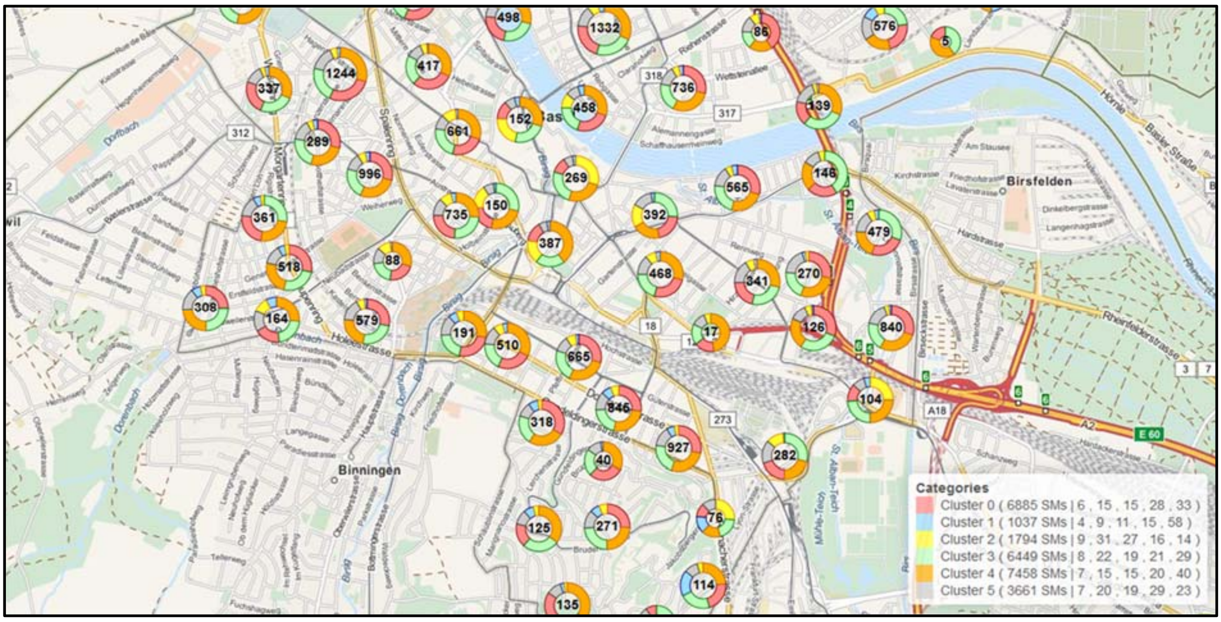
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Department of the Environment, Transport, Energy and Communication DETEC

Swiss Federal Office of Energy SFOE
Energy Research

Final report

Optimized Distribution Grid Operation by Utilization of Smart Metering Data



Date: 21 October 2016

Town: Bern

Publisher:
Swiss Federal Office of Energy SFOE
Research Programme Grids
CH-3003 Bern
www.bfe.admin.ch

Co-financed by:
Industrielle Werke Basel, CH-4002 Basel

Agents:
ETH Zurich, Power Systems Laboratory
Physikstrasse 3
8092 Zurich, Switzerland
www.eeh.ee.ethz.ch/psl

Adaptricity AG
c/o ETH Zurich, Power Systems Laboratory
Physikstrasse 3
8092 Zurich, Switzerland
www.adaptricity.com

Authors:
Dr. Andreas Ulbig, ETH Zurich, Power Systems Laboratory (aulbig@eeh.ee.ethz.ch)
Thierry Zufferey, ETH Zurich, Power Systems Laboratory (thierryz@eeh.ee.ethz.ch)
Dr. Osvaldo Rodriguez Villalon, ETH Zurich, Power Systems Laboratory (rosvaldo@eeh.ee.ethz.ch)
Dr. Stephan Koch, Adaptricity AG (skoch@adaptricity.com)

SFOE head of domain:	Dr. Michael Moser, michael.moser@bfe.admin.ch
SFOE programme manager:	Dr. Michael Moser, michael.moser@bfe.admin.ch
SFOE contract number:	SI/501062-01, SI/501062-02

Summary

The objective of this research project is to investigate methods that utilize Smart Metering data for distribution grid planning and operation. The main motivation for this project is the enhanced visibility and controllability of the electrical distribution system state which can be provided by Smart Metering infrastructure if, and only if, the Distribution System Operators (DSOs) possess suitable methods and tools for data processing, aggregation, analysis, and visualization. To this end, Industrielle Werke Basel (IWB), the utility of the City of Basel, provided extensive anonymized operational datasets from Smart Meter units (40’000+) as well as from commercial load units and distributed generators (PV). A modeling concept for Smart Meters was developed for Adaptricity’s Smart Grid simulation platform *DPG.sim*. A state estimation concept for using Smart Metering data as sensor information for calculating the underlying grid variables has been developed. An extensive data analysis of operational data from IWB has been conducted, including data clustering, data mining and load time-series forecasting methods. This grid operational data has been used for retrospective simulations of the *Kleinhüningeranlage* grid pilot region, a low-voltage distribution grid area in the City of Basel.

Zusammenfassung

Ziel des Forschungsprojekts ist die Untersuchung von neuartigen Methoden, durch welche Smart Metering-Daten speziell für Verteilnetzbetrieb und Verteilnetzplanung nutzbar werden. Die Hauptmotivation für dieses Projekt stammt aus der verbesserten Sichtbarkeit und Kontrollierbarkeit von elektrischen Verteilnetzen, die Smart Metering-Daten für Verteilnetzbetreiber (DSOs) bieten – unter der Voraussetzung, dass DSOs über geeignete Methoden für Datenaufbereitung, Datenaggregation, Datenanalyse und Datenvisualisierung verfügen. In diesem Sinne haben die Industriellen Werke Basel (IWB), der Energieversorger der Stadt Basel, grosse Mengen an anonymisierten Netzbetriebsdaten, u.a. von ihren Smart Metern (40‘000+) als auch von Grosskunden und PV-Anlagen, für dieses Forschungsprojekt bereit gestellt. Im Rahmen des Projektes wurde ein Modellierungskonzept für diverse Smart Meter-Typen für Adaptricity’s Smart Grid Simulationsplattform *DPG.sim* entwickelt als auch ein Netz-Zustandsschätzer (State Estimator) speziell für die Nutzung von Smart Meter-Daten konzipiert. Eine umfangreiche Datenanalyse der von IWB gelieferten Betriebsdaten wurde durchgeführt, inklusive Daten-Clustering und Lastprognosen. Diese Netzdaten wurden für retrospektive Simulationen im Pilotnetzgebiet *Kleinhüningeranlage* in der Stadt Basel durchgeführt.

Résumé

L’objectif de ce projet de recherche est d’examiner des méthodes se basant sur les données de compteurs intelligents pour la planification et la gestion du réseau de distribution électrique. L’infrastructure des compteurs intelligents permet en effet une meilleure visibilité et contrôlabilité de l’état du système de distribution si, et seulement si, le gestionnaire du réseau de distribution possède les méthodes nécessaires pour la préparation, l’agrégation, l’analyse et la visualisation des données collectées. À cette fin, Industrielle Werke Basel (IWB), les services industriels de la ville de Bâle, a fourni une grande base de données opérationnelle provenant de leurs compteurs intelligents (40’000+) ainsi que de consommateurs commerciaux et d’unités de production décentralisée (PV). À l’aide du logiciel *DPG.sim* d’Adaptricity, les données de compteurs intelligents ont été utilisées pour calculer les variables sous-jacentes du réseau de distribution et ainsi estimer l’état de ce réseau. Les violations des valeurs autorisées de tension ainsi que la surcharge de certaines composantes du réseau peuvent donc être anticipées en temps réel, en combinant des prévisions se basant sur les données historiques des compteurs avec le modèle du réseau de distribution. Une analyse complète des données opérationnelles fournies pas IWB a été conduite. Ceci inclut le regroupement de compteurs aux caractéristiques similaires et la prévision de consommation ou de production. Ces données ont permis de simuler rétrospectivement la région pilote *Kleinhüningeranlage* à Bâle.

Contents

List of abbreviations

IWB	Industrielle Werke Basel
kWh	Kilowatt-hour
Wh	Watt-hour
kW	Kilowatt
MW	Megawatt
PV	Photovoltaic(s)
DSSE	Distribution system state estimation
SE	State-Estimation
WLS SE	Weighted-Least Squares State-Estimation
RMSE	Root Mean Squared Error
MAPE	Mean Absolute Percentage Error
CAPEX	Capital Expenditure
OPEX	Operational Expenditure
OLTC	On-Load Tap Changing
BESS	Battery Energy Storage System

1. Introduction

1.1 Trends in Distribution Grid Operation

From passive to active distribution grids

Due to the rise of distributed generation, increasingly controllable thermal load units such as heat pumps and air conditioning as well as the slowly but steadily rising share of grid-connected distributed battery storage, distribution grids cease to be as passive as in the past. Distribution grid behaviour becomes increasingly dynamic, eventually also requiring a more active grid operation and changes to grid planning processes.

Smart Grid technologies are emerging that are able to actively influence electric load and generation profiles in order to improve grid operation and reduce grid loading stress.

Traditional distribution grid as well as transmission grid planning is usually based on a static load flow computation, a so-called snapshot situation, which takes into account the maximum coincident load to determine the necessary grid infrastructure dimensioning. With more distributed generation, particularly wind and solar PV generation, the snapshot of the maximum coincident generation is also becoming relevant for grid infrastructure dimensioning.

The coincidence of load demand and distributed generation in time and space may either reduce or actually impose more stress on the grid than load demand would by itself. Merely considering the grid's overall worst-case load snapshot can lead to highly over-dimensioned parts of the grid infrastructure and rare utilization of the full grid capacity. In fact, the worst-case snapshot approach can also lead to critical under-dimensioning of other parts of the grid infrastructure, as their particular worst-case loading may simple occur at another point in time.

More available operational data in distribution grids

Distributed generation units and Smart Metering systems are also collecting a steadily rising amount of data sets, today usually collected only for billing purposes only. However, these data sets can potentially also be used as sensor data with the purpose of improving distributed grid operation and grid planning.

Industrielle Werke Basel (IWB), the utility of the City of Basel, as the first distribution grid operator in Switzerland decided to make a large-scale Smart Meter roll-out for all its residential customers. The data from these by now more than 40'000 deployed Smart Meter units is a rich source for distribution grid analytics by itself.

Together with additional measurement data from PV units, commercial customers as well as transformer readings for selected distribution grid pilot regions, these data sets enable a wide range of analysis ranging from classical data analytics to detailed grid simulation and state estimation.

Towards distribution grid transparency

The currently emerging role of Smart Metering systems for collecting measurement data form distribution grid operation bring opportunities beyond simply creating electricity bills (meter-to-cash).

Notably, highly localized measurements of electric load profiles can be employed for calculating the current grid operation state via so-called state-estimation, thereby achieving a better grid transparency, i.e. knowledge of the plausible voltage and line loading profiles, as is available in distribution grids today.

Especially on the low-voltage level (400V), distribution grids are still mostly operated in a black box fashion (Figure 1) as there is little to no operational knowledge of what happens at this grid level. At best, grid operators currently track the power consumption at the low-voltage transformer going into a given distribution grid area.

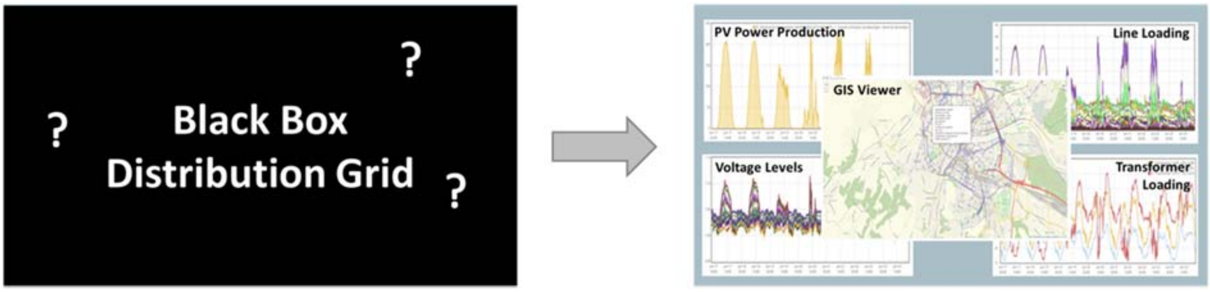


Figure 1 – *Situational awareness in distribution grid operation and distribution grid planning Today: almost no operational knowledge (black box), Future: grid transparency via state-estimation.*

1.2. Smart Grid Simulation Platform *DPG.sim*

Traditional distribution grid as well as transmission grid planning is usually based on a static load flow computation, a so-called snapshot situation, which takes into account the maximum coincident load to determine the necessary grid infrastructure dimensioning.

With the rise of Distributed Generation (DG), particularly wind and solar PV generation, the snapshot of the maximum coincident generation is also becoming relevant for grid infrastructure dimensioning. The coincidence in time and space of load demand and distributed generation may either reduce or actually impose more stress on the grid than load demand would by itself. Merely considering the grid's overall worst-case load snapshot can lead to highly over-dimensioned parts of the grid infrastructure and rare utilization of the full grid capacity. In fact, the worst-case snapshot approach can also lead to critical under-dimensioning of other parts of the grid infrastructure, as their particular worst-case loading may simple occur at another point in time.

Emerging Smart Grid technologies have been proposed that are able to actively influence electric load and generation profiles in order to improve grid operation and reduce grid loading stress. Studies show that innovative operational measures, such as selective renewable energy curtailments, Demand Response, distributed storage, and reactive power control, can potentially make the transition to high shares of renewable energies more cost-effective by reducing otherwise needed grid upgrade costs [dena2012].

In order to reduce grid infrastructure costs by using one or a combination of several Smart Grid technologies, these Smart Grid technologies need to be considered explicitly at the grid planning stage. However, the respective literature shows that existing grid planning tools do in many cases not support the realistic modelling and simulation of electricity grids with Smart Grid elements, a necessary prerequisite for cost-effective grid planning [CIGRE2014].

In order to overcome these limitations, the time-series based Smart Grid simulation platform *DPG.sim* (Distributed Prosumer and Grid Simulation) is developed and commercialized by ETH Zurich spin-off Adaptricity. The detailed simulations and subsequent grid analytics provide valuable qualitative as well as quantitative decision-support for all aspects of distribution grid operation, e.g. the design and performance analysis of active network management operation strategies, as well as distribution grid planning, e.g. the integration of Smart Grid elements into distribution grid planning.

The main advantages of the Smart Grid simulation platform *DPG.sim* are threefold:

- First, its possibility to realistically model and simulate the operation of active distribution grids as well as the temporal evolution of generation, load, storage states, and their operational control algorithms down to the level of individual households and household units. Its unique feature in this respect is its versatile Prosumer modelling approach [Heussen2012], which allows capturing all relevant modelling details and operational constraints of controllable loads, distributed generation, and storage as well as Smart Meter communication infrastructure (Figure 2).
- Second, the ability to perform large-scale time-series simulations and operational (big) data analytics based on heterogeneous sets of grid data as well as end-consumer data sets by tapping into scalable cloud-based computation and data storage resources.

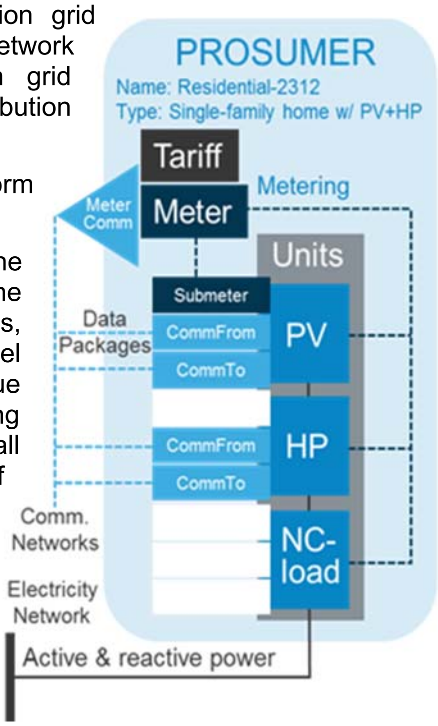


Figure 2 – *Prosumer Modeling Approach.*

- Third, the ability to integrate heterogeneous data sets stemming from grid operation and grid planning as input data for grid simulations and to provide data visualization and statistics tools for the simulation output data. This allows detailed insights into electricity grid operation and enable robust decision support based on already available measurement data for a given grid area, however incomplete (Figure 3).

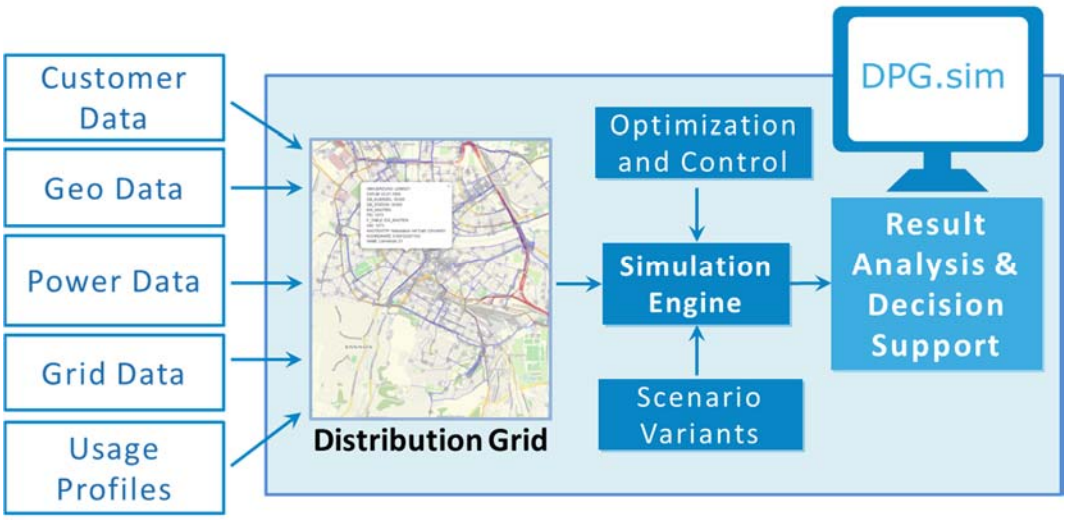


Figure 3 – *Structure of Simulation Platform DPG.sim.*

1.3. Project Goals

The objective of this research project is to investigate methods that utilize Smart Metering data for distribution grid planning and operation. The main motivation for this project is the enhanced visibility and controllability of the electrical distribution system state which can be provided by Smart Metering infrastructure if, and only if, the Distribution System Operators (DSOs) possess suitable methods and tools for data processing, aggregation, analysis, and visualization. To this end, Industrielle Werke Basel (IWB), the utility of the City of Basel, provided extensive operational datasets from their Smart Meter units (40’000+) as well as from commercial load units and distributed generators (PV).

