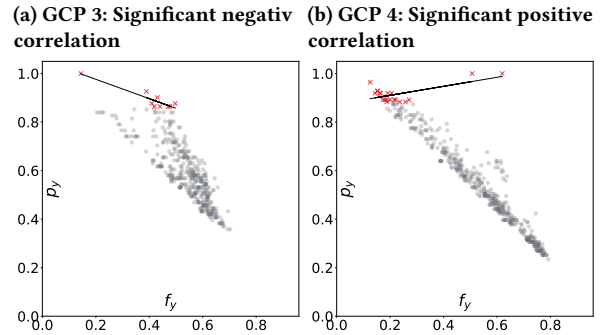


Figure 9: Two examples of the correlation of p_y and f_y of filtered values (red x) that can be significantly different for different GCPs (for a complete overview, see Appendix B)

Table 3 shows the peaks with the lowest flexibility. On a yearly basis, those occur at peaks with a proportion of 95.7% to 100% of the yearly peak value. In six out of eight of the combinations of GCPs and billing methods, the lowest flexibility does not correlate with the highest peak. On a yearly basis, the lowest flexibility occurs in average at a peak with around 98% of the yearly peak power. This imposes the need to include suitable characterization methods for flexibility for all use cases or data-analysis methods. For example, a

control algorithm for operating a charging infrastructure should not only be based on a peak forecast but also, at least implicitly, include a flexibility forecast. Meaning, the most critical day for an optimization may not be those with the highest predicted peak but those with a slightly lower but for longer time constantly high baseline profile.

4.4 Single GCP versus district

As mentioned in section 1, the advantages of combining multiple GCPs to a district need to be critically examined. For this purpose, the flexibility indicator was calculated for three different scenarios (A, B and C). The parameters of interest are the total peak power and the flexibility indicator values.

4.4.1 Scenario A. In the state-of-the-art scenario, billing is carried out for each GCP separately. Therefore, the least flexible peak of each individual GCP is of interest. The relevant baseline load profile in scenario A is the sum of the individual load profiles around these most inflexible peaks. The peak height is defined by the sum of the peak height of all GCPs around these peaks.

4.4.2 Scenario B. In the district scenario, all GCP's baseline load profiles are summed up before applying the analysis workflow. The whole district is thus considered as one large single consumer. Therefore, the flexibility of the district is calculated on the newly formed maxima of the newly calculated load profile. This would require the law to allow billing on a district level. The advantage of district billing in contrast to GCP based billing is the reduction in required peak power. In the real-world data at hand, the district peak is 23.48 % lower than the peak in scenario A (additional plot, see Figure 14). The flexibility in scenario B (0.089), however, is also reduced by almost 25 % compared to scenario A (0.152). The reduction in peak power also led to a reduction in flexibility. Further evaluation is needed to examine, how much the power can be reduced through a district while still keeping the same flexibility as scenario A. It should be examined what the relationship of peak reduction and flexibility provision on a district level looks like.

4.4.3 Scenario C. To bring the reduced peak power and the reduced flexibility in perspective, as a third scenario C, the district is also evaluated with the maximum power of the sum of the GCPs. In the scenario C, the flexibility (0.245) increases by almost 107 % compared to scenario A (0.152). This means, if the same power maxima than in scenario A is assumed, the district outperforms the state-of-the-art. However, if the full peak power reduction potential of 23.48 % should be utilized, the flexibility is also reduced. There appears to be a tradeoff of peak reduction versus flexibility.

Further, since flexibility is always defined relative to the peak, for a fair comparison, also the different absolute peak height should be considered. Therefore, also the comparison of the mean charging power regarding scenario A is included in Figure 10. In the district scenario B, the mean charging power is reduced by more than 40 %. In scenario C the mean charging power is 3.61 times higher than in scenario A. Moreover, the mean available charging power is reduced alongside the peak power. This likewise shows the presumed tradeoff of peak reduction versus flexibility. Data for

additional districts is needed to further examine the relationship of peak reduction and flexibility provision on a district level.

