

# Improving accuracy of energy forecasting through the presence of an electric vehicle fleet

Dejan Ilić<sup>a,\*</sup>, Stamatis Karnouskos<sup>a</sup>, Michael Beigl<sup>b</sup>

<sup>a</sup>*SAP, Karlsruhe, Germany*

<sup>b</sup>*Karlsruhe Institute of Technology (KIT), Germany*

---

## Abstract

Accurate forecasting of energy is of pivotal importance for the Smart Grid vision. Today it is achieved through aggregation of numerous individual energy loads, and hence stochastic behaviour costs are distributed to the entire aggregated stakeholders. However, if one was able to accurately control the internal energy behaviour in order to meet an accurate self-forecast (as seen from external stakeholder's point of view), new business opportunities in the Smart Grid era could be considered. Static storage systems may be used to absorb forecast errors, however these are still costly, and therefore alternatives are sought. Stakeholders, however, can benefit from a plethora of alternative "storages" offered by properly utilizing their assets. We approach the absorption of forecast errors by utilizing a fleet of electric vehicles whose on-premise presence is used to compose a variable energy storage. An empirical assessment with real-world data is provided and results demonstrate the significance of electric vehicles towards helping stakeholders achieve and maintain the accuracy of their self-forecast.

*Keywords:* Smart grids, Demand forecasting, Electric vehicles, Electricity supply industry, Load management

---

## 1. Introduction

Increased adoption of Distributed Energy Resources (DER) and especially volatile Renewable Energy Sources (RES) in electricity grids raised its complexity and their management becomes increasingly difficult [1]. Such trends coupled with the emergence of prosumers (producers and/or consumers of energy), as well as the electrification of transportation [2], increase the unpredictability in the grid and require costly solutions that raise the electricity costs. However, if a prosumer could achieve determinism in his energy signature, via highly accurate load forecast and potentially control over the deviations from that forecast,

---

\*Corresponding author

*Email addresses:* `dejan.ilic@ymail.com` (Dejan Ilić), `stamatis.karnouskos@sap.com` (Stamatis Karnouskos), `Michael.Beigl@kit.edu` (Michael Beigl)

he could act as a reliable resource on the grid [3]. The latter may lower the need for costly mitigation plans that take care of grid reliability.

Emerging business models and roles in Smart Grids call for active participation of the traditionally passive consumers [4]. Such opportunities include for instance active involvement in grid operations [5], participation in local energy markets [6], or demand response programs [7] etc. To participate in such programs it is of key importance to fulfil the prerequisites which for the traditionally passive consumers mean to feature an accurate self-forecast on their energy loads [8]. Although not all stakeholders can achieve it [9], additional benefits are expected for those who can report it to third parties [3] e.g. such that one can verify and measure activity in demand-response programs [7]. The concept of the self-forecasting stakeholders (or groups of them as one stakeholder [10]) is analysed in detail in [11] and is hereby referred to as Self-Forecast EneRgy load Stakeholder (SFERS).

A desired accuracy of the self-forecast is achieved locally and the deterministic energy signature is gained by reporting it to the external stakeholders [3]. To meet the report, forecast errors produced within the self-forecast are absorbed by locally available assets. In past, deployment of storage solutions in highly volatile systems was successfully used to absorb uncertainty in power grids [12]. Similar applications were found in absorbing forecast errors of a smaller neighbourhood [13]. These traditional battery storage systems proved to be efficient but rather expensive solutions [14], thus efforts towards reducing their costs are investigated [15]. Capitalizing on the locally available assets, may lead to a significant cost reduction, thus economically enable business of the stakeholders to profit in the Smart Grid era by acting as SFERS.

This work investigates how a commercial stakeholder can become SFERS and utilise its newly available resources e.g. batteries of electric vehicles (EV) [16] which can act as a variable energy storage. We identify the origin of forecasting errors [17] and assess the suitability of its existing EV fleet for achieving our goals. The empirical investigation is based on real-world data obtained from a commercial building and the EVs attached to its charging stations. We show how the presence of its employees provokes higher forecasting errors, while EVs that are present at that exact point in time can assist in mitigating this error. As such, results show that an EV fleet can be used for self-forecasting, potentially even eliminate a need for a static solution. However, it is also discovered that in intervals of low fleet presence the availability of a static storage unit may be more beneficial than increasing the fleet size.

To validate our concepts we follow these steps: First the intraday forecast accuracy is analysed in [section 2](#). Subsequently in [section 3](#) the storage capabilities are analysed with focus on the EV fleet. In [section 4](#) we provide some insights on the variable storage capacity. Finally in [section 5](#) we assess the actual storage requirements of the stakeholder to achieve different levels of accuracy. All these steps are further complemented with definitions as well as empirical data to make it possible to fully comprehend them. In this manuscript, storage is always referred to as energy capacity (in kWh) and not as power capacity (in kW).

## 2. Intraday Forecast Accuracy

Wide availability of the smart metering data enabled new energy related opportunities even for individual consumers [11]. Their energy consumption data is usually collected at a constant sampling period  $T$ , therefore represented as a discrete-time signal  $y[n] \geq 0$  where  $n$  is an integer. As an example, Figure 1 depicts energy load data of a commercial stakeholder sampled by its smart meter at  $T = 15$  minutes. Although only 5 weeks are shown, this is a representative pattern as the variations in consumption repeat continuously through the entire year. The difference in consumption over days, led us to split the set into working and nonworking days (including holidays). Their load difference, in particularly the intraday behaviour, is expected to affect the errors in the forecasted intervals [17].