The main goals of this project are the development of new strategies and methods, which enable distribution grid operators to exploit Smart Metering data and the advantages that this data brings both in distribution grid operation and distribution grid planning. The identified advantages are, more specifically, the following:

1. **Retrospective load flow analysis and state-estimation inside distribution grids**
   → Leading to improved understanding of load flows in lower-voltage grid levels (LV and MV)

2. **Load prediction in distribution grids by exploiting Smart Metering data**
   → Improvement of highly localized load predictions (i.e. per LV transformer station)

3. **Performance validation of demand response strategies**
   → Validation of actual load demand behavior (by Smart Metering) against load schedules

4. **Distribution Grid Monitoring**
   → Overview and analysis of distribution grid behavior down to the low-voltage grid level

2. Project Work Packages and Results

2.1 Work Package Overview

The project is structured into three main work packages that either focus on global applications (G) or local applications (L), where the Power Systems Lab of ETH Zurich is the main project contributor:

- **ETH-G1 – Retrospective load flow analysis based on Smart Metering-like measurements of load consumption and distributed generation**

  The work package includes the following tasks:

  1. Implementation of retrospective load flow simulations (based on measured transformer, Smart Meter and PV power profiles)
     - Retrospective load flow (based on full measurement information)
     - Retrospective load flow (based on partial measurement information and state-estimation)
  2. Implementation of operational data visualization methods, showing notably
     - Loading of grid components (i.e. transformer, lines)
     - Statistical analysis of voltage and load demand profiles (i.e. box plots, histograms)
  3. Load flow analysis of IWB data
     - Low voltage pilot grid region
     - Medium voltage grid region

  These tasks enable the following applications:

  1. Realization of retrospective high-resolution load flow analysis in distribution grids
  2. Analysis of grid situation specifically in low-voltage grids
  3. Insights into the causes for specific voltage and line loading anomalies (i.e. identifying the responsible load and generation units)
  4. Simulation and analysis of the value of partial information, i.e. only active power measurements versus active & reactive power measurements as well as local voltage measurements

  Potential applications are identified to be located:

  1. Directly with the distribution grid operator (global applications)
  2. Relevant for both distribution grid operation (grid state analysis, state-estimation) as well as distribution grid planning (better insights into grid behavior as well as operational contingencies)

- **ETH-G2 – Improved load predictions in distribution grids by exploiting Smart Metering measurements**

  The work package includes the following tasks:

  1. Predictions of Smart Metering time-series (load demand)
  2. Time-series data analytics via neural networks and clustering methods
  3. Sensitivity analysis of aggregation size (individual household, transformer station level, sub-station level) and its impact on load prediction quality (mean, min/max error)
  4. Comparison of Smart Meter predictions with real-time data from transformer stations via regression methods

These sub-tasks enable the following applications:

1. Usage of various Smart Metering measurements, i.e. both production and consumption for improving prediction quality
2. Day-Ahead predictions on the aggregation level of transformer stations (NE6) as well as sub-stations (NE4)

Potential applications are identified to be located:

1. Directly with the distribution grid operator (global applications)
2. Directly with power procurement and power trading

- **ETH-L1 – Distribution grid monitoring for active distribution grid operation and planning**

The work package includes the following tasks:

1. Concept development for grid-friendly local control strategies
2. Local voltage regulation
3. Reactive power provision
4. PV / load curtailment
5. Battery applications

These sub-tasks enable the following applications:

1. Reduction of grid component loading, specifically
    a. Transformer station loading
    b. Line loading
2. Traffic-light style grid operation concepts

Potential applications are identified to be:

1. Local control strategies applied on
    a. Low-voltage test grid(s)
    b. Medium-voltage test grid(s)

Other project tasks are supporting the above main activities and are led by either IWB or Adaptricity:

- **IWB-1 – Provision of grid models and grid measurements**

The work package includes the following tasks:

1. Grid models
    a. City of Basel (medium-voltage) and
    b. One or two pilot grid regions (low-voltage), e.g. *Kleinhüningeranlage*
2. Grid operational data
    a. Smart Metering measurements, i.e. 15min values, have been provided for a time-period of more than two years
    b. PV production measurements
    c. Industrial/commercial customer measurements

(All data has been provided while observing the existing data protection regulations.)

- **IWB-2 – Validation of simulation and analysis results**

Grid experts from IWB receive all analysis results from ETH Zurich and Adaptricity, including a software license of DPG.sim and evaluate the results for plausibility and relevance.

- **Adap-1 – Software development and provision**

  Adaptricity provides its Smart Grid simulation platform *DPG.sim* and developed additional functionality regarding the modeling of Smart Meters, their functionality as well as actual measurement data.

- **Adap-2 – Integration of project concepts and results**

  Methods developed during the course of the project have been integrated into *DPG.sim* as an add-on toolbox. The software suite *DPG.sim Academic* has recently been released and is available for Smart Grid researchers and industry collaborations (www.adaptricity.com/academic/). The professional version, *DPG.sim Academic professional*, includes all developed Smart Metering functionalities.

- **IDS**

  The German IDS group is an advisor *at-large* during the course of the project, providing valuable insights regarding industry applicability and result validation.

2.2. Work Packages Adapt-1 and Adapt-2 – Developed Software Functionalities in *DPG.sim*

Modelling of Smart Meter data recording capabilities and implementation in DPG.sim

A multitude of Smart Meter systems, offering a wide and often quite differing spectrum of measurement capabilities, is available. More than 100 Smart Meter system datasheets were screened for this purpose. The following measurement capabilities of physical grid-related system states are possible:

- Voltage values and voltage angle,
- Current values as well as
- Frequency values.

In addition to this comes any other information that can be calculated, estimated or otherwise inferred from these base measurements. Not all these values are in fact relevant for the type of time series simulations performed in *DPG.sim*. This includes for instance local frequency measurements.

Capabilities of Smart Meters that are relevant for the simulations in *DPG.sim* include the following:

- Active energy consumed or produced,
- Net active energy consumed,
- Reactive energy consumed or produced,
- Net reactive energy consumed,
- Active power (min/max) consumed or produced,
- Net active power (min/max) consumed
- Reactive power (min/max) consumed or produced and net reactive power (min/max) consumed
- Voltage level (mean and min/max)

(Here, the terms ‘min/max’ and ‘mean’ are meant in reference to the given measurement time period, e.g. 1 min. or 15 min.)

Important for the simulations in *DPG.sim* are also the communication capabilities offered by Smart Meter systems, i.e. communication bandwidth and delay, and the actually implemented communication scheme, ranging in principal from (near) real-time, two-way communication to one-way transmission of consolidated data sets once per day. The latter is the current industry standard and also the status at IWB.

Within the simulator platform *DPG.sim* a freely configurable Smart Meter model was implemented (see parameterization interface shown in Figure 4).

Incorporating all the Smart Meter specifications that are relevant for time-series simulations, including their communication schemes, has become a key feature of *DPG.sim*.

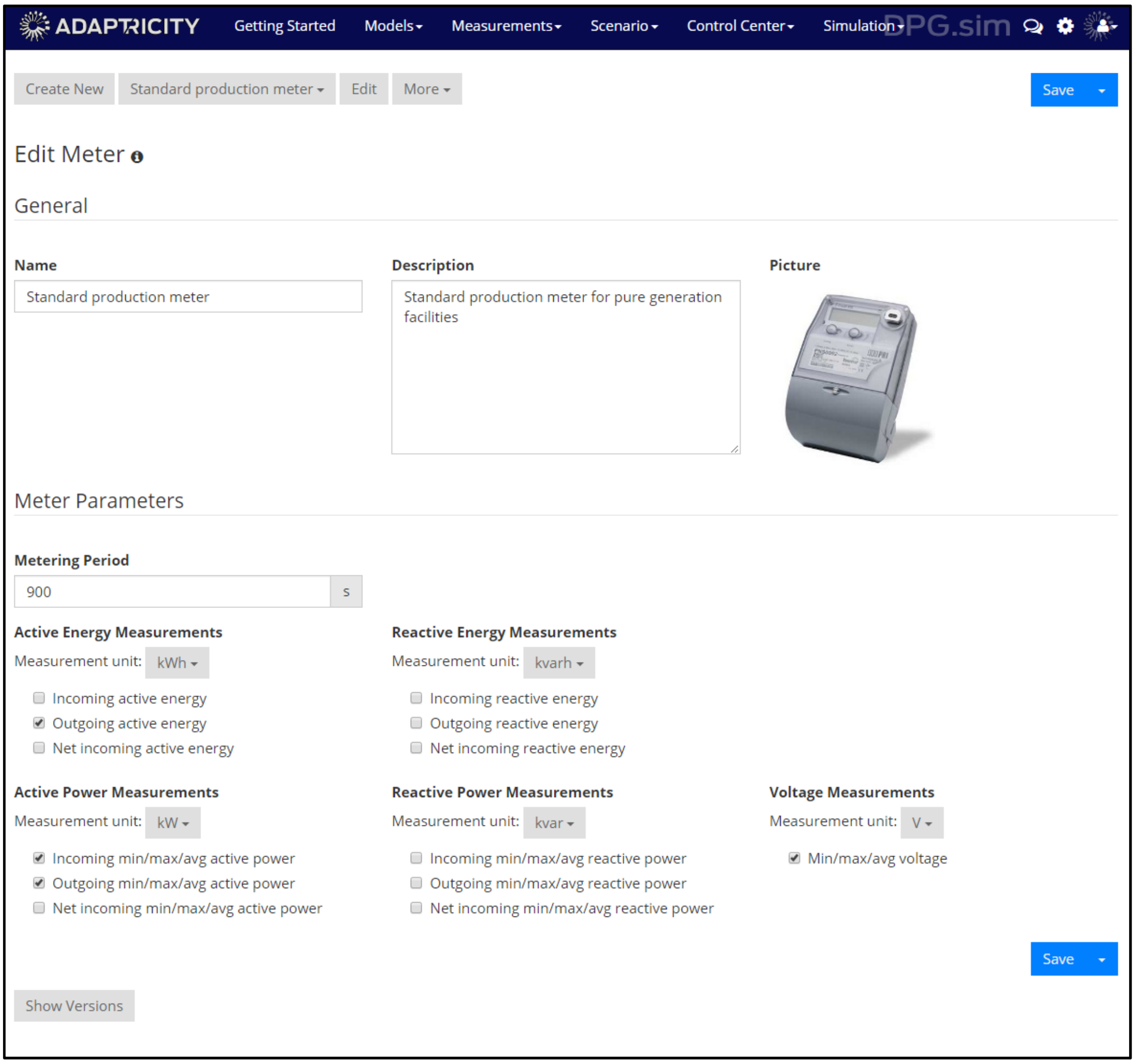


Figure 4 – *Screenshot of Smart Metering parameterization interface in DPG.sim*

Implementation of Smart Meter time series import

An interface has been defined that allows the import of real Smart Metering data into *DPG.sim*.

It is composed of a time series importer with preview functionality (see Figure 5), a function to filter and group Smart Metering time series into subsets for specific purposes, and a module for creating Prosumer objects for the time simulation out of the Smart Metering data.

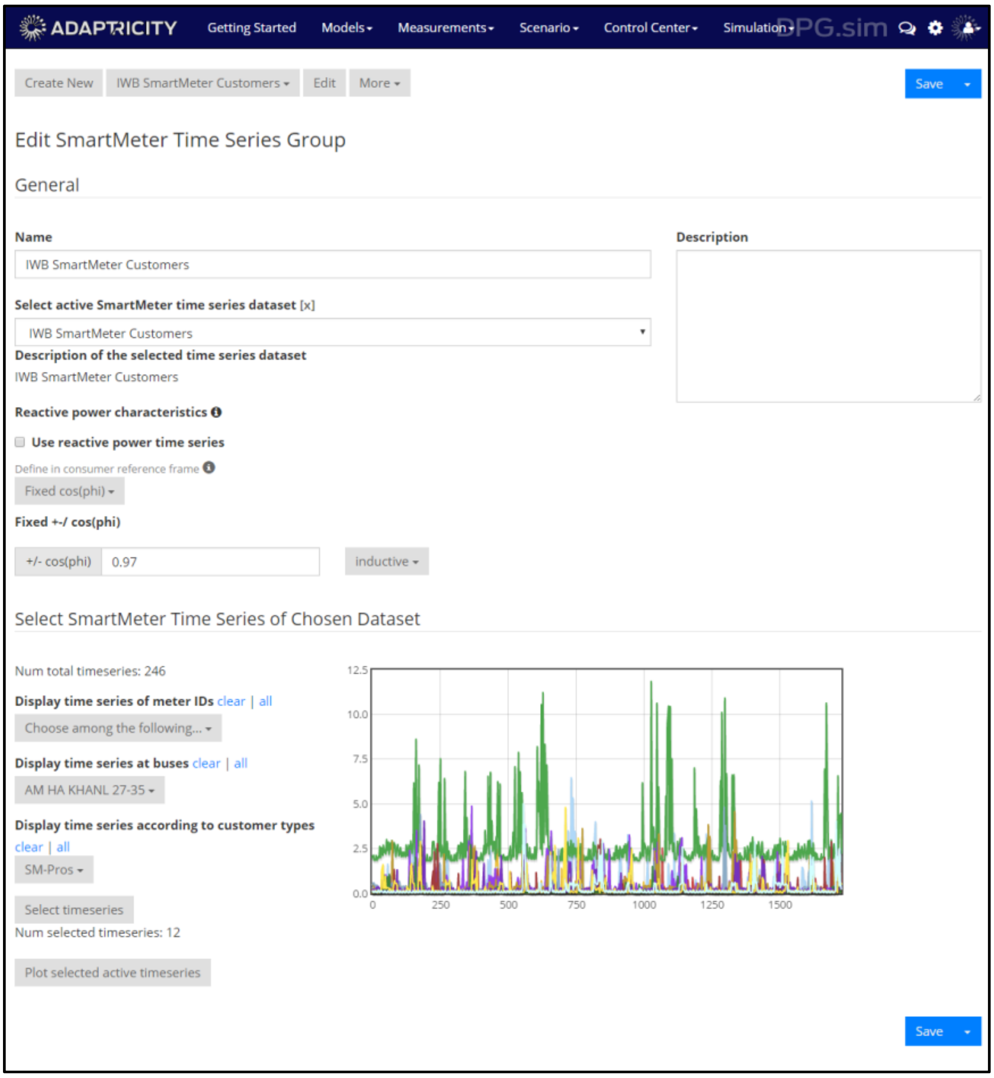


Figure 5 – *Smart Meter time series import and editing in DPG.sim*

Data exports from IWB and import into DPG.sim

For the scope of this project, real-life Smart Metering data from Industrielle Werke Basel (IWB) for the time-frame 2014-15 has been used for running time simulations within *DPG.sim*. As is typically the case, the usage of real-life operational data requires some effort for data checking and data cleaning as well as for data provisioning, import and processing into and inside the simulation platform.

The following data exchanges were initiated between IWB and ETH Zurich/Adaptricity:

- **Grid topology**

  The electrical data of the IWB medium voltage network was provided as a NEPLAN project file by IWB. From there, it was imported into *DPG.sim* by Adaptricity.

- **Geographical Information System (GIS) data**

  The geographical data of the entire medium-voltage network as well as some pilot region parts of the low-voltage network of the City of Basel was exported as a Shapefile from the GIS system of IWB. An import routing was written by Adaptricity in order to display the grid topology on a geographical map (GIS layer) within *DPG.sim*. An exemplary screenshot is depicted in Figure 6.

- **Metering data**

  The Smart Meter data from about 40'000 households (15-min active energy consumption) have been provided by IWB. Furthermore, generation data of photovoltaic (PV) as well as consumption data from large individual customers have also been taken into account.

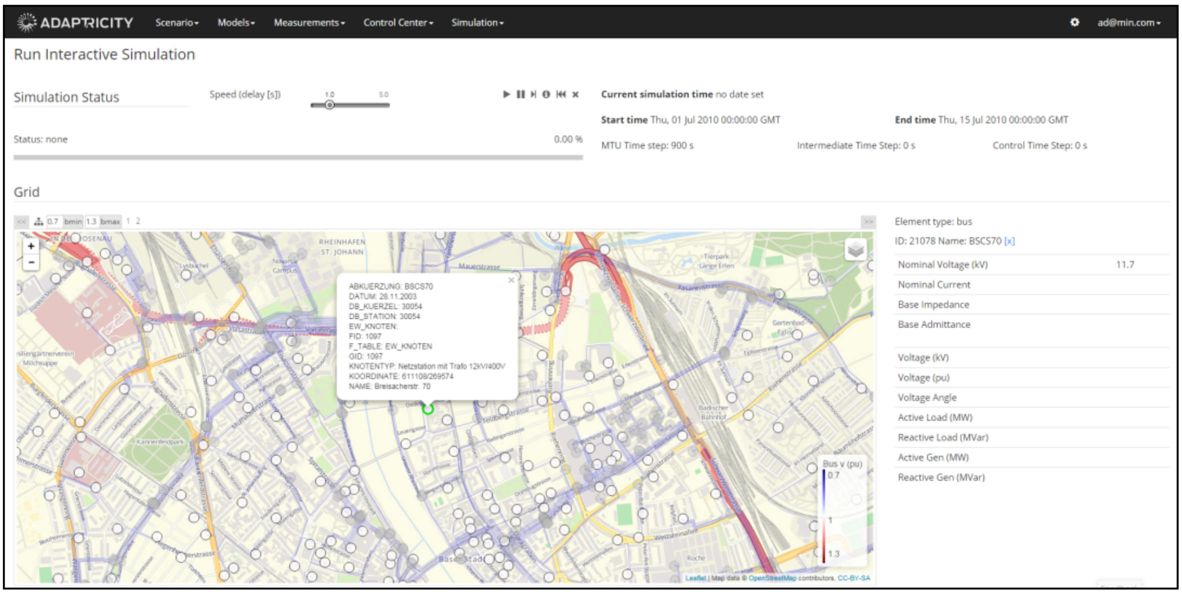


Figure 6 – *Grid topology on GIS layer in DPG.sim*

2.3. Work Package ETH-G1

2.3.1. Predictive state estimation concept for exploiting Smart Metering data

Smart Meters are sensor elements that provide accurate high resolution measurements on both the spatial scale, i.e. on a household level, and the temporal scale, i.e. every few minutes, for parts of the grid topology for which previously only highly aggregate measurements, i.e. on the substation level, were available. The on-going roll-out of Smart Metering equipment and the thereby obtained measurement capabilities on the low-voltage grid level have opened a large opportunity to obtain fruitful operational data which was normally not available without explicit measurement efforts by the grid operator. Some utilities have been used historical data and so-called pseudo-measurements, i.e. synthesized data that plausibly resembles real measurements, in order to determine the network's state or to make the network observable.

Thus, Smart Meters are sensor elements that provide accurate high resolution measurements on both the spatial scale, i.e. on a household level, and the temporal scale, i.e. every few minutes, for parts of the grid topology for which previously only highly aggregate measurements, i.e. on the substation level, were available. In the analysis performed in this project, measurement sampling times of 1 min. and 15 min. have been considered. In principle, these high resolution Smart Meter datasets allow a previously unattainable detailedness in state estimation and other grid analysis functionality, for example high resolution power flow calculations and with it very detailed voltage level and grid constraint assessments.

Normally Smart Meters do not provide real-time information to the grid operator. Instead their measurements are most often collected and transmitted as a consolidated data set, for example, once per day. This is fully sufficient for all retrospective grid simulation and analysis tasks. Nevertheless the information collected by Smart Meters is still valuable even for real-time applications. Combining high resolution but not in real-time available Smart Meter data sets with real-time available but rather coarse (aggregated) substation measurements can still enable real-time applicable state estimation and power flow predictions with unprecedented detail, as will be presented in the following sections.

Within *DPG.sim* real Smart Meter data has been incorporated into a simulation scenario in order to estimate the state of the distribution grid. Medium voltage (MV) and low voltage (LV) grids have been analyzed in the cases studies described in next sections. Whereas in medium-voltage grids it is sufficient to assume aggregated load measurements on a per bus basis, in low-voltage grids more fine-grained localized load measurements, ideally for each prosumer household, are required for a good estimation quality.

With the advent of increasingly smart and, generally, more actively operated distribution grids, new state-estimation techniques have to be capable of dealing efficiently with radial topologies and highly meshed grids; in addition, due to the availability of thousands of real measurements, state-estimation methods can be efficient tools in order to characterize the correct operation of distribution networks [Alimardani et al. 2015]. The tasks in this section of the research project follow this concept via realistic simulations and scenarios of distribution grids with Smart Meters data.

State Estimation Algorithm

The usefulness of state estimation (SE) as a means to provide real time monitoring of the power grid is quite limited. The state of the distribution network is commonly estimated using Weighted Least Squares (WLS) methods, initially applied to the transmission networks analysis and adapted to the case of distribution networks [Manitsas et al. 2012].

Consider an N bus system, with $(2N-1)$ state vector with the form,

$$\mathbf{x} = \left[\theta_2, \theta_3, \ldots, \theta_n, |V_1|, |V_2|, \ldots, |V_n|\right]^T ,$$

where θ_i are phase angles, $|V_i|$ are voltage magnitudes and the phase angle θ_1 is assumed to be known (as it corresponds to the slack bus of the power system).
To estimate the power system state x, a set of $L > 2N-1$ independent measurements z are collected. The measurements are related to the state vector by an over-determined system of linear equations

$$z = h(x) + n \quad ,$$

where $h(\cdot)$ is a set of L nonlinear functions of the state vector.

The number of needed measurements (L) is determined by the length of the state vector x (2N-1), which in turn is determined by the network size (N bus system with an NxN network admittance matrix), n is a zero-mean Gaussian measurement noise vector (i.e. the measurement error).
In the traditional SE approach, the state vector is estimated from the measurement equation using the WLS method.