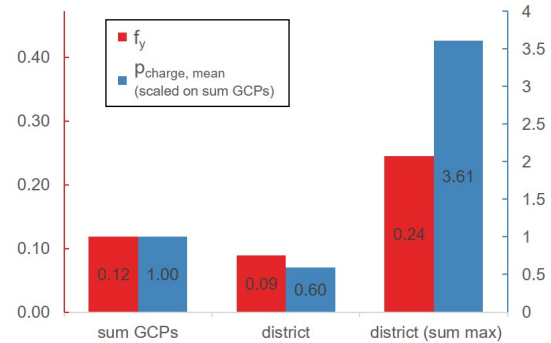


Figure 10: Comparison of characteristic flexibility f_y and mean charging power $p_{charge,mean}$ scaled to the current states of separate GCPs

5 CONCLUSION AND OUTLOOK

The present paper shows that flexibility is not a natural inherent and fixed value of an energy system, but rather describes the suitability of an energy system in performing optimal under specific use cases or boundary conditions.

We present a new flexibility indicator that considers characteristics of charging infrastructure, like limited availability of vehicles and limited charging power, which supports the planning phase of new charging infrastructure. Furthermore, we present the whole data-analysis workflow on how to apply the flexibility indicator. We apply the workflow on a data base from a real-world district and examine the monthly and yearly flexibility and compared the different grid connection points of the district.

The data-analysis in the present paper confirms that peak characteristics are highly influencing the suitability for valley filling. We further demonstrate that it is incorrect to assume that the highest peak during a particular time period is also the most inflexible one. The monthly and yearly flexibility show the same characteristics for the given data base. Furthermore, we show that a surplus in flexibility can be utilized to reduce and optimize the maximum required charging power.

Future work should examine the feasibility of extrapolating flexibility on less than one year of data and validate the results on more than one year of data as well as further real-world datasets. Furthermore, different approaches on dealing with lower measurement resolution could be looked into. Moreover, the identified tradeoff of flexibility versus peak reduction on the district level needs to be further investigated, as well as the potential of flexibility increase when allowing underfulfillment of charging demand in a limited time frame. Finally, since we pointed out that the most inflexible peak is not always the one with the highest peak power, this imposes the need to at least implicitly include flexibility in future works on forecasts and optimizations, and it is not sufficient to only use the highest peak power values as key indicators.

ACKNOWLEDGMENTS

This work was supported by the Helmholtz Association under the program 'Energy System Design' and by the 'Smart East'⁴ project (funding reference L75 21113 by Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg). The authors would like to thank their colleagues from the utility companies Stadtwerke Karlsruhe GmbH and BES - Badische Energie-Servicegesellschaft mbH for supplying the data as well as the colleagues from FZI Research Center for Information Technology, Seven2one, the Energy Lab 2.0 and the Institute for Automation and Applied Informatics (IAI) for the fruitful discussions. Furthermore, we would like to thank Technologiefabrik Karlsruhe GmbH and Hoepfner Bräu Friedrich Hoepfner Verwaltungsgesellschaft mbH & Co KG for their participation in the project.

AUTHOR CONTRIBUTIONS

Author contributions according to Contributor Roles Taxonomy (CRediT, <https://www.casrai.org/credit.html>): *Conceptualization*: J.G.; *Methodology*: J.G.; *Software*: J.G.; *Formal analysis*: J.G.; *Investigation*: J.G.; *Data curation*: J.G.; *Writing - original draft*: J.G.; *Writing - review and editing*: J.G., S.W., V.H.; *Visualization*: J.G.; *Supervision*: S.W., V.H.; *Project administration*: S.W.; *Funding acquisition*: V.H.