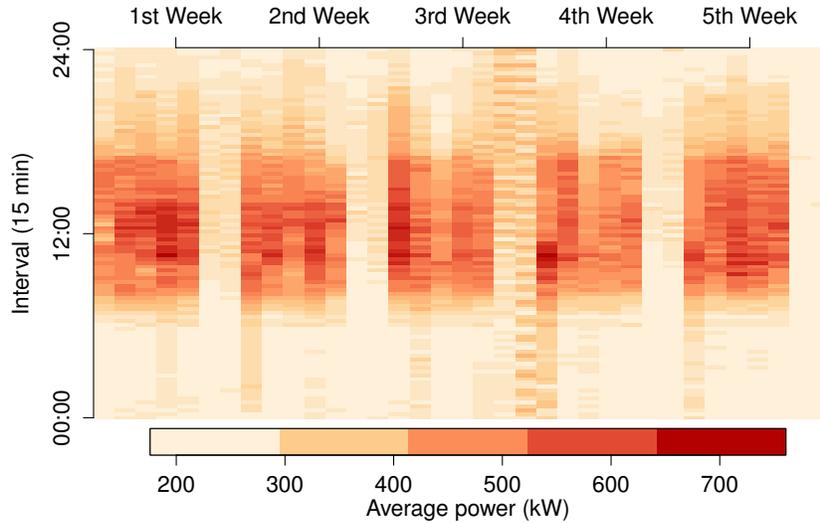


Figure 1: Heatmap of energy consumption of a commercial stakeholder

### 2.1. Quantitative Observation of Forecast Errors

In this work, the consumption self-forecast is done only for a short-term horizon, for one day. Many forecast methods could be applied to the time series data produced by a smart meter [18] and other indicators can further improve the forecast [19]. If an interval forecast is  $\hat{y}[n] \geq 0$ , the total forecast error of that interval is  $w[n] = \hat{y}[n] - y[n]$ . Since that error will be accommodated within an available storage capacity, both the Mean Absolute Percentage Error (MAPE) and quantitative error (kWh) will be observed. If  $X$  is a set of index intervals of interest, e.g. first interval of every work day in 2011, one can measure its average interval error for any

$$W[n]_X = \begin{cases} \frac{1}{|X|} \sum_{k \in X} w[k - n], & 0 \leq n < \frac{\Delta}{T} \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where  $|X|$  is cardinality,  $\Delta$  is the season length and  $k$  represents the element of the set. Since the load varies on intraday basis, it is expected to have variations in the required energy storage capacity to address the self-forecast errors [17].

### 2.2. Intraday Errors

Figure 1 depicts an example of the load produced by this commercial building with offices with 139 working places and its resulting consumption in 2011 was 2440 MWh. As it can be seen, the building is mainly used in between 08:00–17:00 and there is a significant difference in energy load for different days of the week. The average daily power (over the entire year) approximates to 342 kW and 210 kW for working and nonworking days respectively. As such, the interval set  $X$  is divided into sets of first indices for all working and nonworking days,  $X_w$  and  $X_n$  respectively.

For the self-forecast the Seasonal AutoRegressive Integrated Moving Average (SARIMA) model was selected, as it can be used to relatively accurately predict electricity demand [20]. A forecast for next day is done on weekly seasonality and the model is trained with 4 seasons (28 days). The model training is made only with the samples known from 4 seasons up to a forecasted day and the model parameters are extracted from the same set using the "auto.arima" method offered by the ARIMA libraries in the forecast package of R ([www.r-project.org](http://www.r-project.org)). Observations are made for average daily 15-minute intervals of  $\Delta = 1$  day, and Figure 2 depicts the resulting functions of  $W[n]_{X_w}$  and  $W[n]_{X_n}$ .

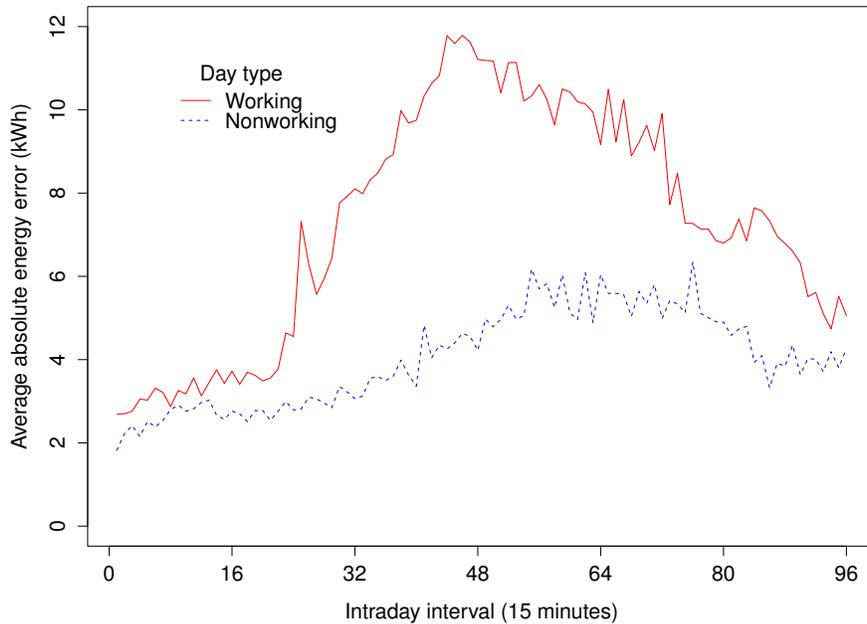


Figure 2: Absolute forecast error averaged over intraday intervals to identify significance of their forecast error contribution

On a daily average, workdays resulted to  $\text{MAPE}_{X_w} = 8.55\%$  and nonworking days to  $\text{MAPE}_{X_n} = 7.56\%$ . Due to the actual consumption difference for  $X_w$  and  $X_n$ , these forecast errors result in average daily error of 702 kWh and 388 kWh respectively. As Figure 2 indicates, quantitatively (and not by percentage) the errors differ significantly for the same intraday intervals. For lower values of  $n$ , working  $W[n]_{X_w}$  and nonworking days  $W[n]_{X_n}$  have a comparable forecast error, however the error of workdays around midday increases significantly. Although the real cause for the error is hard to pinpoint, it appears to be highly correlated with the working hours. Hence, one may expect that also other commercial stakeholders may experience a similar correlation to errors of their self-forecasts.

### 3. Presence of Storage Units

With the electrification of transportation networks [2], we are witnessing an increase in the penetration of highly mobile electricity storage units, e.g. electric vehicles [16]. However, EVs are underutilized as they are idle 96% of their time [21]. For EVs this may imply that the majority of that time they are connected to the grid and can be integrated into a variable energy storage [3]. This holds especially true for company cars parked in garages of the buildings employees work in (as we have in this case). In this section a method to measure the presence of mobile storage units of an entire fleet is proposed.