Figure 7 is showing a flow chart of the algorithm followed to apply the WLS state estimation method. In the case of distribution networks, the WLS could present difficulties with measurement function calculation, since this function is formed with the inverse of the partial derivatives of power equations with respect to angles and voltages and also due to the natural sparsity that these grids could present.

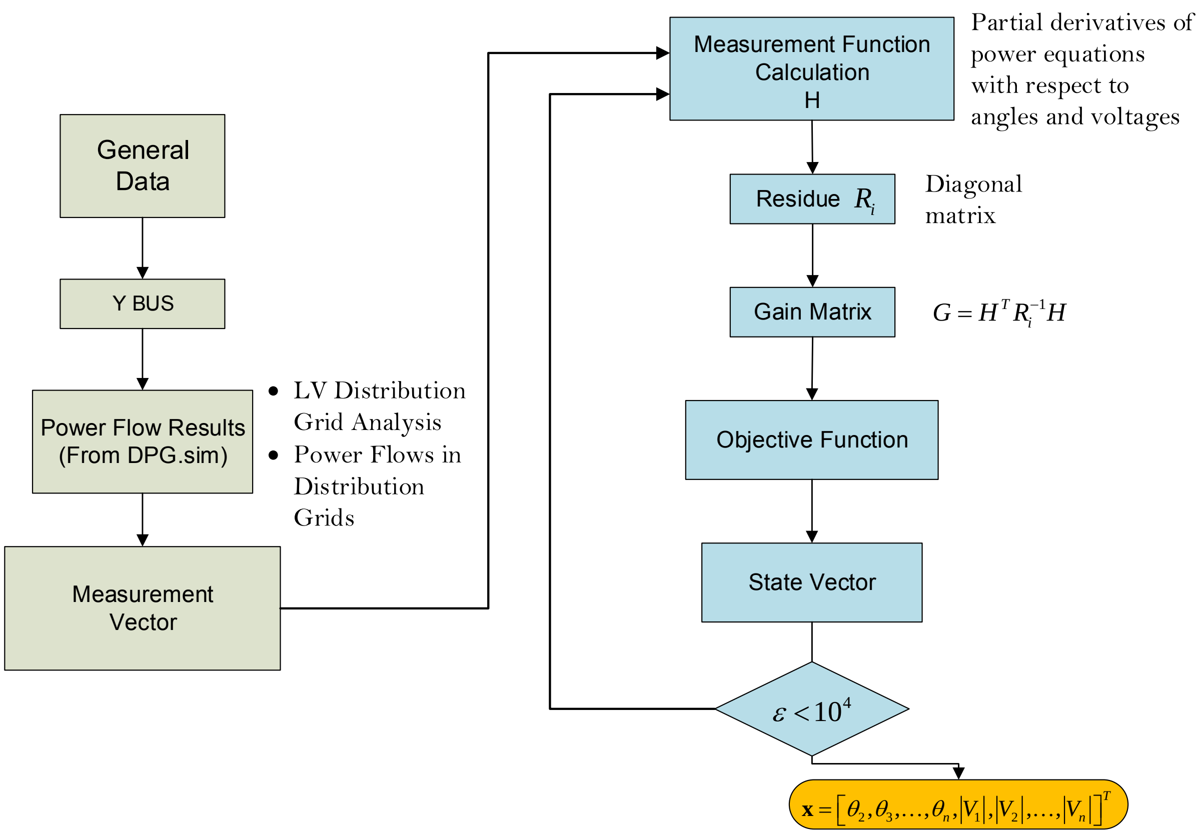


Figure 7 – *State Estimation Algorithm Flowchart*

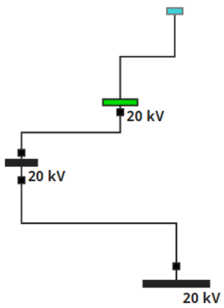
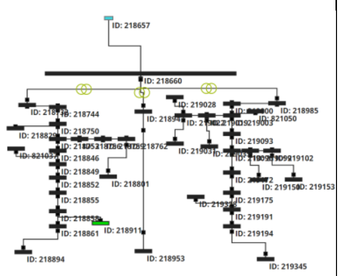
Test Case Scenarios

The interface created to obtain the data from *DPG.sim* and the state estimation algorithm was developed conducting several steps of data analytics in MATLAB. The main variables considered as input measurements for the state-estimation (SE) method are: voltage magnitudes, active and reactive power injected to the grid buses.

The test case scenarios (Table 1) are simulated via *DPG.sim* considering the models of prosumers, and defining a pre-established grid topology**.**

Figure 8 is showing the face page of the scenario in *DPG.sim* for the case of the MV grid of 12 buses [Singh et al. 2009]. From the simulation is obtained the power flow results and the measurements of the loads connected to the grid. In this scenario, prosumer with stochastic loads are selected, due to their similarities with a real load scenario.

Table 1 – Test cases characteristics and simulation scenarios for DSSE

TEST CASE	TEST SYSTEM GRID	PROSUMER TYPES	SMART-METER VARIABLES	STATE ESTIMATION VARIABLES
Test Case *3 Prosumers loads, PV battery* *(MV grid)*		• Household with constant load • Household (random load sampling) with PV, battery and self-consumption controller • Household (random load sampling) with PV, battery and arbitrage controller • Household with non-controllable load with exponential sampling	**Case 1** *Perfect Case* Energy values (P, Q) V (slack bus) **Case 2** Energy values P Bus Voltages (from Smart Meters)	Voltage magnitude $\lvert V_i \rvert$ Voltage angle δ_i
Distribution Test System 12 Buses (MV grid)		• Household (random load sampling) with PV, battery and self-consumption controller • Household (random load sampling) with PV, battery and arbitrage controller • Household with non-controllable load and wind unit • Household with non-controllable load and thermal load	**Case 1** *Perfect Case* Energy values P, Q, V (slack bus) **Case 2** Energy values P Bus Voltages (from Smart Meters)	Voltage magnitude $\lvert V_i \rvert$ Voltage angle δ_i
CIGRE Benchmark Test System (LV grid)		Three different prosumer types: • Domestic prosumers • Industrial prosumers • Commercial Prosumers	**Case 1** *Perfect Case* Energy values P, Q, V (slack bus) **Case 2** Energy values P Bus Voltages (from Smart Meters)	Voltage magnitude $\lvert V_i \rvert$ Voltage angle δ_i

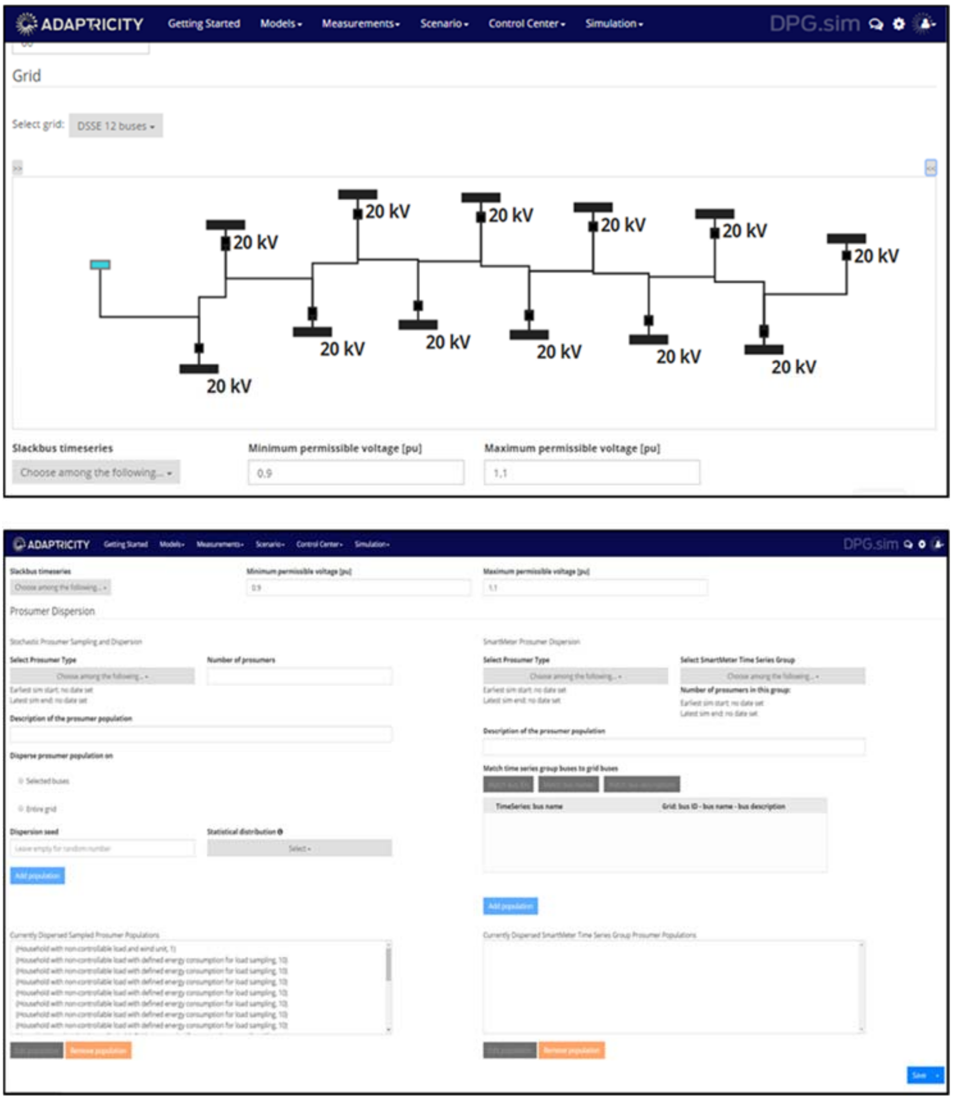


Figure 8 – *Test case scenario in DPG.sim*

The data provision of a simulation run in *DPG.sim* includes the following input parameters for the state estimation routine, i.e. simulation time step, Smart Meter measurements, simulation scenario code, Smart Meter names, measurement type, bus IDs, grid admittance and susceptance matrix, slack bus conditions, active and reactive power injected to each bus (Figure 9).

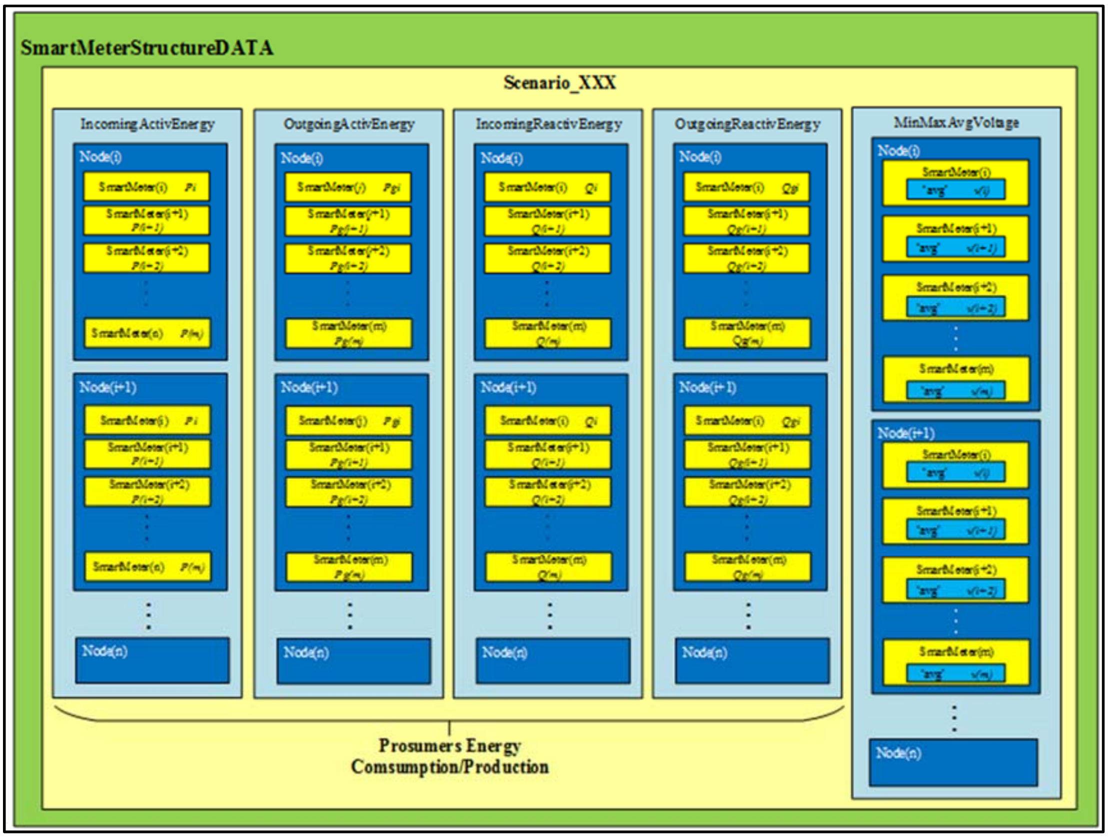


Figure 9 – *Data structure obtained from DPG.sim*

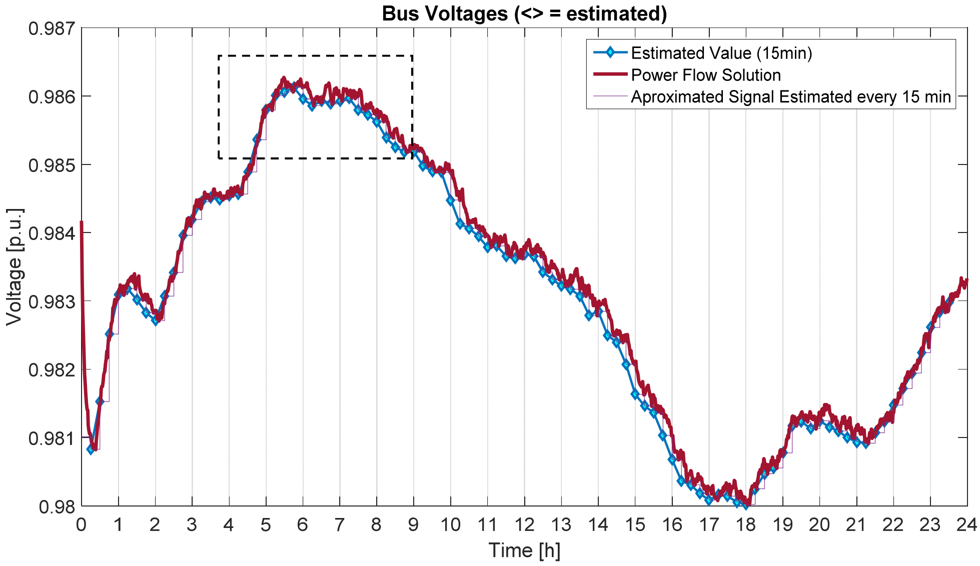
Simulations and results – State estimation with 1 min. vs 15 min. measurement sampling times

The experiment is designed as follows: a set of measurements with time sampling of 15 min. is utilized in order to apply DSSE to the test case scenario, followed by a comparison with respect to the analysis of the network obtained with a time step of 1 min. (power flow solution).

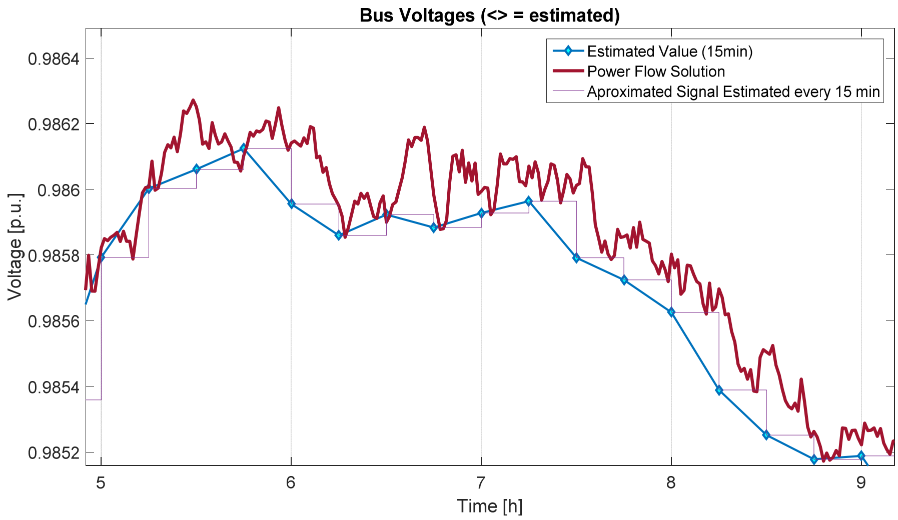
- State estimation is performed using 15 min. Smart Meter voltage measurements, which corresponds to average value of 15 simulation steps (1 min. simulated voltage values).
- The state estimation results are showing the differences between the estimated bus voltages (based on 15 min. Smart Meter measurements) and the instantaneous bus voltages (based on 1 min. simulated voltage values).
- Relative errors are depicted, as well as histogram of the error distribution. It is clearly observed that the error is following a Gaussian distribution, between limits of $\pm 10^{-3}$.

The distribution system state estimation (DSSE) analysis is considering as input data the model of the network in terms of topology and branch parameters, measurement units, types of prosumers, voltage magnitude and active and reactive power injections. These data are provided by DPG.sim according to the structured showed before.

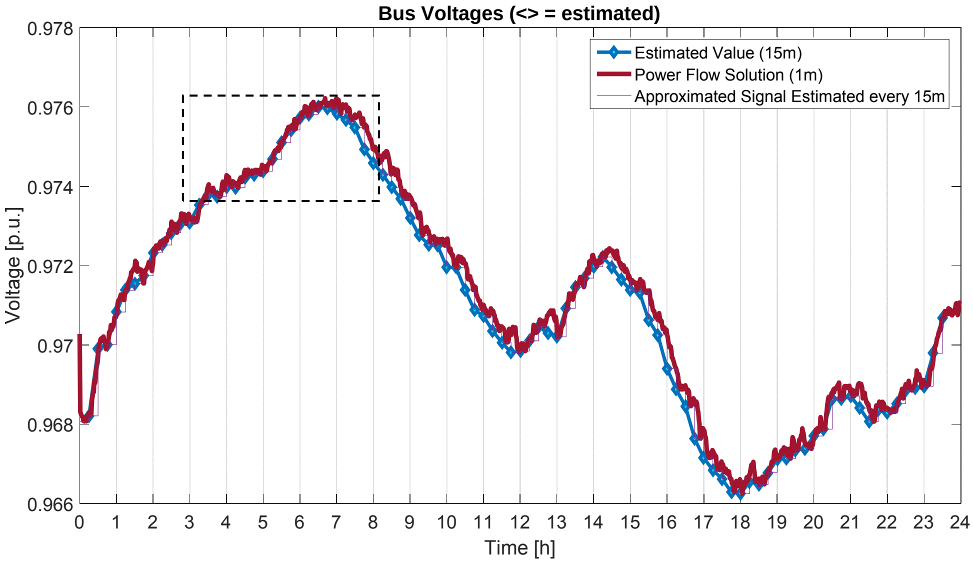
The experiment is designed as follows: a set of measurements with time sampling of 15 min is utilized in order to apply DSSE to the test case scenario, followed by a comparison with respect to the analysis of the network obtained with a time step of 1 min (power flow solution – *“simulated reality”*). Figure 10 is showing the state estimation of a voltage from the test case scenario of 12 buses (MV grid) when energy measurements of active power, reactive power and slack bus voltage reference are selected. The graph *"estimated value (15 min)"* shows the average value of 15 steps with 1 min. step solution. The comparison between power flow solution and estimated values are the core result. The estimated values can be seen as discrete points obtained every 15 min; however, it is possible to join these points in order to build a continuous graph (signal in the plot with blue diamonds). The solid stairs line in the plot is described as *"approximated signal estimated every 15 min"*, which is a discrete signal from every point estimated over the solution. The results clearly show the difference between both solutions, but the estimated graph is showing always the same tendency in amplitude and phase with respect to the power flow solution (simulated reality).