REFERENCES

- [1] Syed Abdullah-Al-Nahid, Tafsir Ahmed Khan, Md. Abu Taseen, and Tareq Aziz. 2020. A Consumer-Friendly Electric Vehicle Charging Scheme for Residential Consumers. In *2020 International Conference on Smart Grids and Energy Systems (SGES)*, 893–897. <https://doi.org/10.1109/SGES51519.2020.00164>
- [2] Bahram Alinia, Mohammad H. Hajiesmaili, Zachary J. Lee, Noel Crespi, and Enrique Mallada. 2022. Online EV Scheduling Algorithms for Adaptive Charging Networks with Global Peak Constraints. *IEEE Transactions on Sustainable Computing* 7, 3 (July 2022), 537–548. <https://doi.org/10.1109/TSUSC.2020.2979854> Conference Name: IEEE Transactions on Sustainable Computing.
- [3] Onur Ayan and Belgin Turkey. 2018. Domestic electrical load management in smart grids and classification of residential loads. In *2018 5th International Conference on Electrical and Electronic Engineering (ICEEE)*, 279–283. <https://doi.org/10.1109/ICEEE2.2018.8391346>
- [4] Lukas Barth, Veit Hagenmeyer, Nicole Ludwig, and Dorothea Wagner. 2018. How much demand side flexibility do we need?: Analyzing where to exploit flexibility in industrial processes. In *Proceedings of the Ninth International Conference on Future Energy Systems*. ACM, Karlsruhe Germany, 43–62. <https://doi.org/10.1145/3208903.3208909>
- [5] Lukas Barth, Nicole Ludwig, Esther Mengelkamp, and Philipp Staudt. 2018. A comprehensive modelling framework for demand side flexibility in smart grids. *Computer Science - Research and Development* 33 (Feb. 2018), 1–11. <https://doi.org/10.1007/s00450-017-0343-x>
- [6] Fatih Birol. 2021. Key World Energy Statistics 2021. (2021), 81.
- [7] Bundesnetzagentur. 2022. Bundesnetzagentur - Elektromobilität: Öffentliche Ladeinfrastruktur. <https://www.bundesnetzagentur.de/DE/Fachthemen/ElekttrizitaetundGas/E-Mobilitaet/start.html>
- [8] Bundesregierung der Bundesrepublik Deutschland. 2020. Energiewirtschaftsgesetz - EnWG. https://www.gesetze-im-internet.de/enwg_2005/index.html
- [9] Bundesregierung der Bundesrepublik Deutschland. 2022. StromNEV. <https://www.gesetze-im-internet.de/stromnev/index.html#BJNR222500005BJNE000603118>
- [10] Georgios Chantzis, Panagiota Antoniadou, Maria Symeonidou, Elli Kyriaki, Efrosyni Giana, Symeon Oxyzidis, Dionysia Kolokotsa, and Agis M. Papadopoulos. 2022. Evaluation of methods for determining energy flexibility of buildings. *Green Energy and Sustainability* 2, 3 (July 2022). <https://doi.org/10.47248/ges2202030006> Publisher: Pivot Science Publication Corp..
- [11] Niangjun Chen, Chee Wei Tan, and Tony Q. S. Quek. 2014. Electric Vehicle charging in Smart Grid: Optimality and Valley-Filling Algorithms. *IEEE Journal of Selected Topics in Signal Processing* 8, 6 (Dec. 2014), 1073–1083. <https://doi.org/10.1109/JSTSP.2014.2334275> Conference Name: IEEE Journal of Selected Topics in Signal Processing.
- [12] Christian Buchmüller and Maximilian Hemmert-Halswick. 2021. Forschungsberichte zum Energiesystem X.0 Nr. 1: Intelligente und effiziente Vernetzung von Energieerzeugern und -verbrauchern auf Quartiersebene. (2021), 221.
- [13] Csaba Csiszár. 2019. Demand Calculation Method for Electric Vehicle Charging Station Locating and Deployment. *Periodica Polytechnica Civil Engineering* (Jan. 2019). <https://doi.org/10.3311/PPci.13330>
- [14] Dominik Danner, Jan Seidemann, Michael Lechl, and Hermann de Meer. 2021. Flexibility Disaggregation under Forecast Conditions. In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems (e-Energy '21)*. Association for Computing Machinery, New York, NY, USA, 27–38. <https://doi.org/10.1145/3447555.3464851>
- [15] Claude Ziad El-Bayeh, Ursula Eicker, Khaled Alzaareer, Brahim Brahim, and Mohamed Zellaoui. 2020. A Novel Data-Energy Management Algorithm for Smart Transformers to Optimize the Total Load Demand in Smart Homes. *Energies* 13, 18 (Sept. 2020), 4984. <https://doi.org/10.3390/en13184984>
- [16] Kevin Förderer, Veit Hagenmeyer, and Hartmut Schmeck. 2021. Automated generation of models for demand side flexibility using machine learning: an overview. *ACM SIGEnergy Energy Informatics Review* 1, 1 (Nov. 2021), 107–120. <https://doi.org/10.1145/3508467.3508477>