#### 3.1. Unit Presence Definition

Every mobile storage unit is able to connect to the electricity grid at some point in time and this connection time frame is called the grid session. Each grid session  $s$  is instantiated by connecting a unit to the grid at time  $t_c$  and is terminated by its disconnection at time  $t_d$ . Sessions of each individual unit can only occur sequentially, where for one session the storage unit is considered to be present for any time  $t$  as  $t_c \leq t < t_d$ . The step function [22] is used to model a single grid session of a storage unit. It is an elementary function denoted by  $u(t)$ , which holds one for positive side and zero for negative. A single grid session  $s$  is represented by two step functions as

$$p^s(t) = u(t - t_c^s) - u(t - t_d^s). \quad (2)$$

As such, the function returns one only if a unit is present on the grid, otherwise zero is returned. Numerous such sessions are actually the components for composition of the unit presence function  $p(t)$ . This function will return the total count of units present at time  $t$ . This is mathematically represented as

$$p(t) = \sum_{s \in S} p^s(t), \quad (3)$$

where  $S$  is the set of sessions from all mobile units considered. As such, the function returns  $\mathbb{N}^0$ , where zero indicates that none of the units is present.

### 3.2. Statistical Presence

In contrast to the unit presence function, one may be interested in understanding how many units are expected to be connected at a selected point in time. Furthermore, without knowing the number of present units of an entire fleet, the presence rate cannot be calculated. Thus, function  $v(t) \in \mathbb{N}^0$  indicates the count of individual units in ownership over time. The presence function is represented by

$$f(t) = \frac{p(t)}{\min(v(t), q(t))}, \quad (4)$$

where  $q(t)$  is reflecting the limited number of charging points on premise. As such, the function is used for a statistical assessment of fleet's behaviour allowing scaling of their presence. If the set  $X$  contains time points of interest (e.g. 00:00 of all working days in year 2012), and  $\Delta$  indicates the season length, then the statistical presence for all points in  $X$  is calculated as

$$F(t)_X = \begin{cases} \frac{1}{|X|} \sum_{\tau \in X} f(\tau - t), & 0 \leq t < \Delta \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

where  $|X|$  is cardinality and the return value is  $\mathbb{R}_0^+$ . Once calculated, the statistical model can be used for any fleet size to estimate the expected presence at a point in time in the form of

$$\bar{f}(t) = \sum_{\forall i} \sum_{\tau \in X_i} F(t - \tau)_{X_i}, \quad (6)$$

where  $X_i \cap X_j = \emptyset$  where all  $i \neq j$ . It is important to note that the model is not prone to errors introduced by inappropriate selection of points in each  $X_i$ . In fact, better selection of these points (e.g. only working days, without holidays) will result to a more accurate statistical model of the fleet's presence.

### 3.3. Presence of a Real-World EV Fleet

After quantitatively identifying forecast errors of the stakeholder, it would be interesting to see if fleet's presence can assist at the times of the highest errors. In this section, the presence curves are produced from 1044 grid sessions  $s \in S$  of a real-world EV fleet. The data is collected from 5 January 2012 to 10 August 2013 (585 days), where 18 working days were marked as holidays (thus nonworking days). The fleet was continuously composed of five "Mercedes-Benz A-Class E-Cell" vehicles ( $v(t) = 5$ ) which were part of the 500 specially manufactured cars built as part of the collaboration between Mercedes-Benz and Tesla Motors. These cars are pure EVs, and the fleet is in full ownership of the same stakeholder presented in [subsection 2.2](#). As EVs were not exclusively assigned to individual employees, different mobility patterns may be expected [\[23\]](#).

Many different variations in session duration were noted (15 minutes  $\leq t_d^s - t_c^s \leq 4$  days), the first observation is done through their duration. The

### 3.3 Presence of a Real-World EV Fleet PRESENCE OF STORAGE UNITS

distribution function of unit sessions from the set  $S$  is depicted in Figure 3. A peak of short sessions can be immediately noticed, however these sessions have no significant impact on the unit presence curve  $p(t)$ . Further investigation revealed that sessions initiated on Friday have much greater impact on the fleet's presence. As Figure 3 indicates, the mean duration from the complete set  $S$  averaged around 10 hours, while for Fridays (depicted from 179 sessions) resulted in more than 13 hours. As it can be observed, small peaks around 72 hours have significant impact as vehicles are present over an entire weekend.

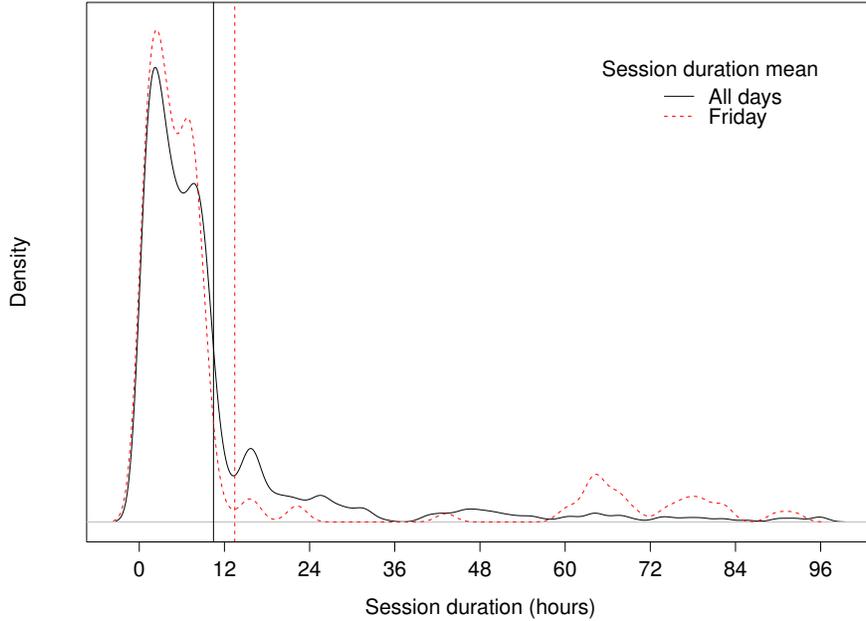


Figure 3: Distribution function of the session duration for the complete set  $S$  and set of session initiated on Friday

A second observation is made for the duration of a grid session  $t_d - t_c$  over its connection time  $t_c$ . Such investigation will help understand intraday behaviour of the units, having  $S' \subseteq S$ , where duration of all sessions is limited to 1 day. In Figure 4 one can see the movement of EVs for all sessions  $s \in S'$ , where most vehicles are connected within the stakeholder's working time (noted in section 2). The trend of availability at the end of working time can be noticed from the drop of hours on the grid if moving along the  $t_c$  axis. All of the connections above this drop are considered as storage units being available over midnight, what appears to be more often for nonworking days.