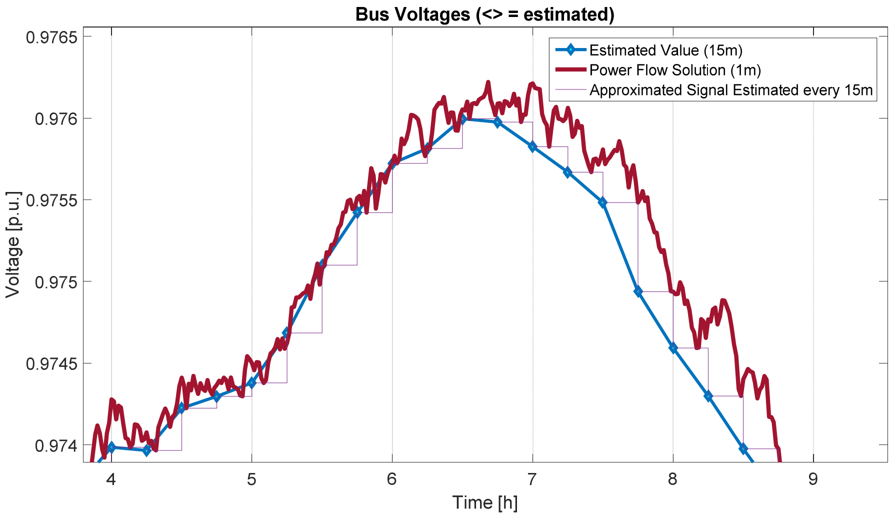
a) Voltage in bus 3 for the 12 bus MV test grid



b) Zoomed area from a) of the voltage in bus 3 for the 12 bus MV test grid



c) Voltage in bus 7 for the 12 buses MV test grid



d) Zoomed area from c) of the voltage in bus 7 for the 12 buses MV test grid

Figure 10 – Voltage graph comparison between 15 min DSSE versus 1 min power flow solution

In Figure 11, the state estimation solution for the case study of the CIGRE LV benchmark (detailed in Table 1) for the voltage of bus 12. The bus analyzed corresponds to the leg that belongs to the domestic load. This case study is in fact a special case, so designed [Papathanassiou et al. 2005] as it includes loads from different types (domestic, commercial and industrial). The plot observed here is depicting the differences between the power flow solutions with respect to the state estimation calculated every 15 min over a one day period. The load possesses characteristics of stochastic behavior, energy measurements of active power, reactive power and a voltage reference (slack bus).

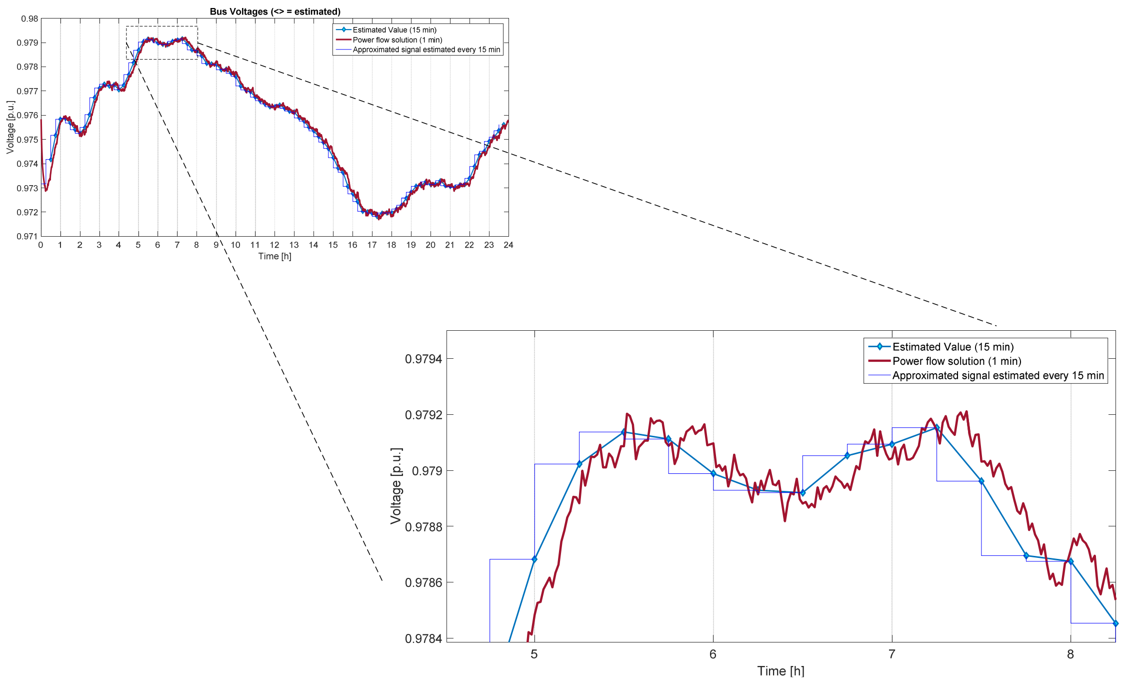
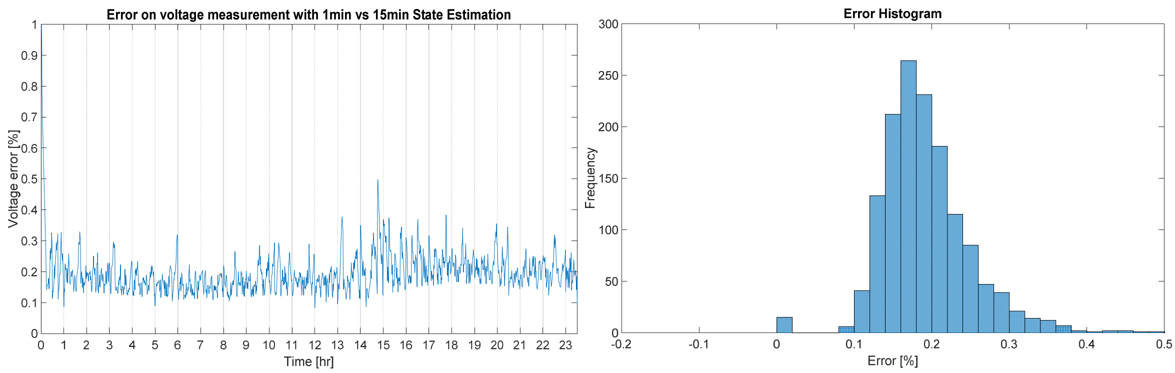
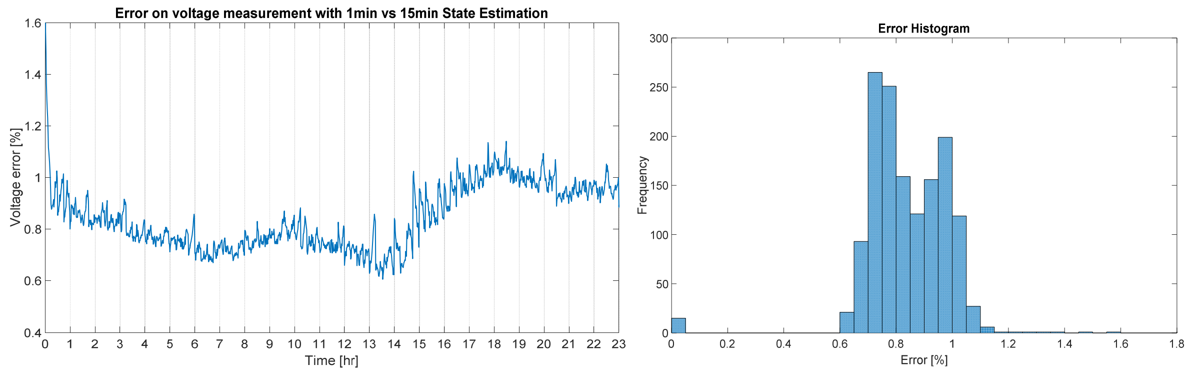


Figure 11 – Voltage graphs comparison between 15 min DSSE with respect to 1 min power flow solution for the LV CIGRE test case study over bus 12 (domestic load)

Additionally, Figure 12 is depicting the errors obtained in the approximation between both solutions (1 min vs 15 min) performed in the test case study described above. Observe that the error oscillates between 0.25% and 0.4% with respect to the power flow solution, representing approximately a Gaussian error in accordance to the histograms also shown in this Figure. The error is calculated by comparison of estimated solution (15 min) with respect to the solution from power flow (1min) but taking one point every 15 steps (15 min comparison).



a) Error and histogram for the estimation of voltage in bus 2



b) Error and histogram for the estimation of voltage in bus 6

Figure 12 – Errors between 15 min DSSE versus 1 min solution for the 12 buses MV test case

Finally, Figure 13 is showing the waveforms of active and reactive power again comparing the 15 min average measurement with respect to 1 min power flow solution for the 12 buses MV test case. It could be concluded that despite the difference on time step, the state estimation is following approximately the profile of the lower time step, with a small error and similar wave feature.

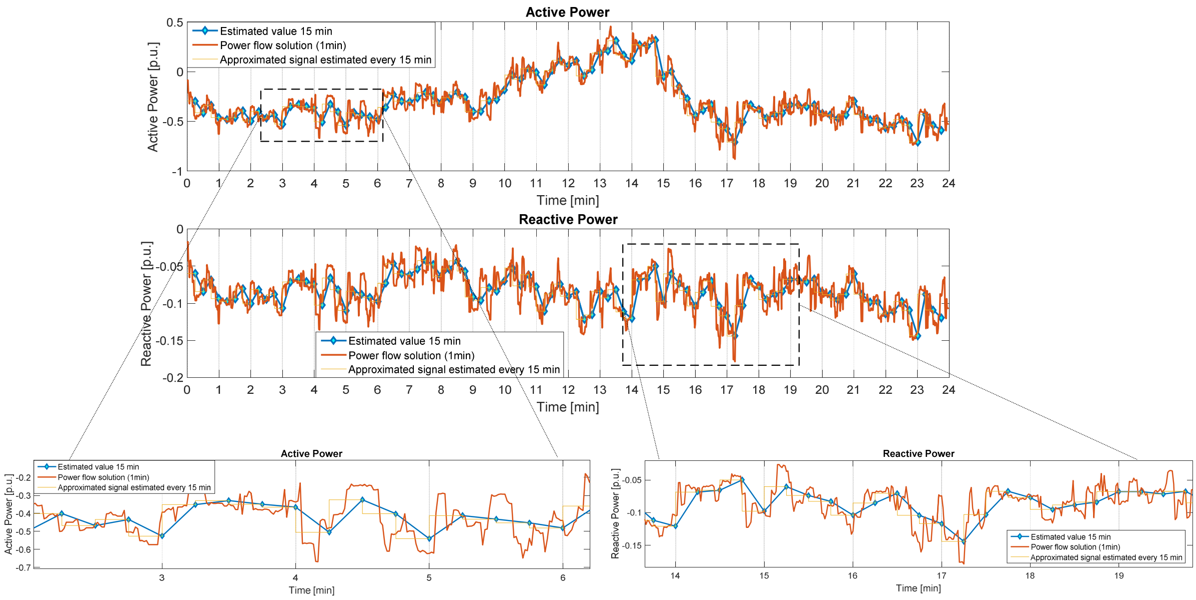


Figure 13 – Active and Reactive power at bus 3.
Comparison between 15 min DSSE with respect to 1 min solution for the 12 buses MV test grid

General State Estimation in Distribution Grids scheme

The characteristics of the DSSE analysis here described can be generalized via the flow chart of the Figure 14. The diagram is structuring the different levels which are followed to complete the procedure for estimating the distribution grid. Starting with the consideration that detailed Smart Meter data set with their respective aggregated measurements provides valuable information of the status of distribution grid, which is further taken as the basis of the SE algorithm, the main task of the concept is to establish a predictive relationship between different days ($d-1, d-2, \ldots, d-n$) in order to predict eventually the power network behavior in terms of loads, voltages, etc.

On the other side, taking advantage on the detailed modeling of prosumers that possess stochastic characteristics through DPG.sim, a realistic DSSE analysis is executed. This allows a variety of estimation performance analysis, i.e. by comparing the sensitivity of the estimation error with respect to different sampling times for power and voltage profiles.

The here presented state-estimation comparison methodology is systematic and may be adapted under any distribution network topology to either MV grids or LV grids analysis.

Figure 14 – *Predictive state estimation concept using Smart Metering data*

Limitations of State-Estimation Schemes

As a final note to this chapter: for state-estimation techniques to be successfully applied to distribution grids, the system must have sufficient measurement redundancy, i.e. more measurement data sets than grid variables that shall be estimated. This is not always possible, thus making it necessary to incorporate pseudo-measurements in order to fulfill the redundancy requirements of state estimation schemes [Singh et al. 2010]. Observability techniques are very useful to improve the state estimation of the distribution grid [Angioni et al. 2016]. The deployment of Smart Meters improves the (real-time) observability of distribution grids, thereby reducing the need for pseudo-measurements and improving the accuracy of state-estimation schemes.

State-Estimation Results for IWB *Kleinhüningeranlage* pilot grid

The above presented SE scheme has also been applied successfully to the Kleinhüningeranlage pilot grid, voltage errors are well-below 0.5% for the given case (Figure 15), although only the Smart Meter measurement data from two buses (out of 13 in total) were used for the state-estimation.

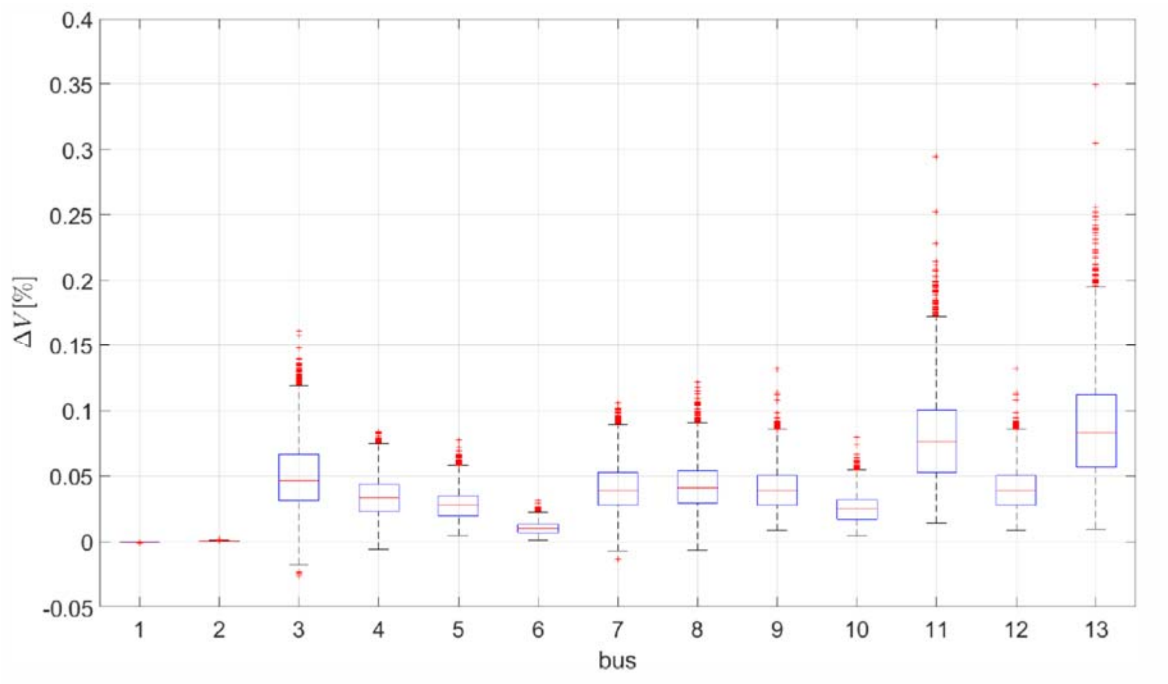


Figure 15 – *IWB Kleinhüningeranlage: instantaneous voltage error per bus for {P, Q, V} measurements at 2 buses*

2.3.2. Modeling and Simulation of *Kleinhüningeranlage* pilot grid region

Distribution grid setup

An urban residential neighborhood in the north-western edge of the City of Basel, the *Kleinhüningeranlage* (Figure 16), is used as a practical case study for employing Smart Metering data as well as other available operational data, i.e. measurements from transformers (secondary substations), commercial customers and distributed generation units (PV) for detailed time-series simulations and analyses of the low-voltage distribution grid.

Such spatially fine-grained time-series grid simulations, obtained via successive power flow calculations, allow a detailed analysis of occurring voltage profiles and line loadings down to the street level – or, if existing grid models (Figure 17) are sufficiently detailed, even down to individual houses (so-called *house connection points* [Hausanschlusspunkte]) and households.

Figure 16 – *Kleinhüningeranlage pilot grid region in City of Basel (system boundaries indicated in red)*

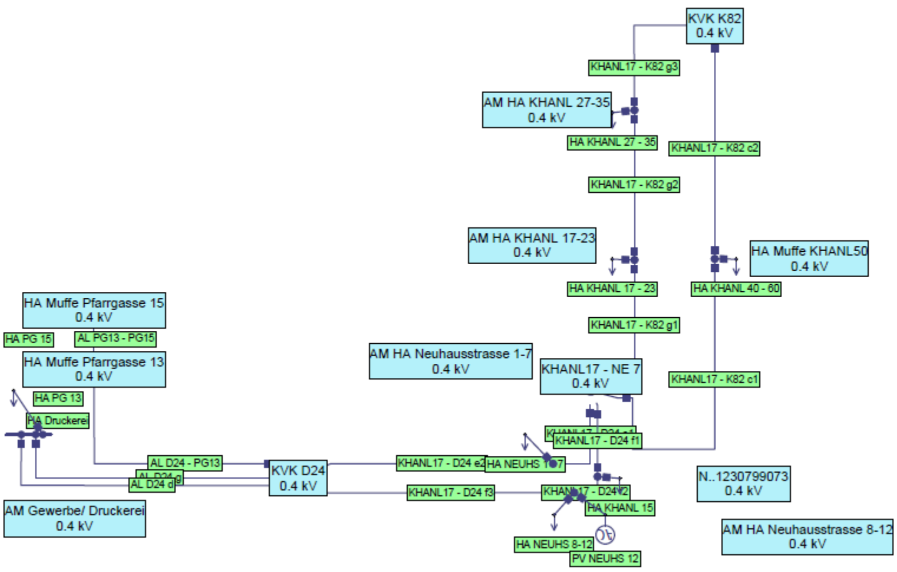


Figure 17 – *Grid Topology of Kleinhüningeranlage pilot grid region*

The *Kleinhüningeranlage* pilot grid region consists of about 200 households with a peak load of 115 kW_p, one commercial customer (a printing shop) with a peak load of 65 kW_p and – at the time of studying – one PV unit with a peak production of 10 kW_p that exhibits a particular production profile due to seasonal shading effects (Figure 18).

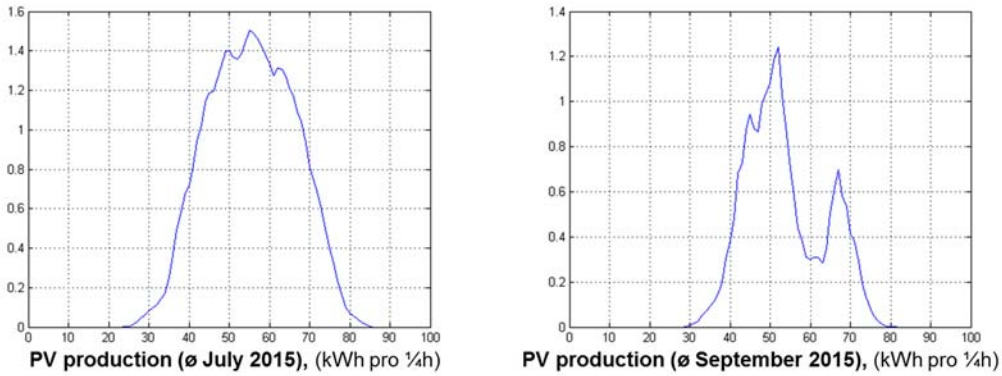


Figure 18 – *Kleinhüningeranlage PV unit (showing seasonal shading effects in autumn months)*

Distribution grid simulations in DPG.sim

The existing grid model of the pilot region was integrated into the simulation platform DPG.sim together with transformer time-series (voltage profiles) as well as load demand and power production profiles from the residential and commercial customers inside the system boundaries (Figure 16). Several distribution grid scenarios were created in order to test and analyze the influence of, for instance, a large-scale PV deployment and the simulation quality when using forecasted Smart Metering data in order to circumvent the typical issue of missing real-time measurement data.

In the following, two scenarios are presented in more detail: a grid scenario that reflects to the best knowledge the current situation of the *Kleinhüningeranlage* distribution grid, *Base Case scenario*, and a *PV Deployment scenario* with 13 large PV units with altogether 130 kW_p, instead of the single PV unit (10 kW_p) that is currently installed in the area. Plausibility of such a large-scale PV deployment in the area *Kleinhüningeranlage* has been checked via the solar cadaster provided by the statistical office of the City of Basel (Figure 19).