- [17] Zechun Hu, Kaiqiao Zhan, Hongcai Zhang, and Yonghua Song. 2016. Pricing mechanisms design for guiding electric vehicle charging to fill load valley. *Applied Energy* 178 (Sept. 2016), 155–163. <https://doi.org/10.1016/j.apenergy.2016.06.025>
- [18] Rune Grønberg Junker, Armin Ghaseem Azar, Rui Amaral Lopes, Karen Byskov Lindberg, Glenn Reynnders, Rishi Relan, and Henrik Madsen. 2018. Characterizing the energy flexibility of buildings and districts. *Applied Energy* 225 (Sept. 2018), 175–182. <https://doi.org/10.1016/j.apenergy.2018.05.037>
- [19] Han Li, Zhe Wang, Tianzhen Hong, and Mary Ann Piette. 2021. Energy flexibility of residential buildings: A systematic review of characterization and quantification methods and applications. *Advances in Applied Energy* 3 (Aug. 2021), 100054. <https://doi.org/10.1016/j.adapen.2021.100054>
- [20] Zhengyi Luo, Jinqing Peng, Jingyu Cao, Rongxin Yin, Bin Zou, Yutong Tan, and Jinyue Yan. 2022. Demand Flexibility of Residential Buildings: Definitions, Flexible Loads, and Quantification Methods. *Engineering* (March 2022). <https://doi.org/10.1016/j.eng.2022.01.010>
- [21] Christoph Molitor. 2015. *Residential city districts as flexibility resource: analysis, simulation, and decentralized coordination algorithms* (1 ed.). E.ON Energy Research Center, RWTH Aachen University, Aachen. OCLC: 1073862535.
- [22] Amr A. Munshi and Yasser Abdel-Rady I. Mohamed. 2018. Extracting and Defining Flexibility of Residential Electrical Vehicle Charging Loads. *IEEE Transactions on Industrial Informatics* 14, 2 (Feb. 2018), 448–461. <https://doi.org/10.1109/TII.2017.2724559> Conference Name: IEEE Transactions on Industrial Informatics.
- [23] Bijay Neupane, Torben Bach Pedersen, and Bo Thiessen. 2015. Evaluating the Value of Flexibility in Energy Regulation Markets. In *Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems (e-Energy '15)*. Association for Computing Machinery, New York, NY, USA, 131–140. <https://doi.org/10.1145/2768510.2768540>
- [24] Frauke Oldewurtel, David Sturzenegger, Goran Andersson, Manfred Morari, and Roy S. Smith. 2013. Towards a standardized building assessment for demand response. In *52nd IEEE Conference on Decision and Control*. IEEE, Firenze, 7083–7088. <https://doi.org/10.1109/CDC.2013.6761012>
- [25] Presse- und Informationsamt der Bundesregierung. 2019. *Klimaschutzprogramm 2030 der Bundesregierung zur Umsetzung des Klimaschutzplans 2050*. Technical Report. Presse- und Informationsamt der Bundesregierung, Berlin. 173 pages.
- [26] Seyedfarzad Sarfarazi, Marc Deissenroth-Uhrig, and Valentin Bertsch. 2020. Aggregation of Households in Community Energy Systems: An Analysis from Actors' and Market Perspectives. *Energies* 13, 19 (Oct. 2020), 5154. <https://doi.org/10.3390/en13195154> Place: Switzerland Publisher: MDPI AG.
- [27] Jonas Schlund, Marco Pruckner, and Reinhard German. 2020. FlexAbility - Modeling and Maximizing the Bidirectional Flexibility Availability of Unidirectional Charging of Large Pools of Electric Vehicles. In *Proceedings of the Eleventh ACM International Conference on Future Energy Systems (e-Energy '20)*. Association for Computing Machinery, New York, NY, USA, 121–132. <https://doi.org/10.1145/3396851.3397697>
- [28] Johannes Schäuble, Thomas Kaschub, Axel Ensslen, Patrick Jochem, and Wolf Fichtner. 2017. Generating electric vehicle load profiles from empirical data of three EV fleets in Southwest Germany. *Journal of Cleaner Production* 150 (2017), 253–266. <https://doi.org/10.1016/j.jclepro.2017.02.150> Place: Netherlands Publisher: Elsevier BV.
- [29] Smart East Karlsruhe. 2021. Smart East - Das Quartier aus der Vogelperspektive. https://www.youtube.com/watch?v=hjhpDf_-9Y
- [30] Keenan Valentine, William G. Temple, and K. Max Zhang. 2011. Intelligent electric vehicle charging: Rethinking the valley-fill. *Journal of Power Sources* 196, 24 (Dec. 2011), 10717–10726. <https://doi.org/10.1016/j.jpowsour.2011.08.076>
- [31] Ziming Yan, Tianyang Zhao, Yan Xu, Leong Hai Koh, Jonathan Go, and Wee Lin Liaw. 2020. Data-driven robust planning of electric vehicle charging infrastructure for urban residential car parks. *IET Generation, Transmission & Distribution* 14, 26 (2020), 6545–6554. <https://doi.org/10.1049/iet-gtd.2020.0835> _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1049/iet-gtd.2020.0835>