Points in set  $X$  can be set to many different variations, however for experiments with a commercial stakeholder the weekly points should be used due to the significant difference over weekdays. For a better understanding of fleet's behaviour, the defined sets  $S$  and  $S'$  can be both observed through their statistical presence function  $F(t)_X$ . As previously depicted in Figure 3, such limitations

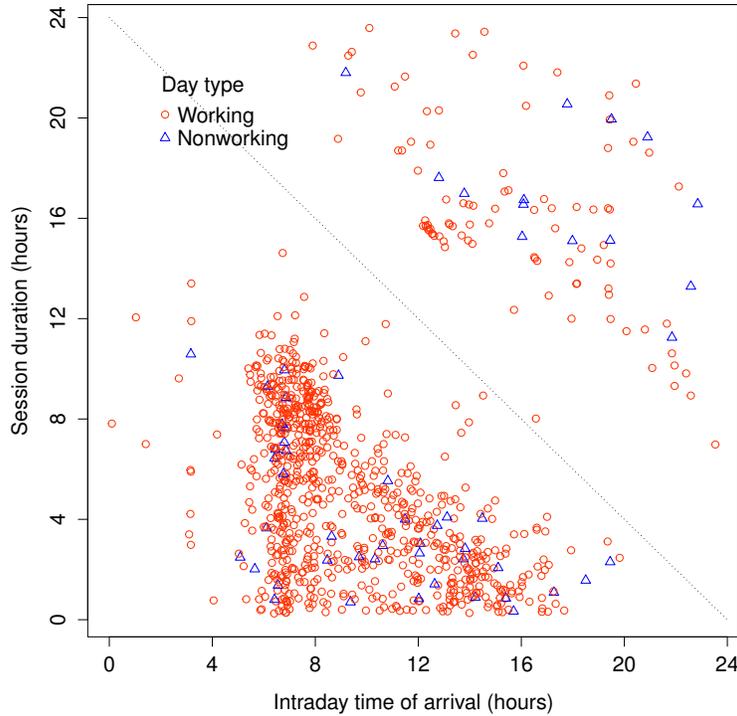


Figure 4: Duration of the grid connection session in respect to time of their intraday initialization

are expected to be significant in the overall presence of the fleet. Figure 5 depicts the resulting statistical presence for both the complete set  $S$  and the reduced one  $S'$  for  $\Delta$  of 1 week. A significant improvement of the presence for the complete set is observed.

Interestingly all workdays of the week look alike (in average 18.9% for  $S$  and 11.3% for  $S'$ ), while significant drop is noticeable over weekend days (in average 9.1% for  $S$  and 1.8% for  $S'$ ). Such a small difference allows the distinction of  $X$  to working days  $X_w$  and nonworking days  $X_n$  (including holidays). Their statistical presence appears to have good fit to the forecast errors (of the stakeholder described in subsection 2.2) and will be scaled to achieve a required self-forecasting accuracy of the stakeholder.

#### 4. Variable Storage Capacity

The definition of the presence curves in the previous section are further used to address the specific stakeholder's forecasting errors. As an example, the analysis of the stakeholder in subsection 2.2 and its EV fleet presence from subsection 3.3 appears to be a good fit. Once the presence curves are computed, they are used to calculate the resulting energy capacity (in kWh) composed of

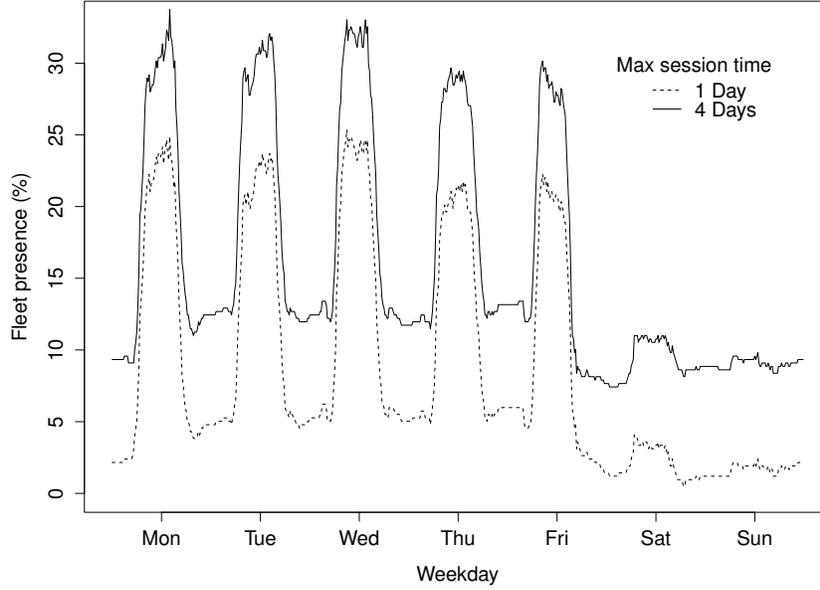


Figure 5: A real-world EV fleet statistical presence at weekly  $\Delta$  for both complete  $S$  and reduced set  $S'$

storage units. These units will be observed as the dynamic part of the variable storage. If the capacity of a single unit is denoted as  $c$ , simple multiplication as

$$p_c(t) = cv(t)f(t) = cp(t), \quad (7)$$

will give the total capacity available over time. From the experiment described in subsection 3.3, we have in Figure 6 the capacity availability from the fleet that is calculated for  $c = 36$  kWh. The statistical capacity present over  $X_w$  and  $X_n$  resulted with an average capacity of 36.9 kWh and 16.6 kWh respectively for the complete session set  $S$ . As such,  $p_c(t)$  can be used in the assessment simulation for improvement through an existent EV fleet. For further scaling of its fleet, statistical presence curves  $F(t)_X$  can be used only for one classification of storage characteristics  $c$ . Since many fleets are expected to have different  $c$ , the equations need to be further expanded.