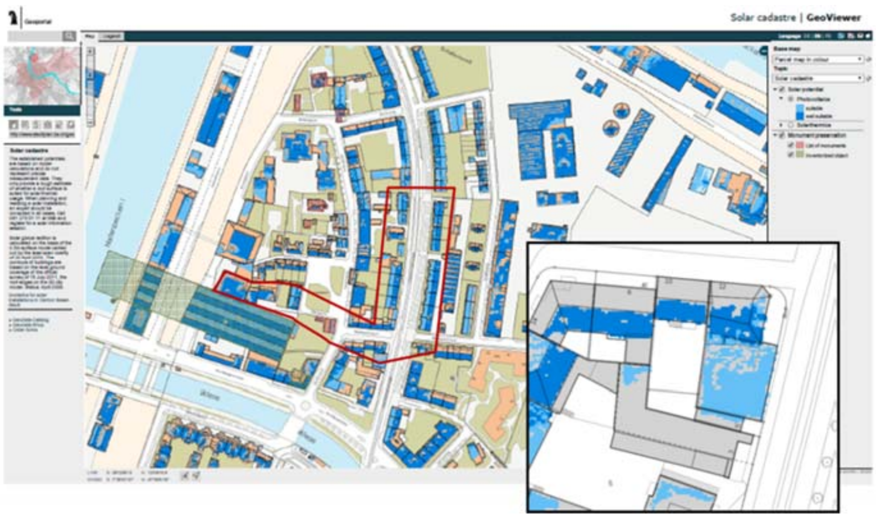


Figure 19 – *Solar Cadaster showing the potential for roof-top PV*

An illustration of a grid simulation run of the *Kleinhüningeranlage* pilot grid region, showing some key simulation parameters such as bus voltages and active power losses, is given below by Figure 20.

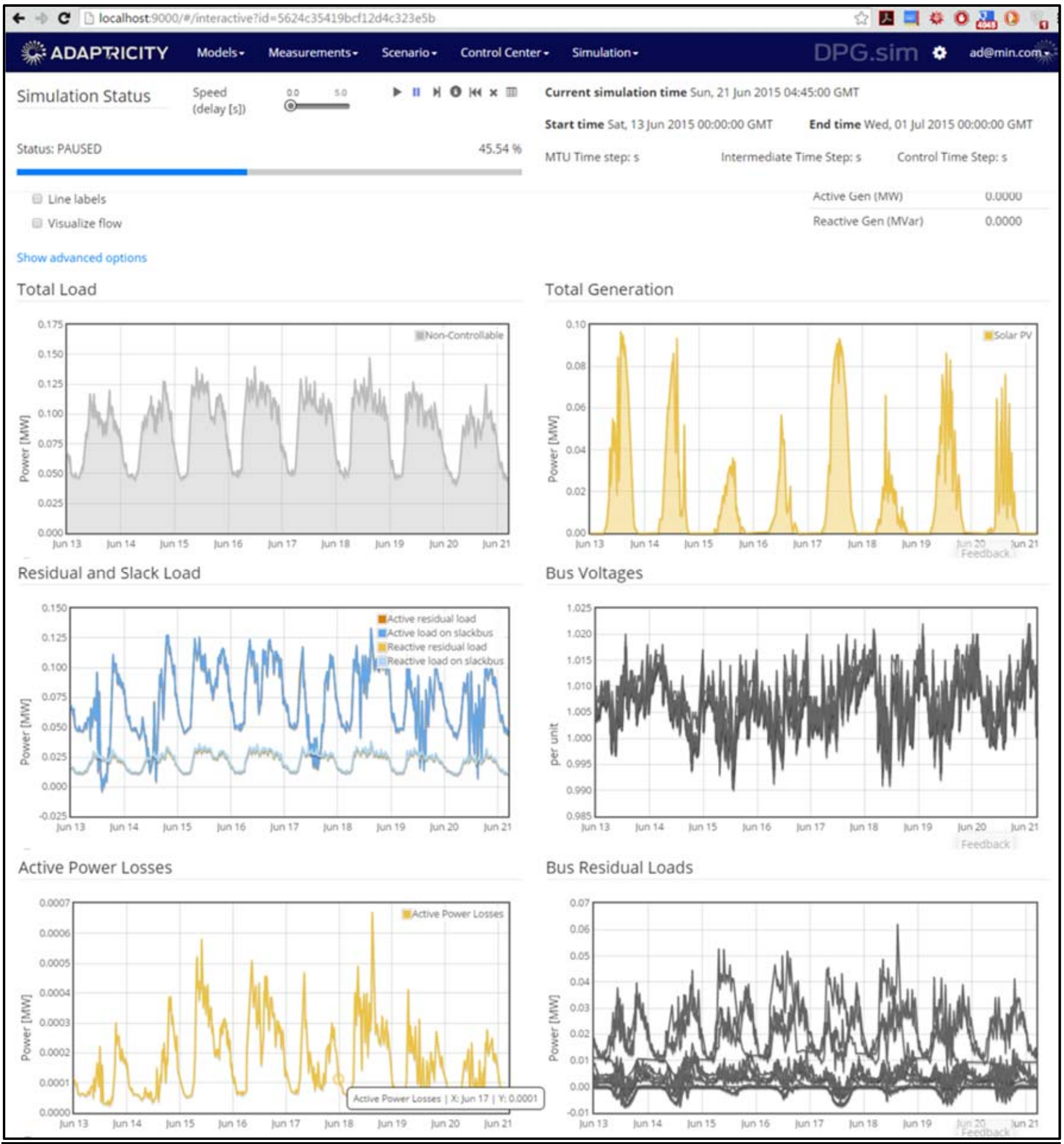


Figure 20 – *Simulation of Kleinhüningeranlage in DPG.sim*

A comparison of the (low voltage) line loadings for the Base Case and the PV Deployment scenarios is given by Figure 21: The line loading in the PV scenario is – at least globally – somewhat lower than in the Base Case. This is due to the very homogenous PV deployment. The locally produced PV power can serve the local electricity consumption, thereby lowering to a certain degree the electricity transport inside the distribution grid.

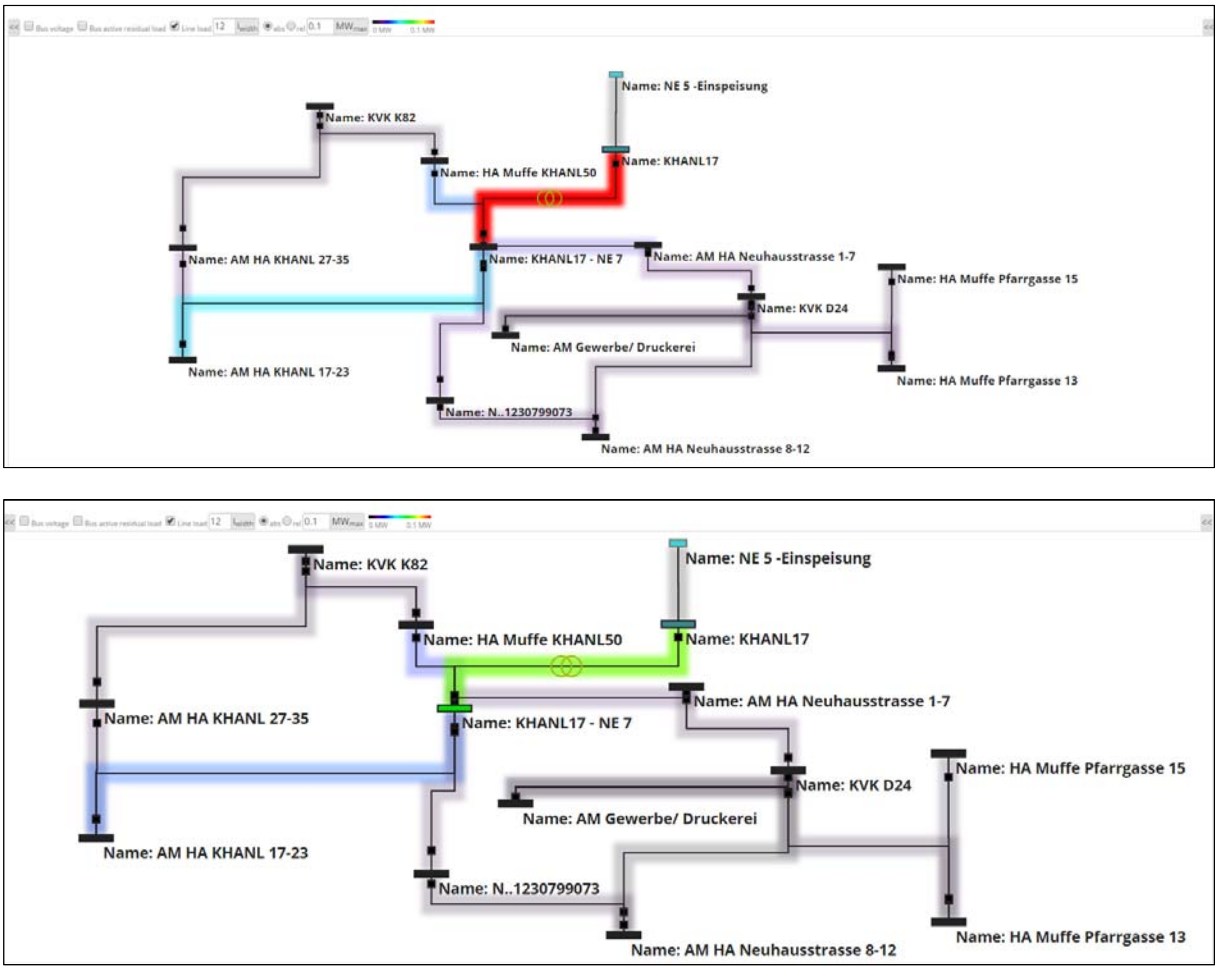


Figure 21 – Snapshot of *line loading situation in the IWB Kleinhüningen Grid for the Base Case (top) and the PV Deployment case (bottom)*

Figure 22 shows the differences between the total load, i.e. the *nominal* load demand of all residential and commercial customers, and the residual load, i.e. the *effective* load demand when accounting for the significant local PV production inside the pilot grid region. This clear difference validates the above made conjecture about the differences in load flow patterns between the two grid scenarios.

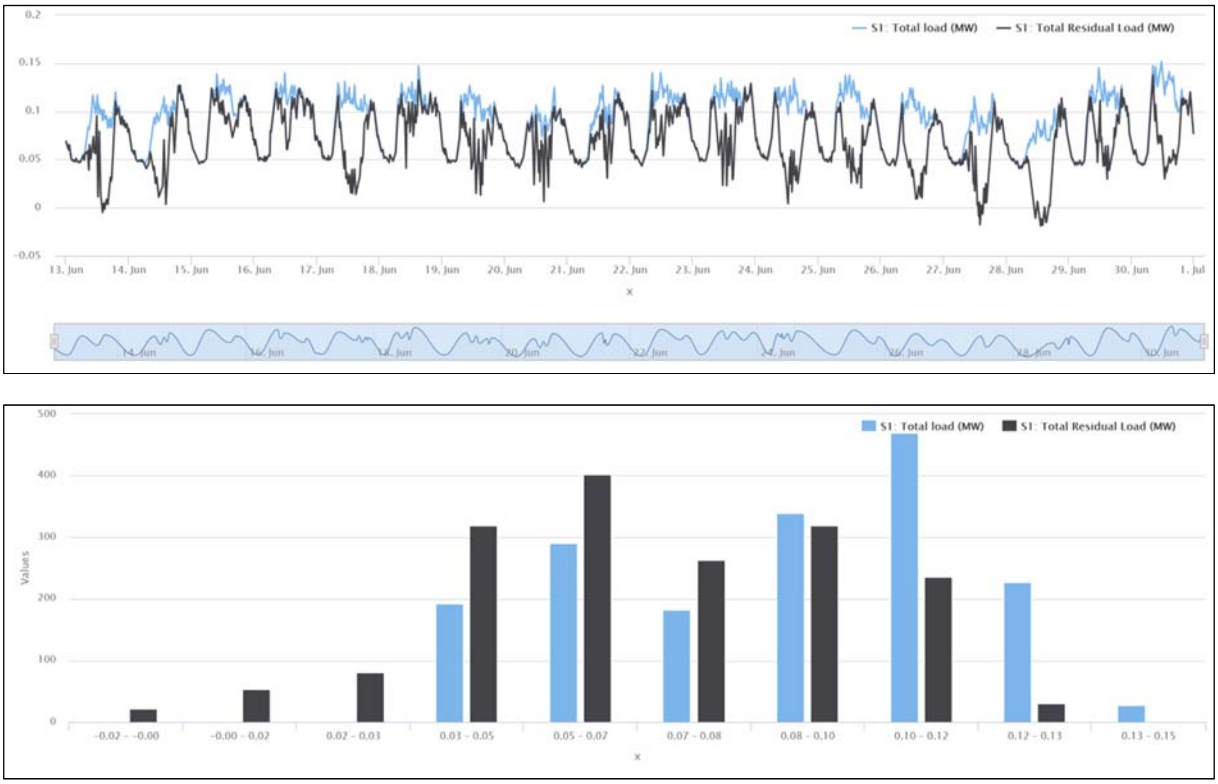


Figure 22 – *PV Deployment scenario exhibits clear difference between total and residual load demand (blue: total load [MW] and black: residual load [MW])*

Furthermore, it makes sense to use forecasted time series from smart metering devices in order to predict the behavior of active distribution grids. A power flow scenario has been simulated in DPG.sim (Figure 23) for the grid pilot region of *Kleinhüningeranlage* and shows that despite the fact that the prediction in blue fails to match all small fluctuations measured by individual smart meters and tends to slightly shrink the original time series, it is still accurate enough not to alter significantly bus voltages (Figure 24) and power line flows (Figure 25).

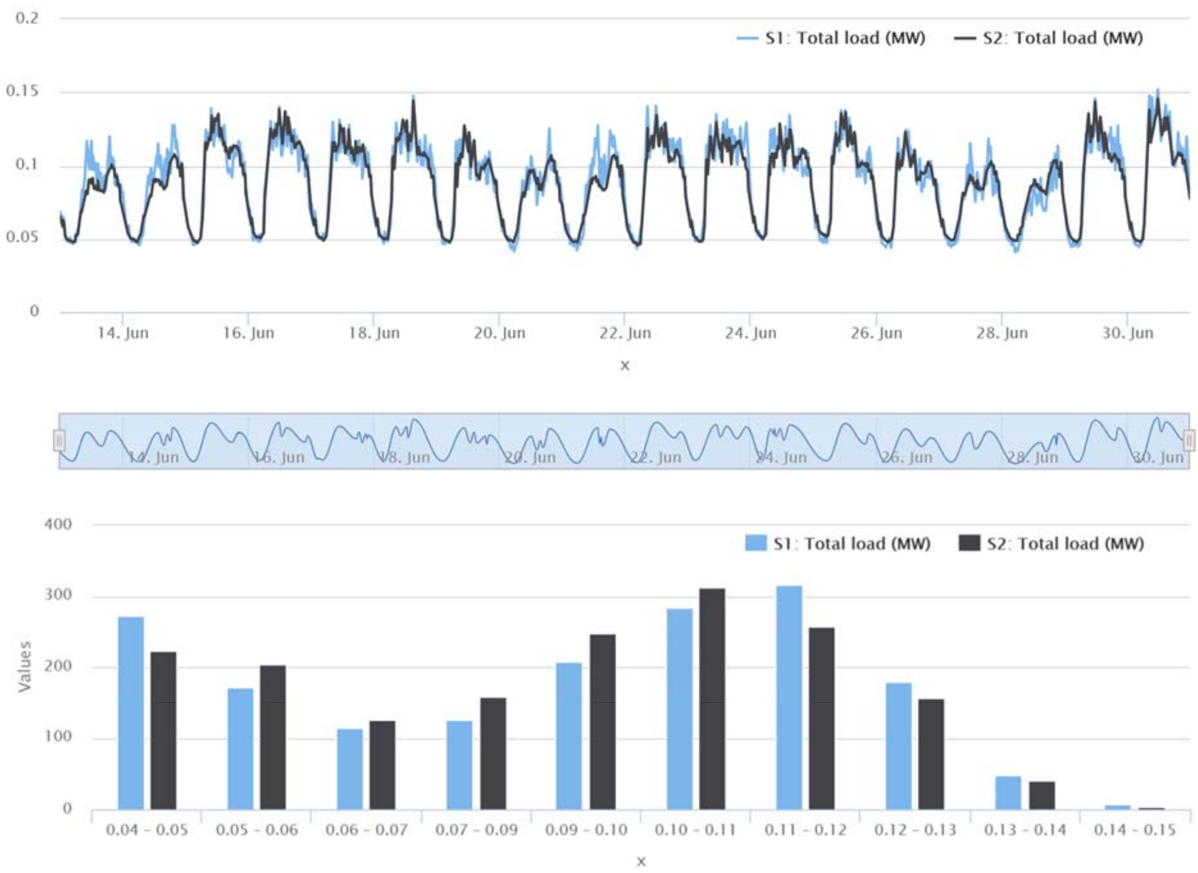


Figure 23 – *Active Power Flows at a Transformer & Histogram with Measured (blue) and Predicted (black) Data*