⁴<https://smart-east-ka.de/>

- [32] Kaiqiao Zhan, Zechun Hu, Yonghua Song, Ning Lu, Zhiwei Xu, and Long Jia. 2015. A probability transition matrix based decentralized electric vehicle charging method for load valley filling. *Electric Power Systems Research* 125 (Aug. 2015), 1–7. <https://doi.org/10.1016/j.epsr.2015.03.013>
- [33] Li Zhang, Faryar Jabbari, Tim Brown, and Scott Samuelson. 2014. Coordinating plug-in electric vehicle charging with electric grid: Valley filling and target load following. *Journal of Power Sources* 267 (Dec. 2014), 584–597. <https://doi.org/10.1016/j.jpowsour.2014.04.078>
- [34] Yanchong Zheng, Ziyun Shao, Yitong Shang, and Linni Jian. 2019. Distributed Charging Control of Smart Electric Vehicle Chargers for Load Valley Filling in Distribution Networks. In *2019 IEEE Vehicle Power and Propulsion Conference (VPPC)*. 1–6. <https://doi.org/10.1109/VPPC46532.2019.8952459> ISSN: 1938-8756.
- [35] Hüseyin K. Çakmak and Veit Hagenmeyer. 2022. Using Open Data for Modeling and Simulation of the All Electrical Society in eASiMOV. In *2022 Open Source Modelling and Simulation of Energy Systems (OSMSES)*. 1–6. <https://doi.org/10.1109/OSMSES54027.2022.9769145>
- [36] Laurynas Šikšnys, Torben Bach Pedersen, Muhammad Aftab, and Bijay Neupane. 2019. Flexibility Modeling, Management, and Trading in Bottom-up Cellular Energy Systems. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems (e-Energy '19)*. Association for Computing Machinery, New York, NY, USA, 170–180. <https://doi.org/10.1145/3307772.3328296>