#### 4.1. Mobile Units of Different Classification

Although the calculation of the presence curves can be done through capacity, it is not applicable to fleets with units of different capacities. For example, if only two vehicles of capacity  $c$  and  $10c$  are available, the presence of smaller unit may jeopardize the estimation of the capacity available. With that in mind, every session  $s$  is expanded with the classification  $j$  of invariable capacity  $c_j$ . All sessions  $s_j$  are therefore populating the set of  $S_j \subseteq S$ . The classified statistical presence is expanded from Equation 5 as  $F_j(t)_X$ , where only  $s \in S_j$  are

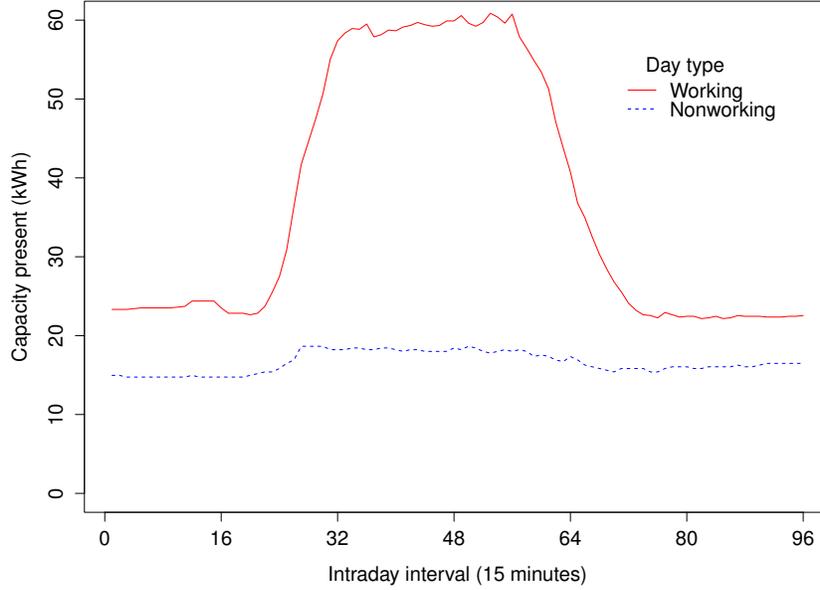


Figure 6: Storage capacity present from the fleet's statistical presence

considered. Although  $c_j$  is considered to be invariable, the total count of classified mobile units is to be scaled in simulations and is represented as  $v_j \in \mathbb{N}^0$ . Scaling  $v$  of each classification will contribute to the total capacity present and is mathematically described as

$$\bar{p}_c(t) = \sum_{\forall j} c_j v_j \bar{f}_j(t). \quad (8)$$

The total capacity available is expected to grow by an increasing number of units within the stakeholder's fleet, so one can assess their impact on the achievement of a higher forecast accuracy. As such, individual variation of  $v_j$  can be used for assessment of the SFERS simulations with scenarios utilizing different vehicle classifications.

#### 4.2. Constantly Present Storage

For highly variable fleets, a continuously present storage can be critical towards reaching a required forecast accuracy [17]. For instance, if a fleet suffers from low presence intervals (as observed in Figure 6), forecast accuracy may converge to  $\text{MAPE} > 0$ . For this particular example, even a significant fleet scaling might still result to insufficient capacity to cover the forecast errors in the intervals of low presence. Hence the model should adopt a static storage capacity in parallel with the variable one to compensate for such gaps. The total capacity present can be noted as

$$\bar{P}_c(t) = C + \bar{p}_c(t), \quad (9)$$

### 4.3 Property of Assessment on Actual Storage Requirement

where  $C$  is the static part in the variable storage and  $\bar{p}_c(t)$  its dynamic part. Scaling  $C$  and unit count  $v_j$  is therefore used for assessment of impact in absorbing the forecast error.

#### 4.3. Property of a Variable Storage

In the context of the statistical presence curves, the overall storage capacity is calculated from many different mobile units e.g. EVs, that connect and disconnect at any point in time. This introduces increased complexity of unit management, in particular towards estimating the connection and disconnection State of Charge (SOC) of an individual asset [24]. As this complexity management goes beyond the scope of this work, it is important to mention that SOC is not treated individually. If  $t_1 < t_2$ , the overall SOC is expressed as

$$SOC(t_2) = \frac{SOC(t_1)\bar{P}_c(t_1) + a(\bar{P}_c(t_2) - \bar{P}_c(t_1))}{\bar{P}_c(t_2)}, \quad (10)$$

where state at  $t_2$  is inherited by its previous one (at  $t_1$ ) and forecast error is added thereafter. As the SOC per unit is not available, the variable  $a$  is introduced for the adjusting the SOC of expanded/reduced storage units. In this work, variable  $a$  will result to

$$a = \begin{cases} 50\%, & \bar{P}_c(t_2) \geq \bar{P}_c(t_1) \\ SOC(t_1), & \text{otherwise.} \end{cases}$$

Management of connected vehicles can be based on numerous factors [21] e.g. selecting a unit to be charged. This is similar to the power plant management, where "dispatch" refers to the timing turning on and off power plants to match grid's needs. Although considered interesting, evaluation on an individual SOC of units is left for future work.

## 5. Assessment on Actual Storage Requirement

In this section, the stakeholder analysed in subsection 2.2 is evaluated with respect to the presence of the fleet as it is analysed in subsection 3.3 that is further scaled by applying the methods from section 4. Since the energy data of the stakeholder is a discrete-time signal of  $T = 15$  minutes, the presence curves are sampled at the same frequency. This is important as a forecast error  $w[n]$  is quantitatively absorbed by an estimated capacity  $\bar{P}_c[n]$  for the same interval  $n$ . With help of the discrete-time signals, the assessment of SFERS via  $\bar{P}_c$  is done by variation of the mobile units in its ownership  $v_j$  (or their capacity  $c_j$  for any classification  $j$ ) and the static storage capacity  $C$ .

### 5.1. Individual Capacity Scaling

As the forecast algorithm utilized in section 2 resulted to greater errors within the working hours, the capacity available from the dynamic storage in subsection 3.3 is expected to be highly relevant towards absorbing them [17].

The efficiency of the dynamic storage shape  $P_c$  cannot be directly compared to  $C$  on the efficiency to absorb the errors. As such, the dynamic storage shape available is to be averaged and scaled by a constant over all working  $X_w$  and nonworking  $X_n$  days in  $X$ . The average capacity presence resulted respectively in 20.5% and 9.2% for the overall shape efficiency, having an average (on weekly basis) of 17%. As described in [17], for their direct comparison equal shape areas need to be considered as:

$$P'_c = m \int_a^b \bar{P}_c(t) dt \equiv \int_a^b C dt = C(b - a), \quad (11)$$

where  $m$  is the scaling factor used to align the areas of the dynamic and static shapes. For the follow-up experiments,  $m = 100\%/17\% = 5.88$  is used to calculate the efficient dynamic storage. This overall shape efficiency is used to depict the approach comparison in Figure 7, by individually scaling of  $v_j$  and  $C$  in  $\bar{P}_c$ .