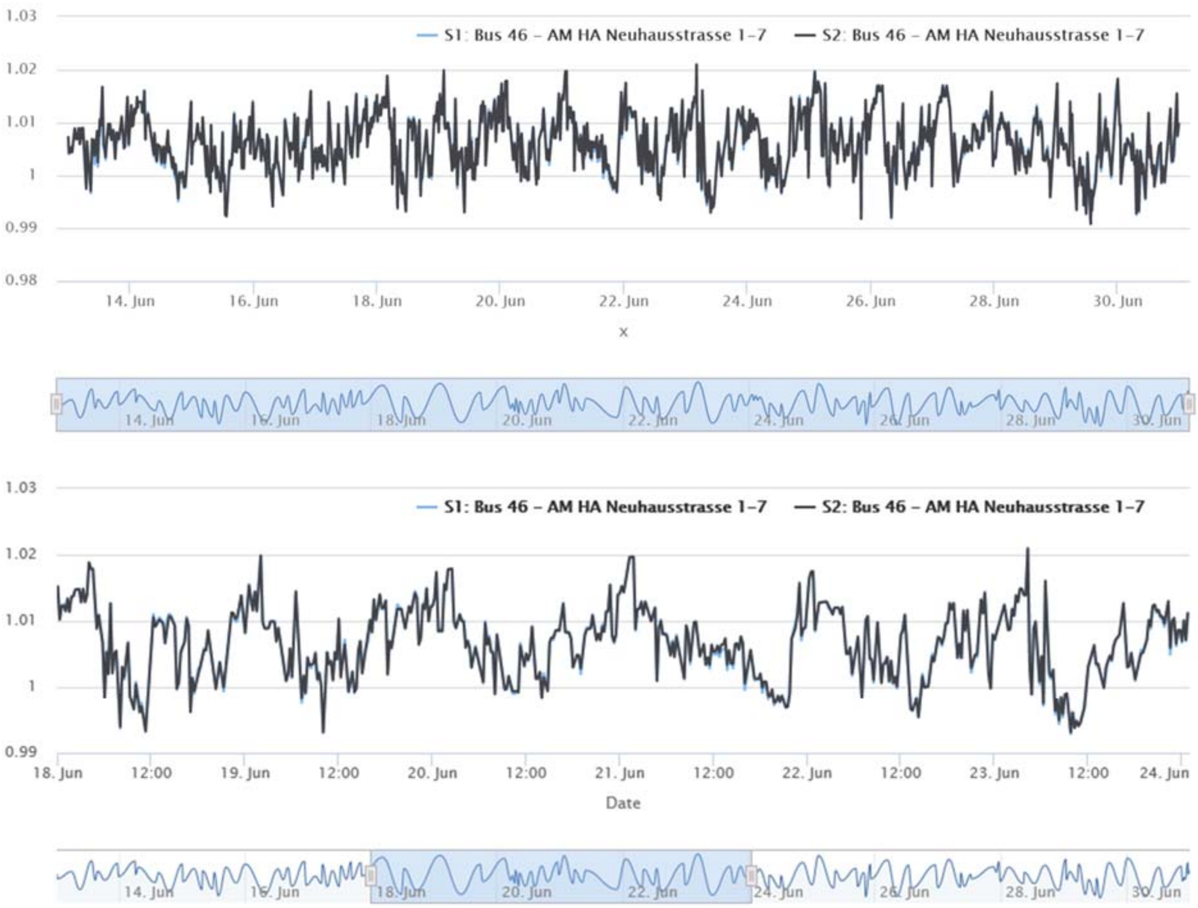


Figure 24 – *Bus Voltages with Measured (blue) and Predicted (black) Data*

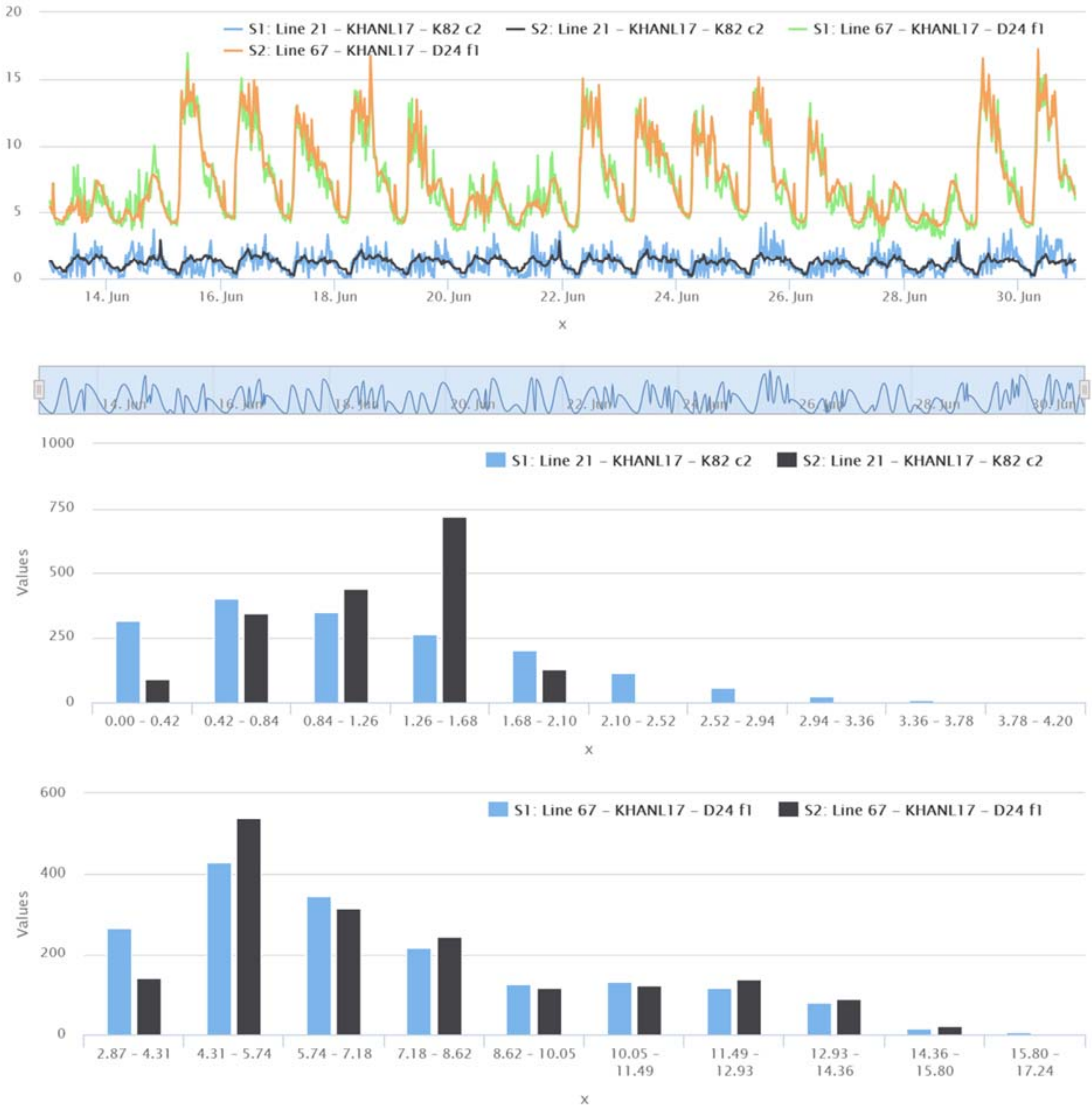


Figure 25 – *Active Power Line Flows & Histogram with Measured (blue) and Predicted (black) Data*

2.4. Work Package ETH-G2

2.4.1. Investigation of suitable methods for load profile generation and predictions

An important part in understanding the ability of Smart Meter data to describe the occurrences in the grid is the accuracy of representing the time evolution of the actual load and generation. It is clear that by aggregating the instantaneous values on 15-minute intervals, information is lost. Although the energy content of both the instantaneous values and the Smart Metering data is the same by definition, the Smart Meter data profile is smoothed considerably.

The load curve in Figure 26, upper plot, has been obtained by using the Load Profile Generator [Pflugrad 2014] from TU Chemnitz, Germany, using 100 randomly parameterized households of different type. This tool can be regarded as a realistic representation of the occurrences in residential households that lead to the well-known residential load profile shape.

In Figure 26, lower plot, the 15-minute energy aggregates of the load curve which have been rescaled to average power values (4 * kWh/15 min = kWh/h) are shown. The upper plot is considerably spikier and exhibits a peak load that is about 20% higher than the corresponding Smart Metering data.

From this result, it becomes clear that the smoothing and peak-reducing effect of the 15-minute aggregation procedures of Smart Meters has to be taken into account when Smart Metering data shall be used for accurate grid simulation and analysis purposes.

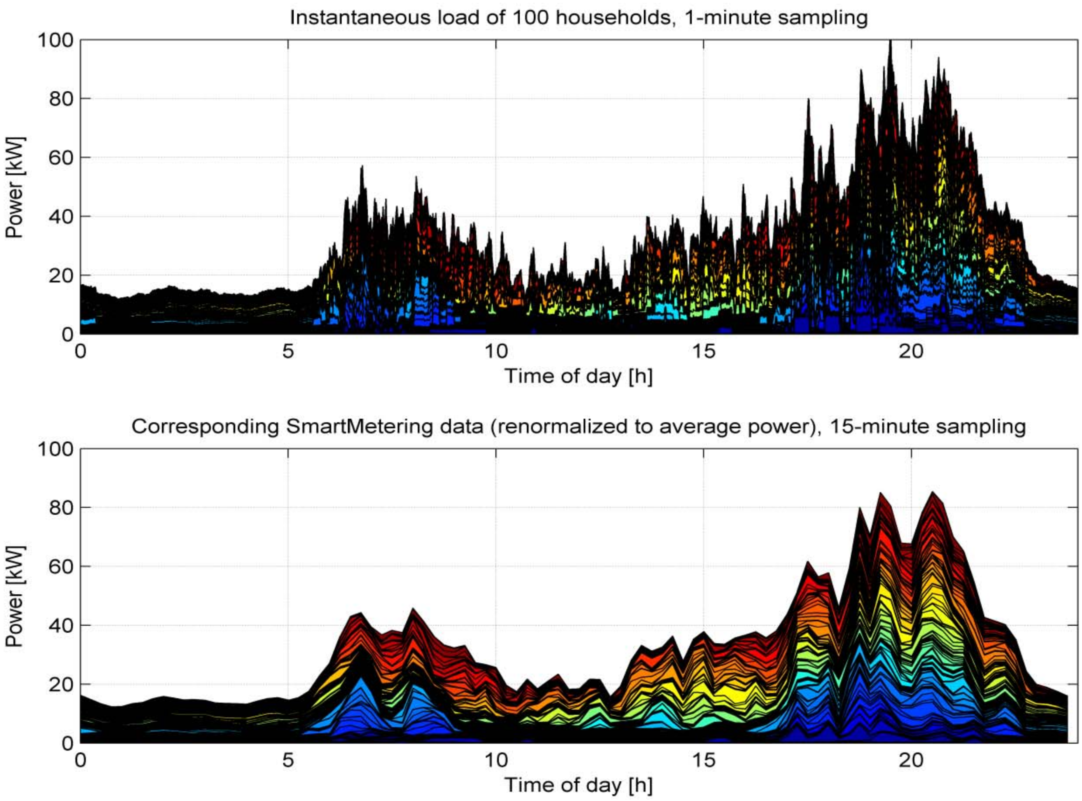


Figure 26 – *Load profile (1-minute sampling) and resulting 15-minute Smart Meter data*

2.4.2. Smart Meter data analytics

The following section focuses on the analysis and exploitation of the operational distribution grid data that has been provided in an anonymized fashion by IWB. The project tasks ranged from the analysis of the available Smart Metering data sets via data mining and data clustering techniques, the refinement and application of the developed Smart Meter state-estimation concept as well as the exploitation of IWB's *real* operational data for modeling and simulation of the distribution grid pilot region *Kleinhüningeranlage* down to the geographical granularity of the street and household level.

Overview publications of the performed Smart Meter data analytics tasks and obtained results have recently been published in [Zufferey2016] in an international conference and in [Vogel2016] on the national level (SEV Bulletin).

Smart Meter Data Set

The available smart meter datasets from the City of Basel, coming from over 40'000 households, almost 1'000 commercial customers as well as around 400 PV systems between April 2014 and September 2015, have been preprocessed and analyzed using so-called Big Data technologies that allow handling and processing of very large datasets. **All household datasets have been anonymized by IWB.** The general approach to this analysis includes different stages (Figure 27):

- Data preprocessing, which consists of following steps:
    - Data integration (formatting, checking and validation),
    - Missing values imputation, i.e. extrapolation of missing or false data points by using correction algorithms based on plausible assumptions,
    - Features extraction, which prepares the data for subsequent clustering and forecasting methods;
- Grouping or clustering of sub-groups of households based on similar and characteristic behavior in terms of load consumption;
- Interactive visualization-based data mining which gives valuable insights into the distribution of load and PV production in the City of Basel;
- Time-series forecasting at different aggregation levels, e.g. hour-ahead and day-ahead consumption predictions can be accomplished.

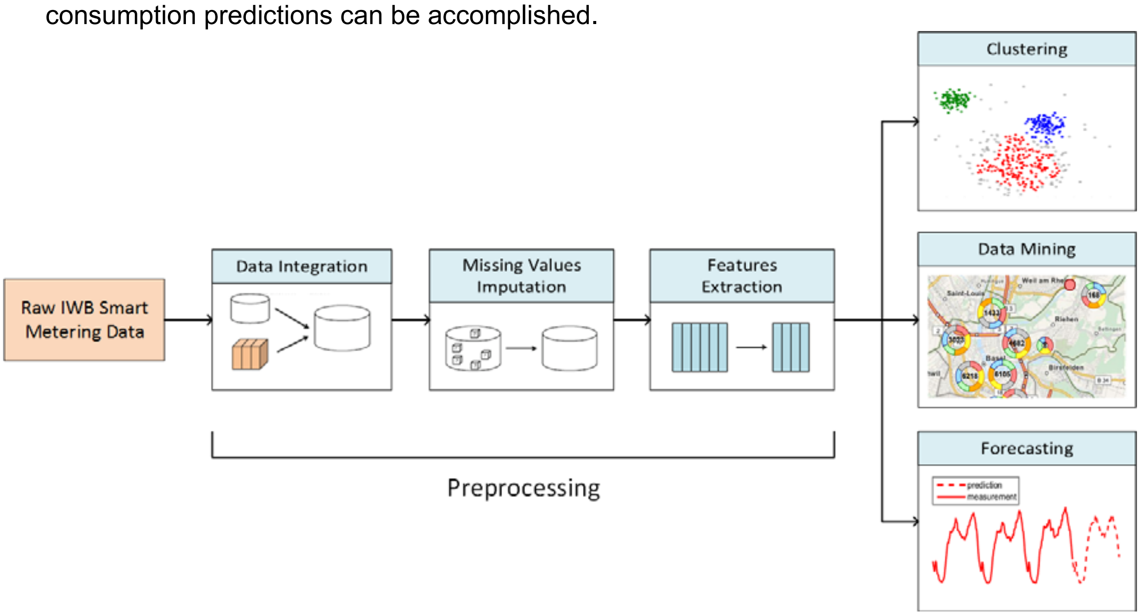


Figure 27 – *Smart Meter Data Analytics Approach*

Data Preprocessing

The data provided by IWB consists of many files, i.e. one per day in the case of households and one per month for industries and PV systems. The first step of a comprehensive data analytics consists thus of integrating all relevant sources together to obtain one single active power time series per smart meter for the whole analysis period, i.e. from April 2014 until August 2015. Smart meters whose time series is not available every day, respectively every month, have been discarded. Furthermore, exogenous information like meteorological data from the weather station of Binningen provided by MeteoSwiss has also been considered.

In addition, a simple anomaly detection has been performed to eliminate amongst others duplicates, households with a mean daily consumption lower than 100 Wh (Figure 28), industries that consume less than 100 kWh per month or with a share of zero values higher than 20%, and PV systems exhibiting a night production. This first stage leads to a total of:

- 26’337 valid household time series
- 808 valid commercial customer time series
- 351 valid PV system time series

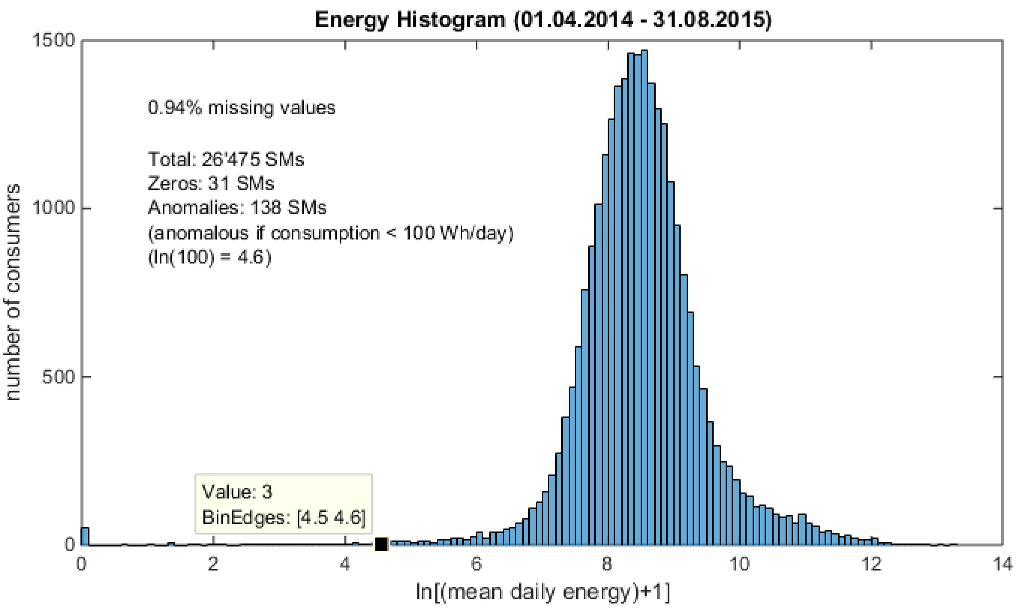


Figure 28 – *Energy Histogram of measured Households (energy readings in Watt-hours)*

Concerning household time-series, it appears that 0.94% of active power values are missing in the time-period of April 2014 to August 2015. This has, however, no impact on customer billing as a separate dataset of cumulated energy consumption, here a separate data register, is used instead. There are two extreme cases of missing data:

- 28th August 2014: except the first 8 values (until 2am), data is missing for all Smart Meters
- 17th February 2015: except the first 4 values (until 1am), data is missing for the vast majority of Smart Meters

Since incomplete datasets can drastically decrease the performance of machine learning algorithms, a method called *“k-nearest neighbors”* has been used to substitute missing data points with plausible values of similar datasets. Indeed, for each incomplete time-series, the algorithm looks in the dataset for the *k nearest* time-series, i.e. the most similar time-series to the time-series with missing values, and then replaces every missing measurement points by the mean value from these *k neighbors*.

Note that all data preparation tasks have been implemented in *Java* and time-intensive computations such as missing values imputation and features extraction have been supported by the engine *"Apache Spark"*, which is suitable for large-scale parallel data processing.

Household Clustering

Although households have been anonymized, on can still extract useful information from their load profile and distinguish multiple types of consumers by means of a classical cluster analysis, which allows to find structure and regularities in otherwise non-categorized data (unsupervised learning) [Jain2010]. The first step is to obtain so-called features that serve as benchmarks for the group formation. An intuitive MATLAB-based graphical user interface (Figure 29) has therefore been implemented in order to select features and define clustering criteria. Amongst others, typical daily or weekly patterns, the average load, the consumption share of a certain period in the day, the week or the year as well as correlations with meteorological time series are available. In addition, it provides the flexibility to take into account only a part of the whole dataset by selecting desired hours, weekdays and months.

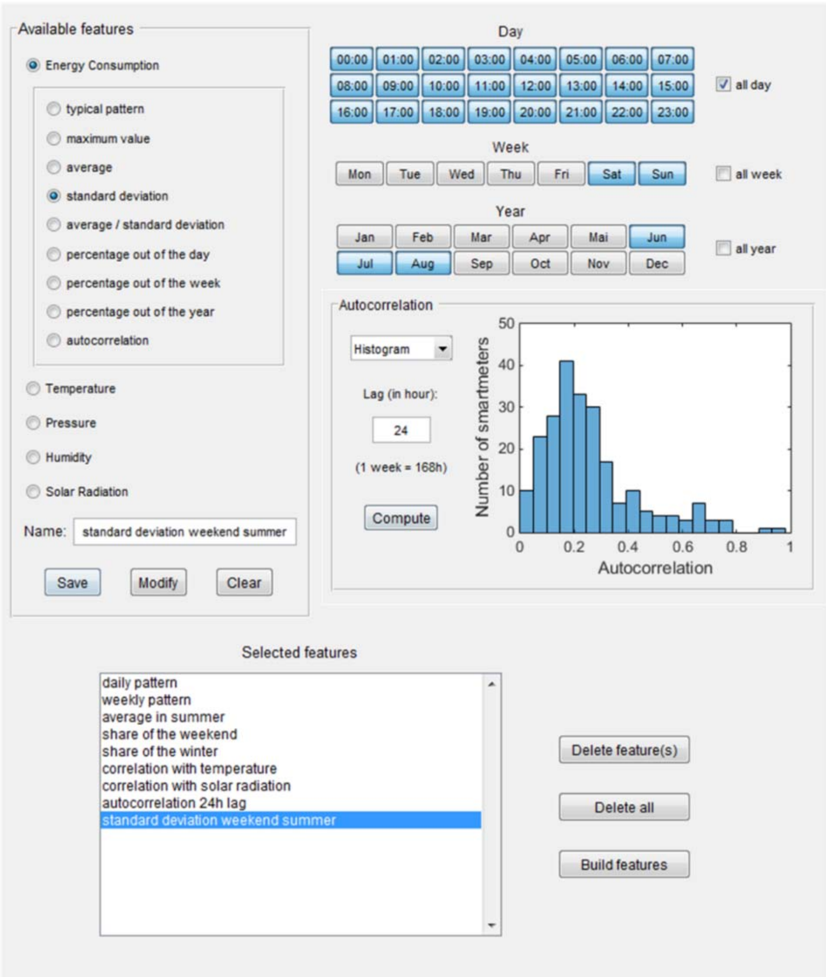


Figure 29 – *Graphical User Interface for Features Selection*

In a second step, extracted features are supplied to a popular clustering algorithm called *"K-Means"* implemented in the cloud computing software H_2O. The number of clusters depends on the data diversity, on selected features and on the clustering purpose.

Figure 30 illustrates three groups of consumers clustered according to their typical daily load profile. In this case, cluster 0 likely consists of residential loads whereas cluster 1 represents night owls and cluster 4 is typical for offices or shops. A further interesting example is the evolution of the typical load profile in relation with the mean power consumption, as shown by Figure 31.

On average, small consumers exhibit a typical household pattern even if their individual profile can be of any shape. Nevertheless, the more energy customers are consuming, the more rectangular the mean load profile looks like.