A REAL-WORLD BASELINE LOAD PROFILES

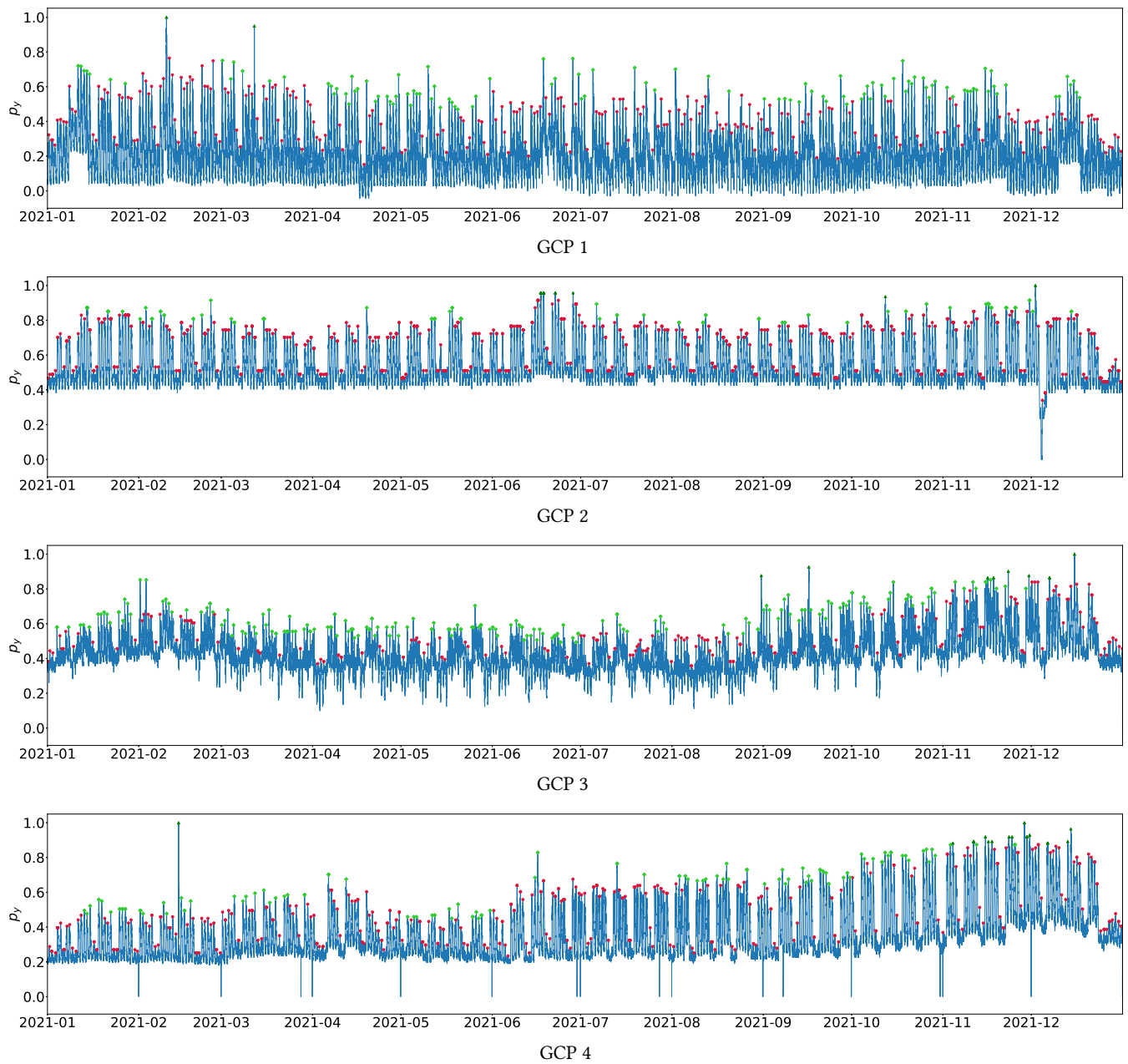


Figure 11: Baseline load profiles of all GCPs

B CORRELATION OF FLEXIBILITY (f) AND PEAK POWER (p)