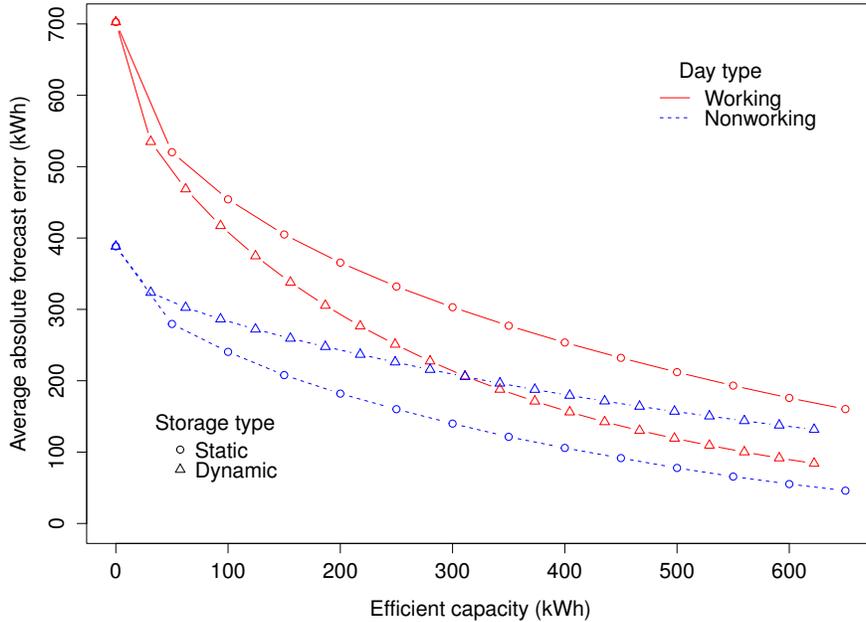


Figure 7: Efficiency of static and dynamic storage shapes to absorb the forecast errors

As initially suspected, the fleet presence successfully overlaps with the working hours in  $X_w$  and over-performs when compared to the static storage. However, the error reduction on nonworking days is significantly higher with the static storage approach, due to the capacity availability when low presence of dynamic storage is observed [13]. Equation 11 is not to be omitted, as average shape efficiency of the dynamic storage resulted only in 17% for the complete set  $S$ . Combining these two approaches may reduce costs [16], and their individual advantages are expected to complement each other.

### 5.2. Interdependent Storage Relationship

Individual scaling in subsection 5.1 already emphasizes weaknesses and strengths of the two approaches. Their combination is expected to fill-in the performance gaps of the other approach. The results from Figure 7 strongly indicate that the two approaches are advantageous either for  $X_w$  or  $X_n$ . Applying Equation 9, the storage scaling is performed for both  $C$  and  $a_j$  together, thus affecting the total capacity available  $P'_c$ . Figure 8 depicts how their scaling reduces the absolute forecast error for working  $X_w$  and nonworking  $X_n$  days, as well as their difference.

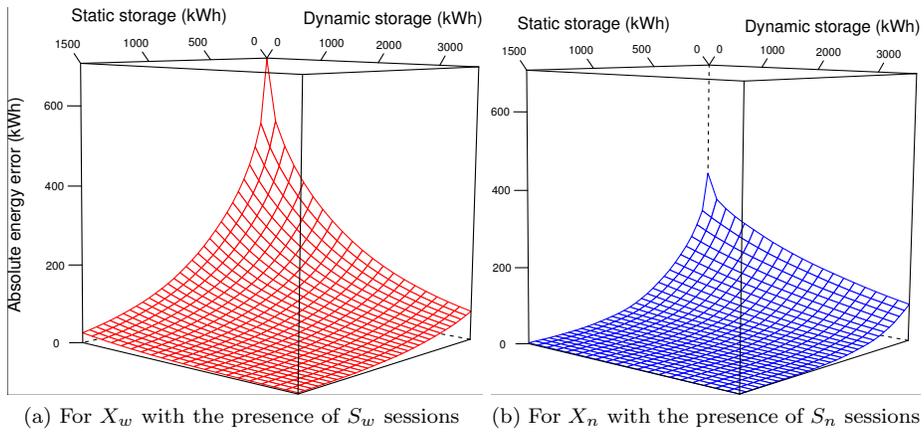


Figure 8: The forecast error reduction on daily average with the interdependent storage approaches

The values in Figure 8 are selected so that one can observe the advantage of the two datasets, having  $0 \leq C \leq 1500$  and  $a_j = 100$  (or 72% of employees). Aligned axis values of 8a and 8b are important for a quantitative observation in forecast error reduction from their daily consumption (of 8.2 MWh and 5 MWh respectively). The capacity selected from both approaches is insufficient for  $X_w$ , while the static storage was embolden for  $X_n$ . As expected, both figures show high convergence towards the average forecast error of zero once approaches are combined. The difference between forecast errors from  $X_w$  and  $X_n$  is interesting, as one can see how the error reduction of  $X_w$  prevails over the improvement rate of  $X_n$ , whereby for the low values of  $C$  the error of  $X_w$  overcomes  $X_n$  on greater values of  $a_j$ .

It is important to point out Equation 11, where  $C$  of the constantly present storage takes 100% of capacity availability while dynamic capacity efficiency is measured to 20.5% for  $X_w$  and 9.2% for  $X_n$ . Understanding the benefits of the presented approaches to address the uncertainty of self-forecasting can help obtaining a most economical settings to achieve an adequacy of SFERS becoming a resource [25]. Omitting the potential of stakeholder's EV fleet would

increase the requirement of a constantly present storage, which is expected to rise the overall system costs [26].

5.3. Achieving Forecast Accuracy Levels

Envisioning a longer-term role of SFERS with their variable storage as compensation of their forecast errors, and opportunity (dependent on the actual SOC) to accommodate an intermittent energy resource, can lead us to rethink the roles in the energy systems of today. A question to a stakeholder becoming a SFERS is where to draw a line between the two approaches presented, particularly on economical terms [26]. Perhaps their performance dependency needs to be evaluated for the levels of forecast accuracy (e.g. 2% or 3%) relevant to the most economical position for the stakeholder. In dependence to their goal, stakeholders can achieve sufficient accuracy for daily, weekday or even intraday requirements [27]. Hereby only continuous accuracy is observed, where the forecast error levels will be always below a certain MAPE limit. Figure 9 depicts a few example results from the data of the experiment in subsection 5.2.