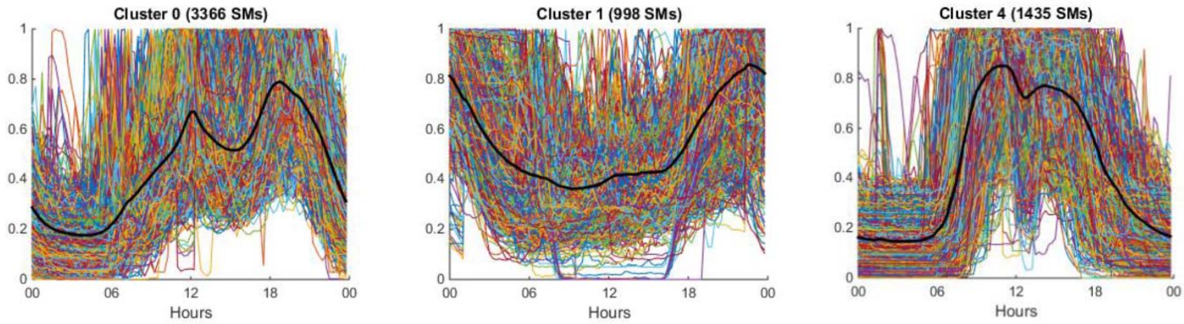


Figure 30 – *Clustering Outcome based on the Daily Load Profile*

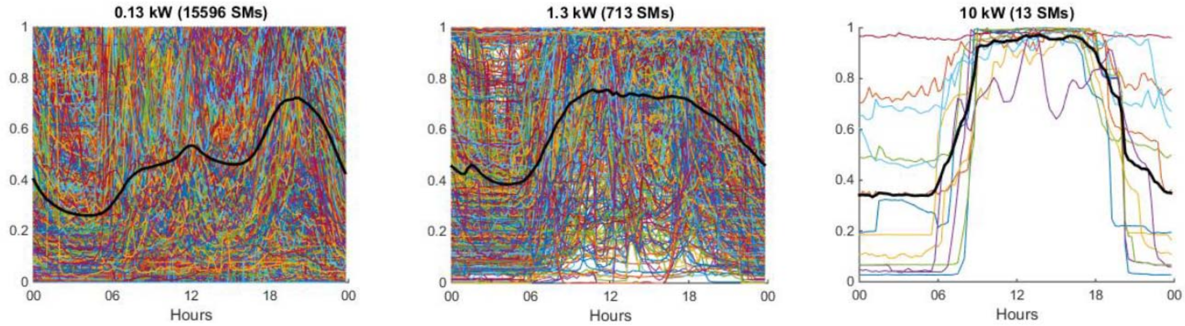


Figure 31 – *Clustering Outcome based on the Mean Power Consumption*

Visualization Tool

Since the information transmitted by IWB also includes post addresses, residential and industrial loads as well as PV systems have been visualized on the map of the City of Basel.

Note that only data concentrator's addresses are available for households. According to the chosen clustering features (daily share, weekly share, yearly share, mean consumption, correlation with temperature or weekly autocorrelation), one can observe the distribution of different types of consumers in each region where smart metering devices are installed (Figure 32). For instance, the city center stands out from other areas, showing a greater share of customers mainly active during working hours and whose load is highly correlated with the outside temperature. In addition, the visualization tool provides the user with valuable information (Figure 33) for each Smart Meter such as its clustering features, its physical address (due to the data anonymization this is here the address of the transformer station / data concentrator), the mean power, the day with highest energy, typical daily patterns (Figure 34) as well as energy shares amongst weekdays or seasons.

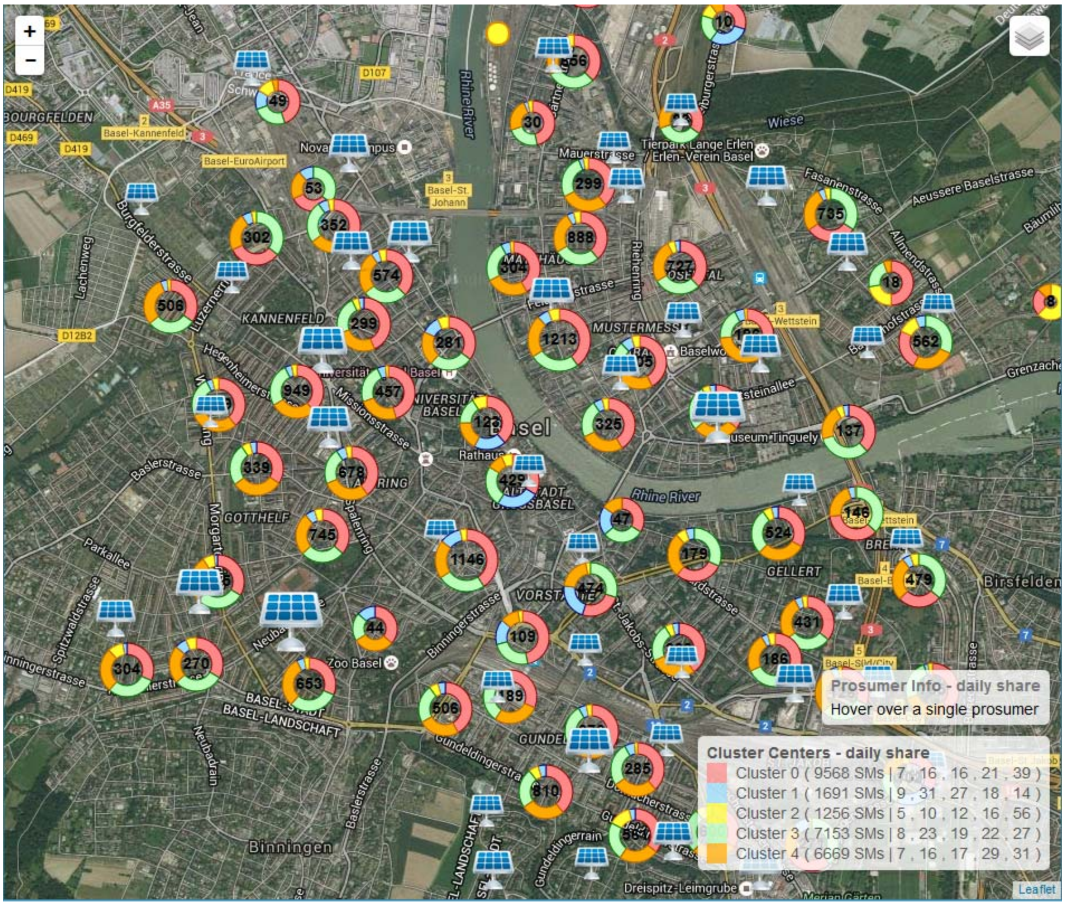


Figure 32 – *Visualization Example for Smart Metering Data*

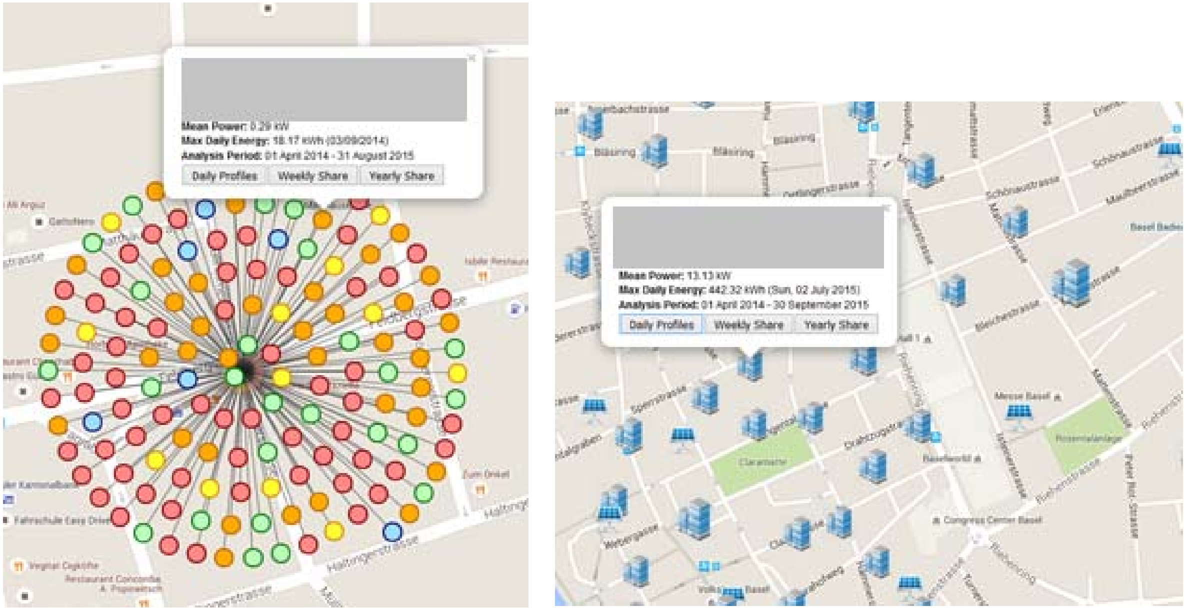


Figure 33 – *Information Boxes for a commercial and a residential customer*

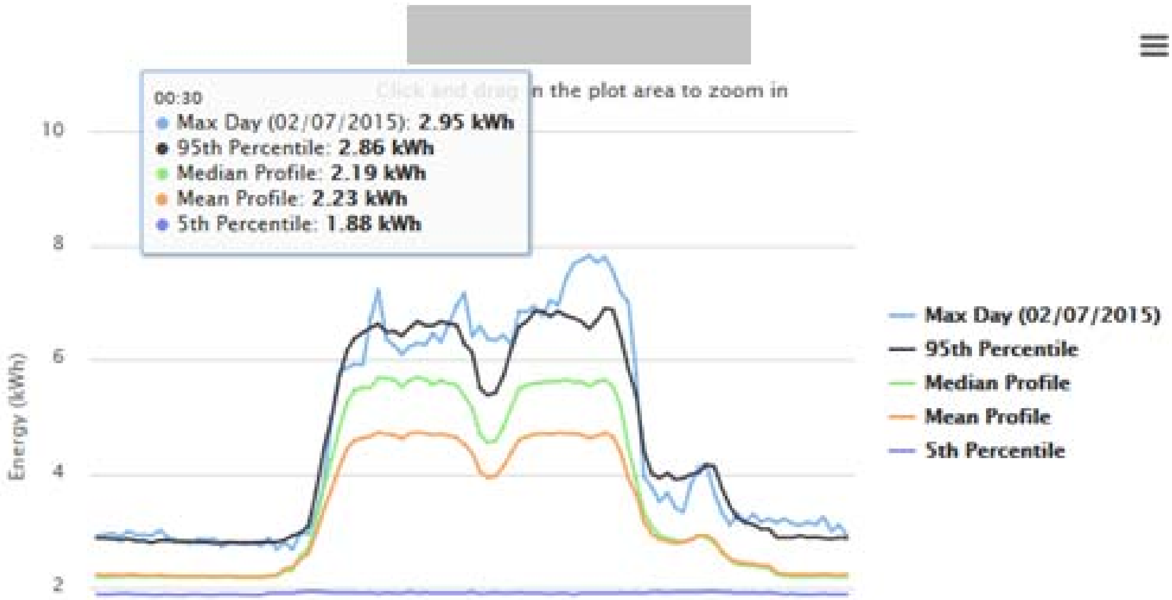


Figure 34 – *Typical Daily Pattern of a Specific Industrial Customer*

Load and PV Forecasting

Besides clustering, machine learning techniques are increasingly being used in the field of Short-Term Load Forecasting, i.e. the extrapolation of recently observed loads to the nearest future, which is nowadays primordial in power system planning and operation. A so-called “artificial neural network” (more precisely a “200-neuron Multilayer Perceptron”) implemented in H_2O has thus been trained with three feature types:

- Historical smart metering time series: mean consumption of previous day, energy value at the same time step one day or several days ahead …
- Exogenous variable: temperature forecast
- Human cycles or activities: hour, weekday, month, holiday …

Different levels of aggregation have been tested, going from the single consumer to the total load, and evaluated with the Mean Absolute Percentage Error (MAPE) which compares the absolute error with the measured value. Individual time series (Figure 35) are mostly very spiky and only the main tendency can be foreseen. However, the forecasting considerably improves for larger loads like commercial customers (Figure 36) or when multiple load profiles are aggregated, e.g. for the typical aggregation level of a transformer station (≈100 Smart Meters) or a sub-station (≈10’000 Smart Meters).

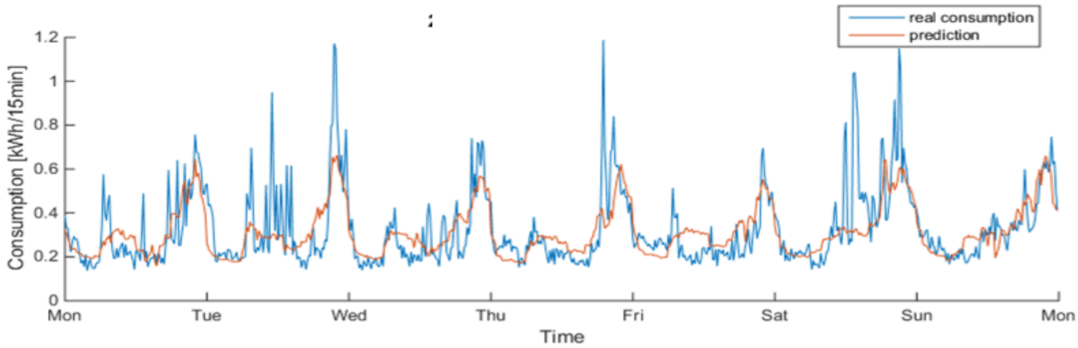


Figure 35 – *Day-Ahead Load Demand Forecast for a Single Household*

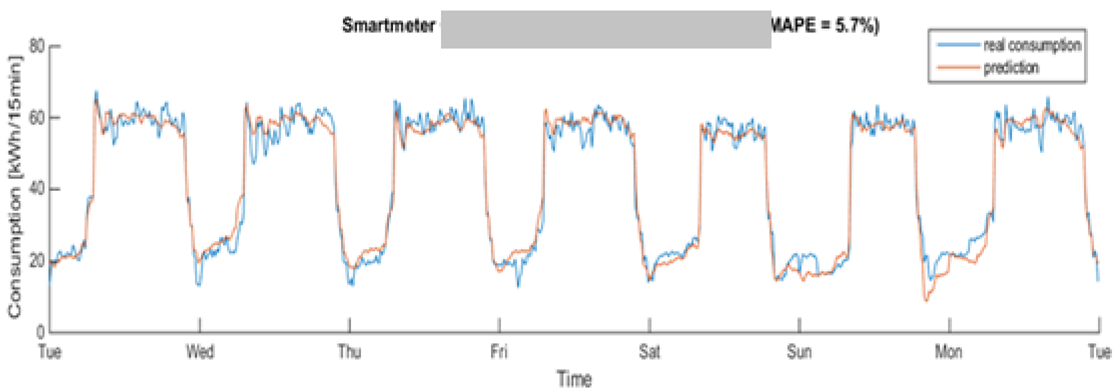


Figure 36 – *Forecasting Outcome for a Commercial Customer*

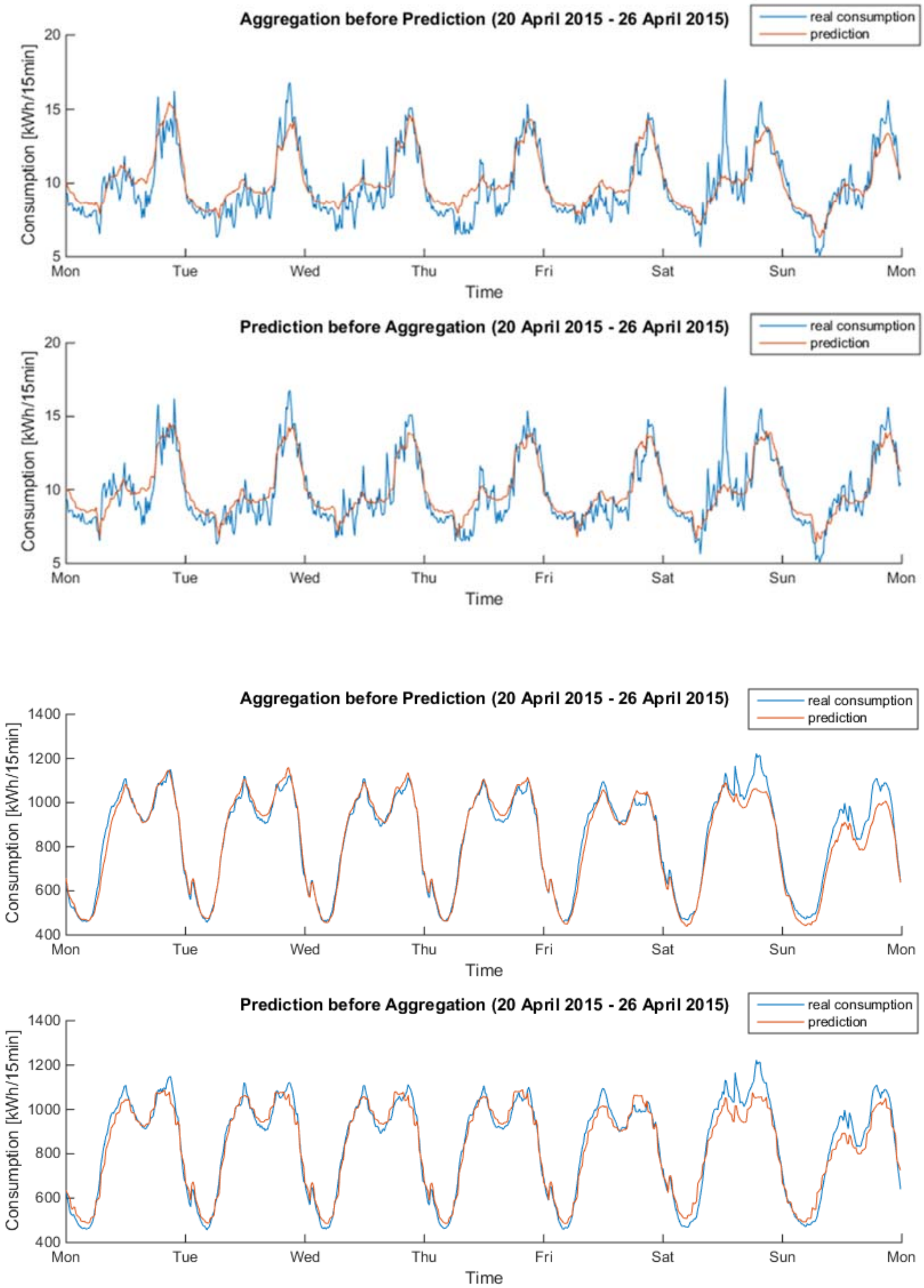


Figure 37 – *Day-Ahead Load Demand Forecast for Aggregations of Smart Meter Customers*
above: aggregation level of a transformer station (≈100 Smart Meters)
below: aggregation level of a sub-station (≈10'000 Smart Meters)

Aggregation-before-Prediction versus Prediction-before-Aggregation

Notice that there is no flagrant disparity in the algorithm performance if the aggregation occurs before or after the forecasting process whereas hour-ahead predictions appear to be slightly better than day-ahead predictions at a high level of aggregation (Figure 38).

In both cases, prediction accuracy increases (mostly) monotonically for larger aggregation sizes. Only in the case of the mean daily energy accuracy does an aggregation-before-prediction perform better than the other way round (Figure 38 – bottom-right plot).