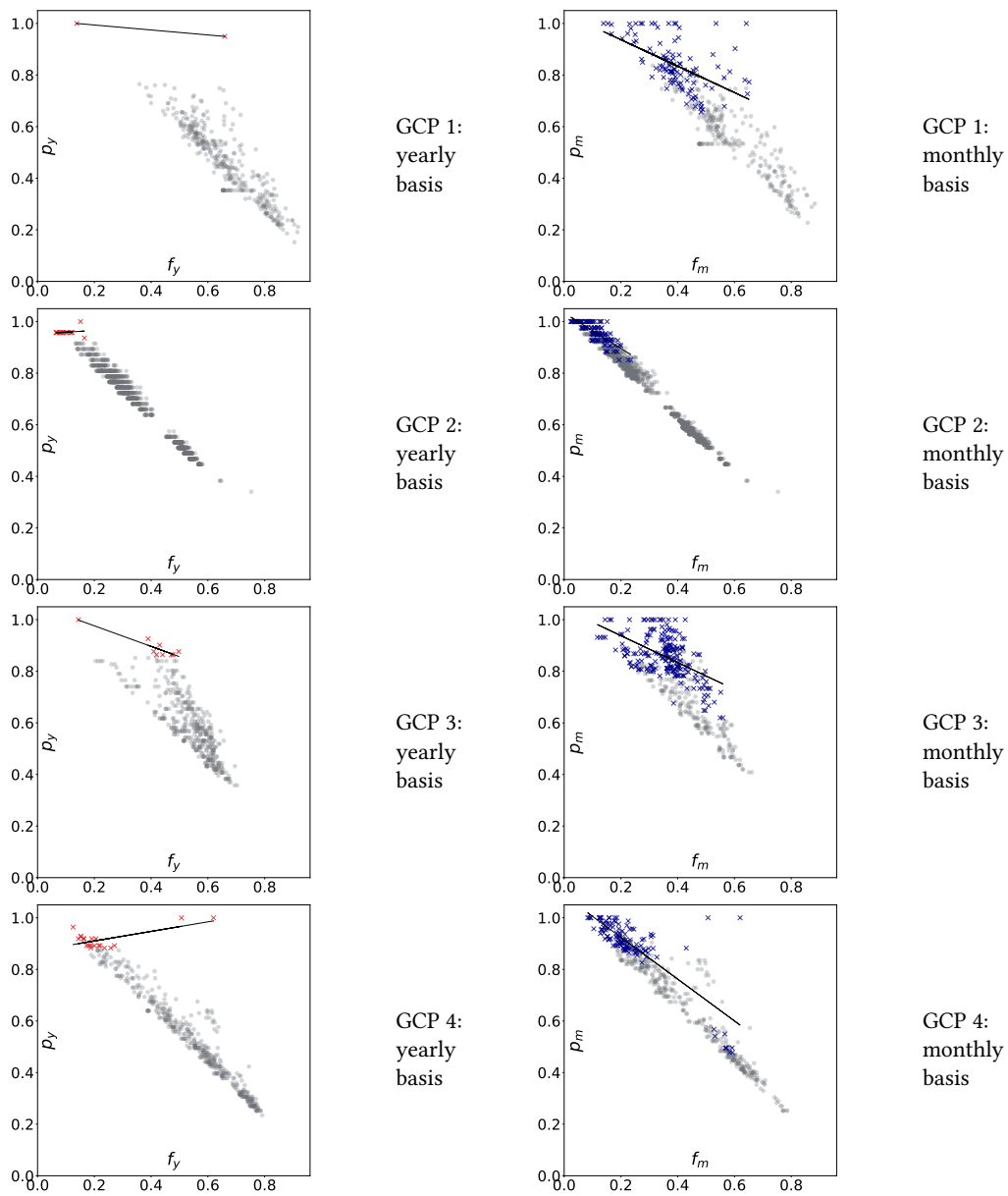


Figure 12: Correlation of flexibility (f) and peak power (p) of all GCPs on a yearly and a monthly basis black lines marking linear regression of correlation of filtered values, x marking relevant and o marking irrelevant peaks