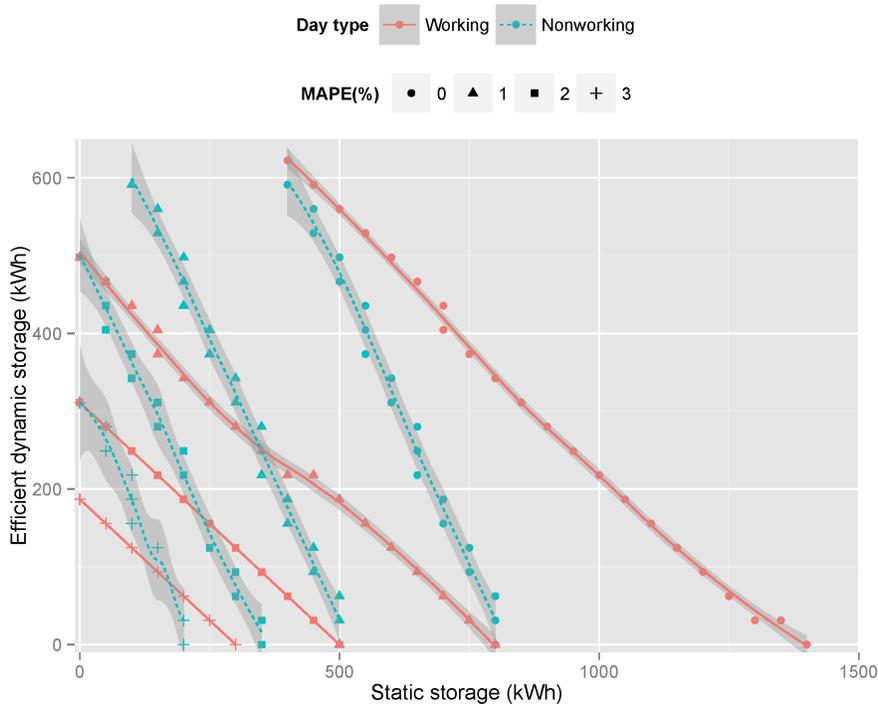


Figure 9: Example of forecast accuracy levels of the commercial stakeholder

The limits are minimal for each depicted accuracy level of MAPE, thus on the right side of the limits a lower MAPE is expected (for both  $X_w$  and  $X_n$ ). Furthermore, losses due to the storage efficiency  $\xi$  can be omitted if the observation is made through the entire consumption of a stakeholder as

$\frac{(1-\xi)}{2}|w[n]| + y[n]$ . For example, if forecasted load is  $\hat{y}[n] = 150$  kWh and the actual energy consumed is  $y[n] = 160$  kWh (having MAPE of 6%), with an  $\xi = 90\%$  the actual energy consumed deviates only 0.3% (or 160.5 kWh) from the stakeholder's original consumption. With this in mind, the storage efficiency has almost no impact to costs of a stakeholder; hence stakeholders can apply the proposed methodology to evaluate their dependency on static and dynamic storage in their target to become an economically sustainable SFERS.

## 6. Discussion

There are several limitations present, including the off-the-shelf forecasting algorithms used in this work, which can significantly improve results presented here through further optimization and inclusion of additional customization [19]. Another limitation is the usage of the statistical curve where an approximation of the significance of the fleet to absorb forecast error is considered. In this way, one cannot observe the actual SOC of a battery, which led to the respective discussion in [subsection 4.3](#). Although this is an interesting aspect, it is not within the scope of this work. As an extension of this work, future simulations of individual units need to be considered, as well as technical limitations (e.g. charging power in kW) of the storage units.

If a company utilizes an EV fleet as a storage solution, the system has to make sure that the individual driver requirements are met e.g. each car is adequately charged for its next trip. As an example, if a desired SOC for SFERS with a static storage solution would be at 50%, the clustered available storage from EVs would have to also be at that level. However this does not mean necessarily 50% SOC for all individual EVs as this might conflict with the owner's goals which are e.g. to be at least 80% in order to cover his travel plans. Furthermore, each unit can provide a certain percentage of the battery capacity to the variable storage and still can guarantee that the EV is ready whenever the user needs it. Understandably, more research is needed towards local and global constraint consideration, however such strategies are not considered part of this assessment here and are seen as future work.

## 7. Conclusion

Significant research in the Smart Grid is devoted towards improving the stability of electricity grid and their quality of service. By becoming SFERS, the stakeholders can slip into new roles as these are envisioned by the Smart Grid and benefit business-wise. This work demonstrates an empirical approach towards achieving higher accuracy of self-forecast by the usage of the variable energy storage e.g. made available via the presence of EVs. The results show that the forecast errors (in kWh) of the commercial stakeholder grow within the working hours, thus higher grade of absorption needs to take place in those time frames. We identified the presence of EVs at those time frames, and demonstrated empirically how the forecast improvement progresses along with

size of the EV fleet. A requirement analysis is conducted by scaling the fleet size and capacity of a static energy storage at the stakeholder's premises.

Finally, for the demonstration of a variable energy storage we adopt one composed of the batteries of the EV fleet. However, clearly this can in the future include any asset that can help absorb the errors produced by a self-forecast. Such assets can somehow appear to "store" energy directly or indirectly (e.g. via transformation to kinetic) or via its selective control strategy (e.g. supermarket freezers, data servers or even emergency power systems). The approach depicted in this work may additionally help realizing energy-autonomous infrastructures in the future as in conjunction with using variable storage, also control and rescheduling or adjustment of processes can complement it, and potentially lead to better energy management.

### Acknowledgement

The authors thank the European Commission, and the partners of the EU FP7 project SmartKYE ([www.SmartKYE.eu](http://www.SmartKYE.eu)) for the fruitful discussions.

### References

- [1] P. Palensky, D. Dietrich, Demand side management: Demand response, intelligent energy systems, and smart loads, *Industrial Informatics, IEEE Transactions on* 7 (3) (2011) 381–388. doi:10.1109/TII.2011.2158841.
- [2] W. Su, H. Eichi, W. Zeng, M.-Y. Chow, A survey on the electrification of transportation in a smart grid environment, *Industrial Informatics, IEEE Transactions on* 8 (1) (2012) 1–10. doi:10.1109/TII.2011.2172454.
- [3] D. Ilić, S. Karnouskos, P. Goncalves Da Silva, S. Detzler, *A system for enabling facility management to achieve deterministic energy behaviour in the smart grid era*, in: 3<sup>rd</sup> International Conference on Smart Grids and Green IT Systems (SmartGreens), Barcelona, Spain, 2014. doi:10.5220/0004861101700178.  
URL <http://dx.doi.org/10.5220/0004861101700178>
- [4] European Commission, SmartGrids SRA 2035 – Strategic Research Agenda, Tech. rep., European Technology Platform SmartGrids, European Commission (Mar. 2012).
- [5] J. Mathieu, P. Price, S. Kiliccote, M. Piette, Quantifying changes in building electricity use, with application to demand response, *Smart Grid, IEEE Transactions on* 2 (3) (2011) 507–518. doi:10.1109/TSG.2011.2145010.
- [6] D. Ilić, S. Karnouskos, P. Goncalves da Silva, M. Griesemer, *An energy market for trading electricity in smart grid neighbourhoods*, in: Digital Ecosystems Technologies (DEST), 2012 6th IEEE International Conference on, 2012, pp. 1–6. doi:10.1109/DEST.2012.6227918.  
URL <http://dx.doi.org/10.1109/DEST.2012.6227918>

- [7] M. L. Goldberg, G. K. Agnew, Measurement and verification for demand response, Tech. rep., DNV KEMA Energy and Sustainability (Feb 2013).
- [8] P. Goncalves da Silva, D. Ilić, S. Karnouskos, [The impact of smart grid prosumer grouping on forecasting accuracy and its benefits for local electricity market trading](#), IEEE Transactions on Smart Grid 5 (1) (2014) 402–410. doi:10.1109/TSG.2013.2278868. URL <http://dx.doi.org/10.1109/TSG.2013.2278868>
- [9] R. Sevlian, R. Rajagopal, Value of aggregation in smart grids, in: Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on, 2013, pp. 714–719. doi:10.1109/SmartGridComm.2013.6688043.
- [10] D. Pudjianto, C. Ramsay, G. Strbac, Virtual power plant and system integration of distributed energy resources, Renewable Power Generation, IET 1 (1) (2007) 10–16. doi:10.1049/iet-rpg:20060023.
- [11] D. Ilić, [Self-Forecasting Energy Load Stakeholders for Smart Grids](#), Ph.D. thesis, Karlsruhe Institute of Technology (KIT) (Jul. 2014). URL <http://digbib.ubka.uni-karlsruhe.de/volltexte/1000042781>
- [12] C. Rasmussen, Improving wind power quality with energy storage, in: Sustainable Alternative Energy (SAE), 2009 IEEE PES/IAS Conference on, 2009, pp. 1–7. doi:10.1109/SAE.2009.5534860.
- [13] D. Ilić, S. Karnouskos, P. Goncalves Da Silva, Improving load forecast in prosumer clusters by varying energy storage size, in: IEEE Grenoble PowerTech 2013, Grenoble, France, 2013.
- [14] P. Poonpun, W. Jewell, Analysis of the cost per kilowatt hour to store electricity, Energy Conversion, IEEE Transactions on 23 (2) (2008) 529–534. doi:10.1109/TEC.2007.914157.
- [15] V. Alimisis, N. Hatziargyriou, Evaluation of a hybrid power plant comprising used EV-batteries to complement wind power, Sustainable Energy, IEEE Transactions on 4 (2) (2013) 286–293. doi:10.1109/TSTE.2012.2220160.
- [16] J. Tomić, W. Kempton, Using fleets of electric-drive vehicles for grid support, Journal of Power Sources 168 (2) (2007) 459–468.
- [17] D. Ilić, S. Karnouskos, [Addressing energy forecast errors: An empirical investigation of the capacity distribution impact in a variable storage](#), Energy Systems, Springer (2014) 1–14doi:10.1007/s12667-014-0119-3. URL <http://dx.doi.org/10.1007/s12667-014-0119-3>
- [18] M. Cho, J. C. Hwang, C.-S. Chen, Customer short term load forecasting by using arima transfer function model, in: Energy Management and Power Delivery, 1995. Proceedings of EMPD '95., 1995 International Conference on, Vol. 1, 1995, pp. 317–322vol.1. doi:10.1109/EMPD.1995.500746.

- [19] S. Aman, Y. Simmhan, V. K. Prasanna, Improving energy use forecast for campus micro-grids using indirect indicators, in: Proceedings of the 2011 IEEE 11<sup>th</sup> International Conference on Data Mining Workshops, ICDMW '11, IEEE Computer Society, Washington, DC, USA, 2011, pp. 389–397. doi:[10.1109/ICDMW.2011.95](https://doi.org/10.1109/ICDMW.2011.95).
- [20] S. Zhu, J. Wang, W. Zhao, J. Wang, A seasonal hybrid procedure for electricity demand forecasting in China, Applied Energy 88 (11) (2011) 3807–3815. doi:[10.1016/j.apenergy.2011.05.005](https://doi.org/10.1016/j.apenergy.2011.05.005).
- [21] W. Kempton, J. Tomić, Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy, Journal of Power Sources 144 (1) (2005) 280–294.
- [22] A. V. Oppenheim, A. S. Willsky, S. H. Nawab, Signals and systems, Vol. 2, Prentice-Hall Englewood Cliffs, NJ, 1983.
- [23] M. Dogru, M. Andrews, J. Hobby, Y. Jin, G. Tucci, Modeling and optimization for electric vehicle charging infrastructure, in: 24th annual POM conference, Denver, Colorado, USA, 2013.
- [24] P. Sanchez-Martin, G. Sanchez, G. Morales-Espana, Direct load control decision model for aggregated EV charging points, Power Systems, IEEE Transactions on 27 (3) (2012) 1577–1584. doi:[10.1109/TPWRS.2011.2180546](https://doi.org/10.1109/TPWRS.2011.2180546).
- [25] PJM, [PJM Manual 18: PJM Capacity Market](http://www.pjm.com), Tech. rep., PJM (www.pjm.com) (Nov. 2013). URL <http://www.pjm.com/~media/documents/manuals/m18.ashx>
- [26] M. Korpaas, A. T. Holen, R. Hildrum, Operation and sizing of energy storage for wind power plants in a market system, International Journal of Electrical Power & Energy Systems 25 (8) (2003) 599–606, 14th Power Systems Computation Conference. doi:[10.1016/S0142-0615\(03\)00016-4](https://doi.org/10.1016/S0142-0615(03)00016-4).
- [27] M. Blank, S. Lehnhoff, Assessing reliability of distributed units with respect to the provision of ancillary services, in: IEEE 11<sup>th</sup> International Conference on Industrial Informatics (INDIN), Bochum, Germany, 2013.