C MOST INFLEXIBLE PEAKS AND CHARGING POWER

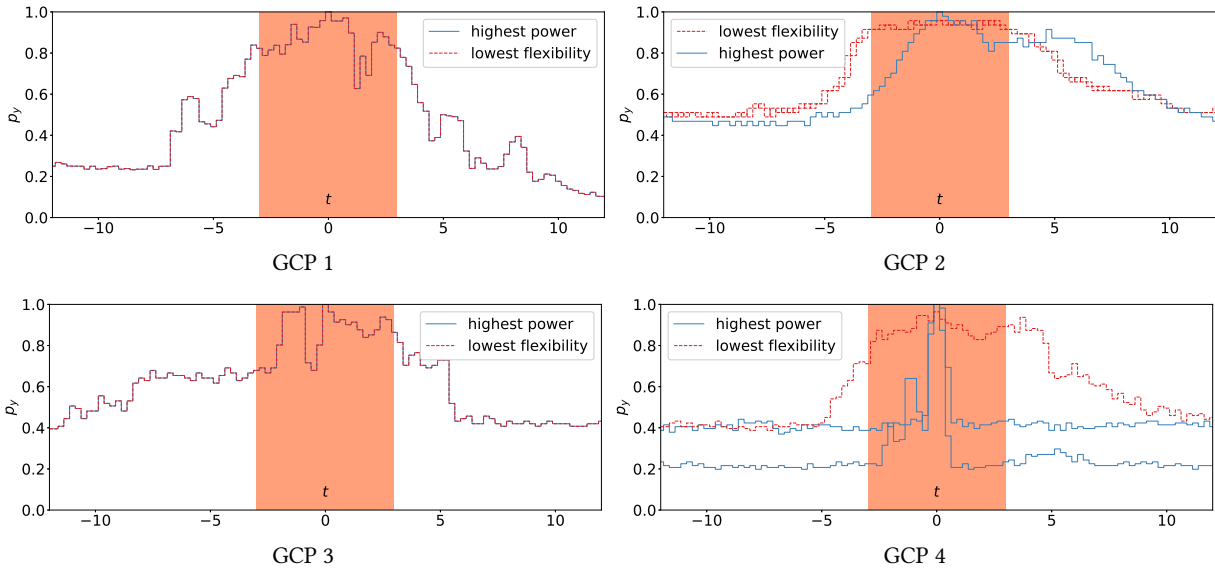


Figure 13: For GCP1 and GCP3 the highest year peak corresponds with the most inflexible peak. Regarding GCP2 and GCP4, the year peak is significantly more flexible, than other peaks

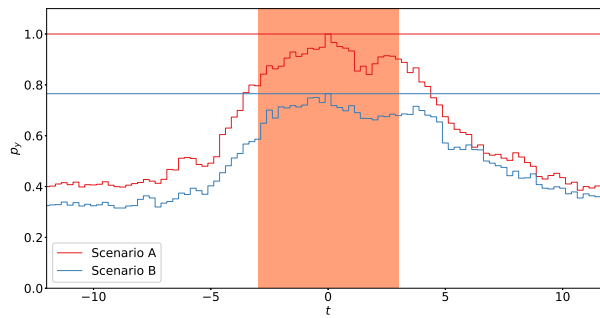


Figure 14: Sum of the most inflexible peaks of all GCPs (sce. A) versus the most inflexible peak of the district profile (sce. B)

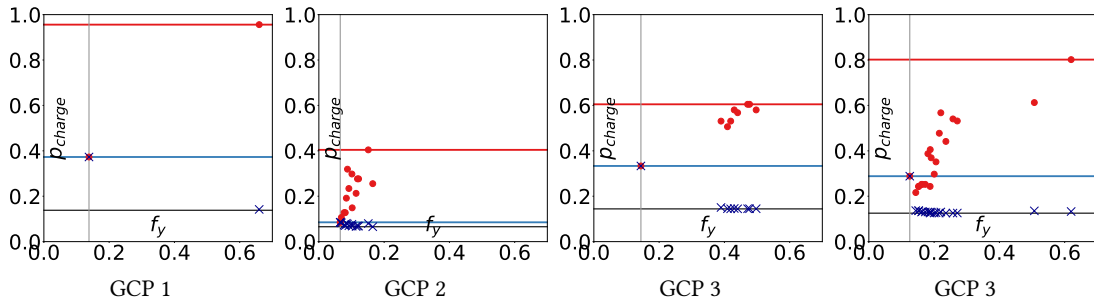


Figure 15: The most inflexible peak determines the maximum charging power. In the unoptimized case (red o), higher charging powers occur at all peaks exhibiting greater flexibility. The excess flexibility can be used to reduce the charging power (blue x) down to the mean charging power (black line).