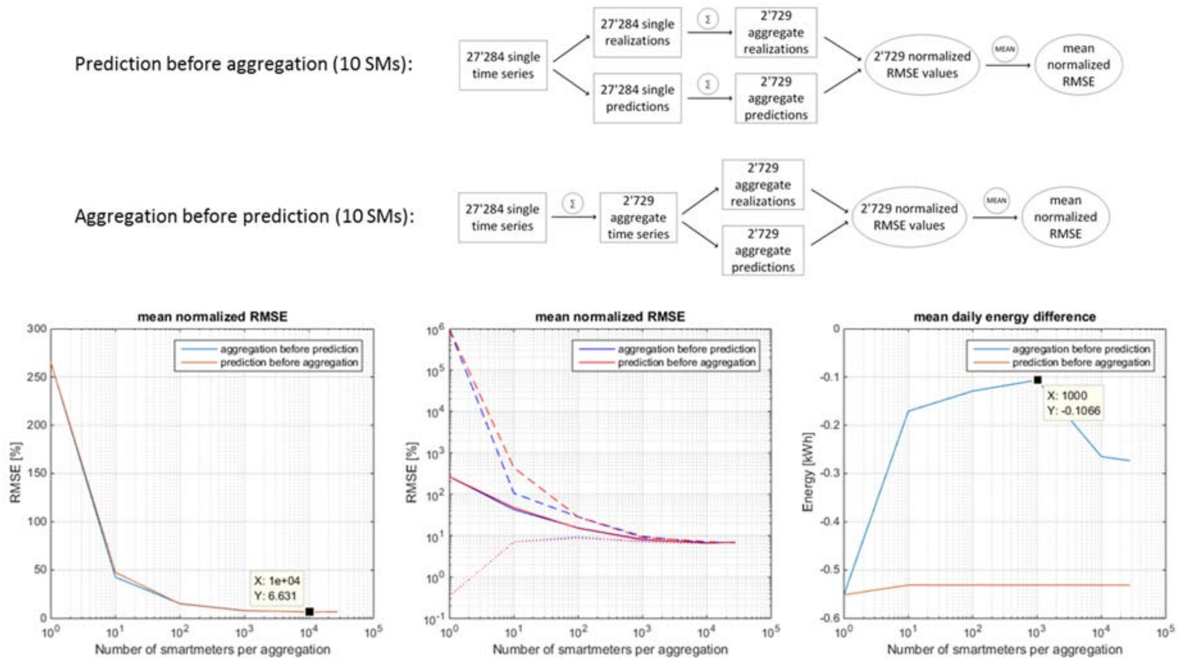


Figure 38 – *Performance Evaluation: Prediction-before-Aggregation vs. Aggregation-before-Prediction*

In Figure 39, it can be clearly seen that the household forecasting performance increases for Data Concentrators that collect consumers with a higher mean daily energy since their load profile is less spiky and more similar from one day or one week to another.

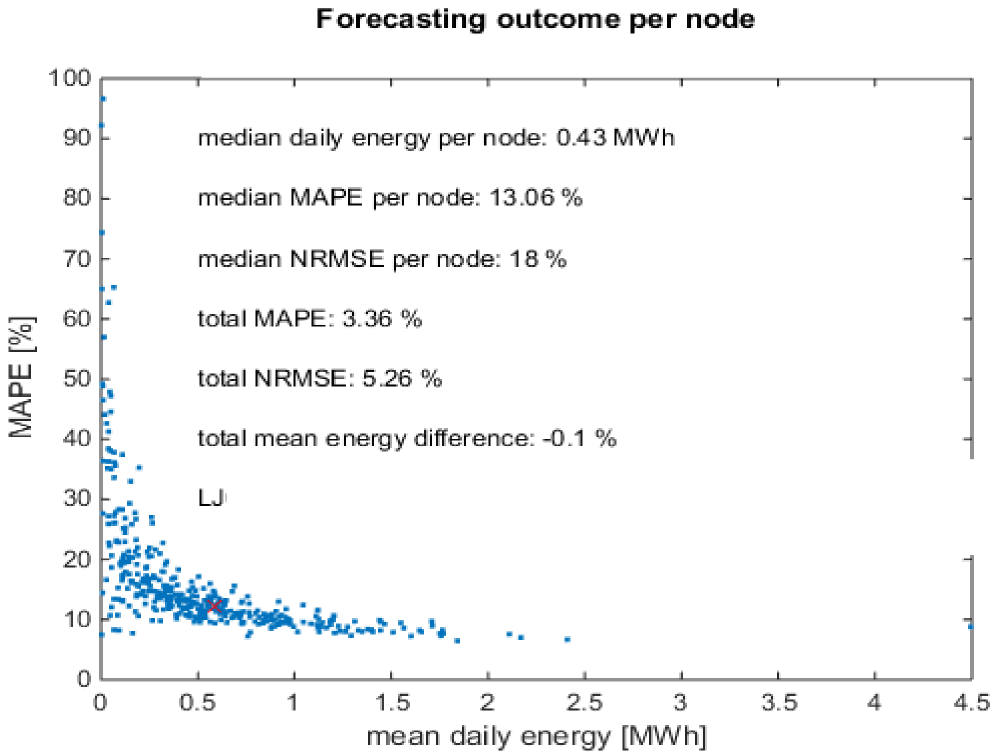


Figure 39 – *MAPE Dependency on Mean Daily Energy Consumption per data concentrator node)*

In addition to consumption predictions, the algorithm has been extended to PV forecasts and notably relies on the solar radiation. An example can be seen in Figure 40 with a relatively high performance for a single smart meter time series.

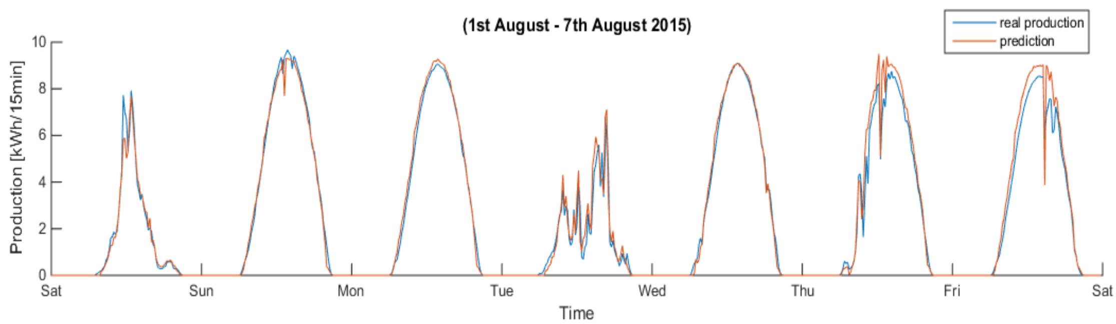


Figure 40 – *Day-Ahead Production Forecast for a PV System*

2.5. Work Package ETH-L1

As part of the Smart Metering project, several concepts for grid-friendly local control strategies for improving the grid operation in *stressed* distribution grids were developed and also employed in realistic Swiss distribution grids.

The modelling and analysis approach taken by means of the Smart Grid simulation platform *DPG.sim* is outlined below (Figure 41):

1. After an initial modeling stage, in which all available grid operational data is integrated into a so-called Base Case scenario.
2. Different Grid Scenarios with more or less severe operational problems, e.g. as caused by large-scale PV deployment, are then modeled and simulated.
3. In case of observed grid stress, e.g. voltage band violations as well as line or transformer overloading, different distribution grid upgrade options can be modeled and assessed.
4. The evaluation of a variety of grid upgrade options is performed via a technical assessment, e.g. analyzing energy losses or voltage band quality, as well as a financial evaluation, e.g. CAPEX and OPEX costs.

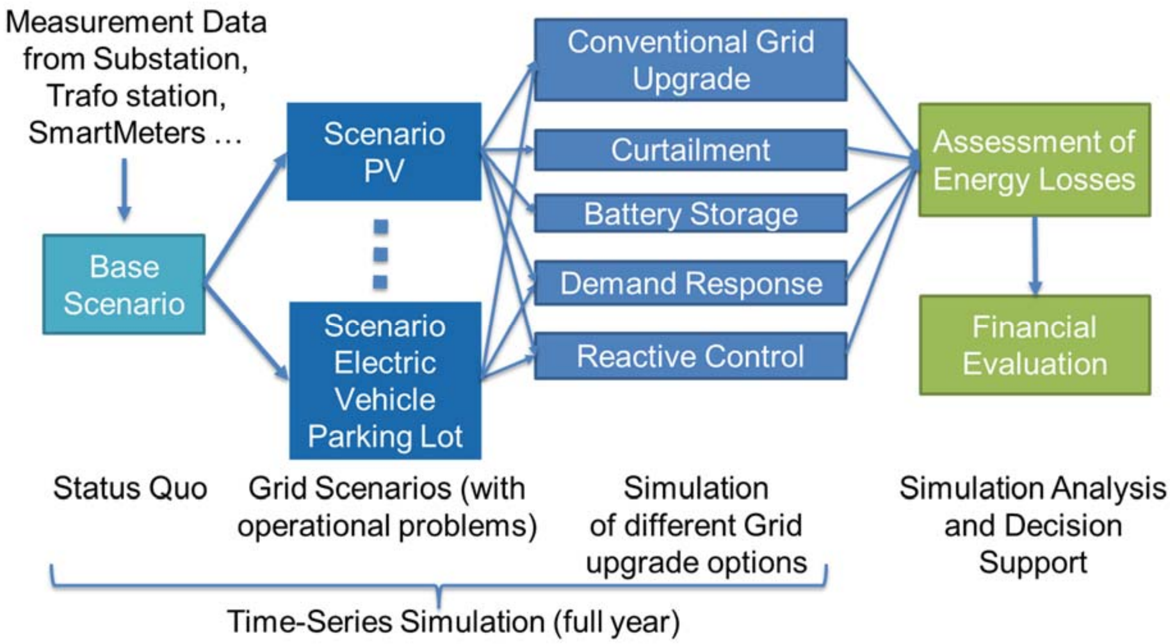


Figure 41 – *Smart Grid Upgrade Option Evaluation Methodology*

Base Case

Time-series simulations in DPG.sim for a Swiss distribution grid section are performed for a reference year. The voltage quality during grid operation is evaluated according to the ± 10% voltage band criterion, as stipulated in EN 50160.

Based on the measured, reconstructed, and estimated grid data time-series, the following grid situation arises for a representative week in June (Figure 42). While the steady load demand profile of the data center makes up the largest share of overall load demand, the load profiles of the three transformer stations serving the mixed commercial and residential area exhibit a typical daily load pattern with pronounced evening hour peaks.

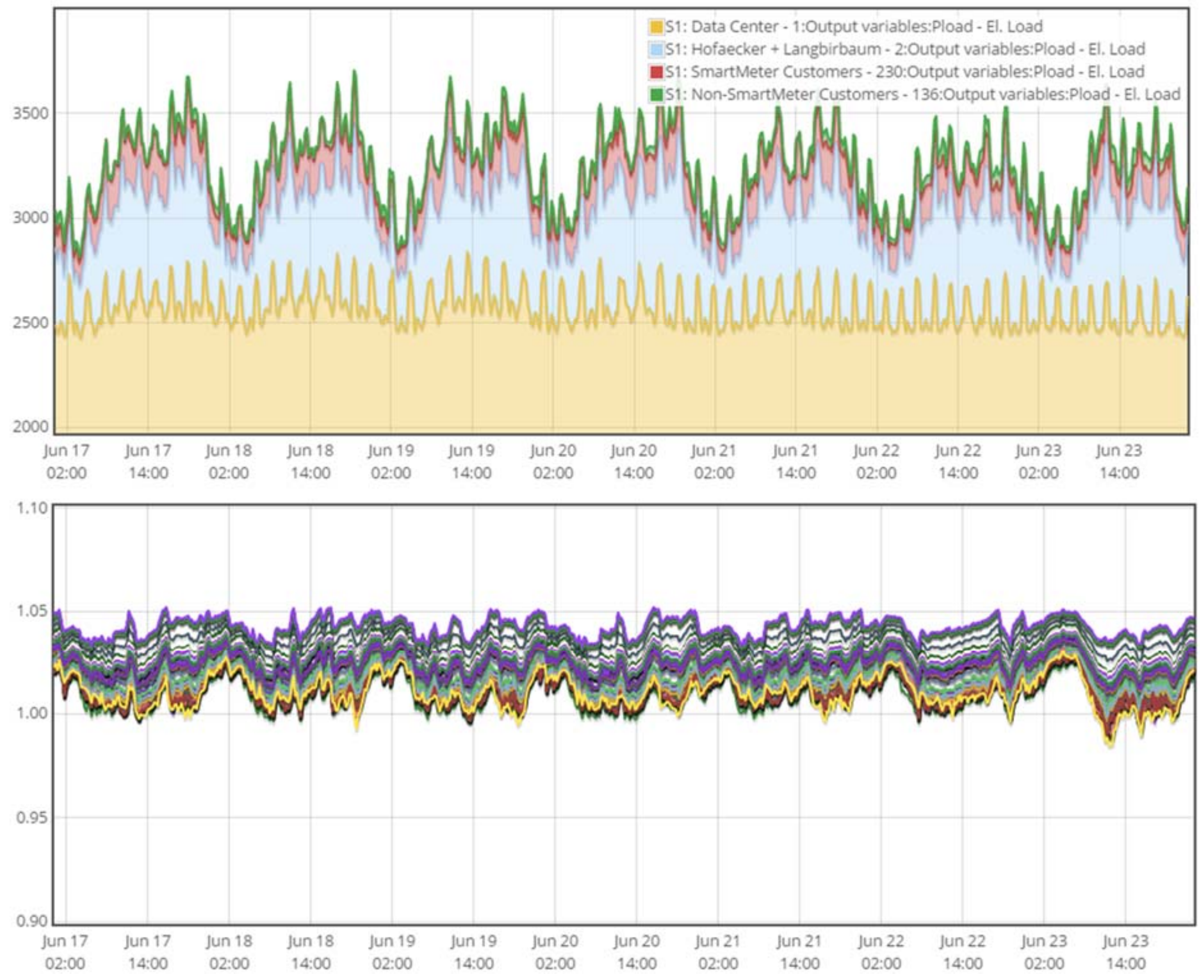


Figure 42 – *Base Case – Grid Situation (one summer week in June)*
above: Load consumption medium-voltage grid level, below: Voltage band evolution (no violations)

Grid Scenario – PV Deployment

A relatively sizeable PV unit (≈280kW), compared to the peak load demand of the local transformer station (less than 400 kW), has significant impacts on the LV grid. A significant voltage level rise (above 1.10 pu) as well as higher line loadings are induced by the PV power in-feed on sunny days as shown for a representative summer week (Figure 43). This leads to a violation of the permissible voltage corridor of ± 10%. The loading level of the Luberzen transformer is decreased on average but subjected to more fluctuations due to the PV in-feed.

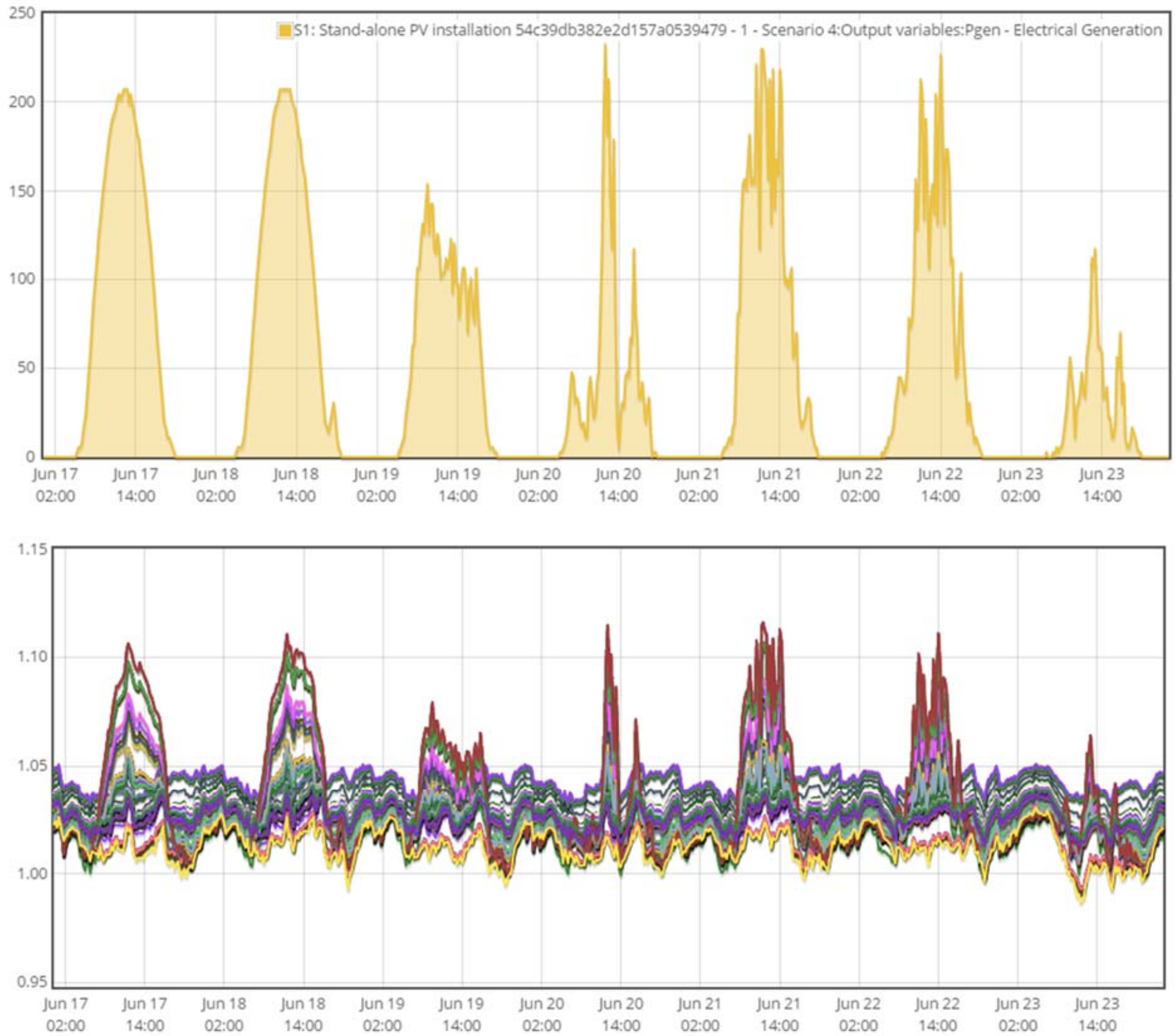


Figure 43 – *PV Deployment Case – Grid Situation (one summer week in June)*
above: Photovoltaic power production, below: Voltage band evolution (clear overvoltage violations)

Conventional Grid Upgrade Case

The conventional grid reinforcement would in this case be to replace all lines from the PV in-feed point to the transformer by underground cables with larger cross-section; this means in this case a change from 3x95 line diameter to a larger 3x150 line diameter. In comparison to the base case, the voltage rise is contained below the threshold of 1.10 pu (Figure 44).

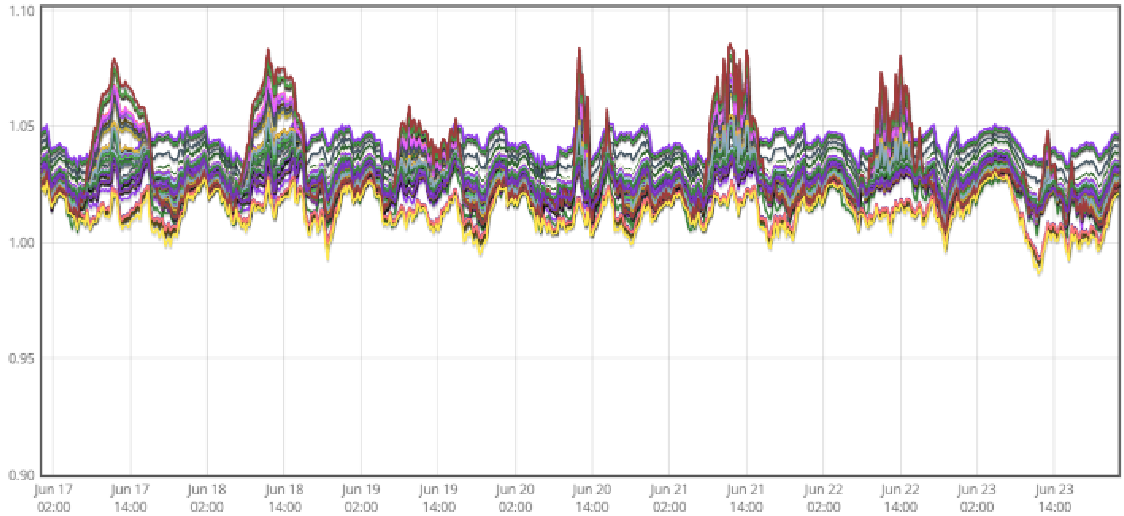


Figure 44 – *Conventional Grid Upgrade Case – Grid Situation (one summer week in June) Voltage band evolution (no visible overvoltage violations)*

PV Power In-Feed Curtailment Case

An effective alternative to line reinforcement is the controlled curtailment of available PV power in-feed. In many countries this has become the state-of-the-art option for coping with PV in-feed overflow in case all other available means fail. Here, PV in-feed is curtailed to 60% of $P_{installed}$ (typical rule used for larger German PV installations) when the bus voltage is above 1.10 pu. The PV curtailment pattern for a representative summer week is shown below (Figure 45).

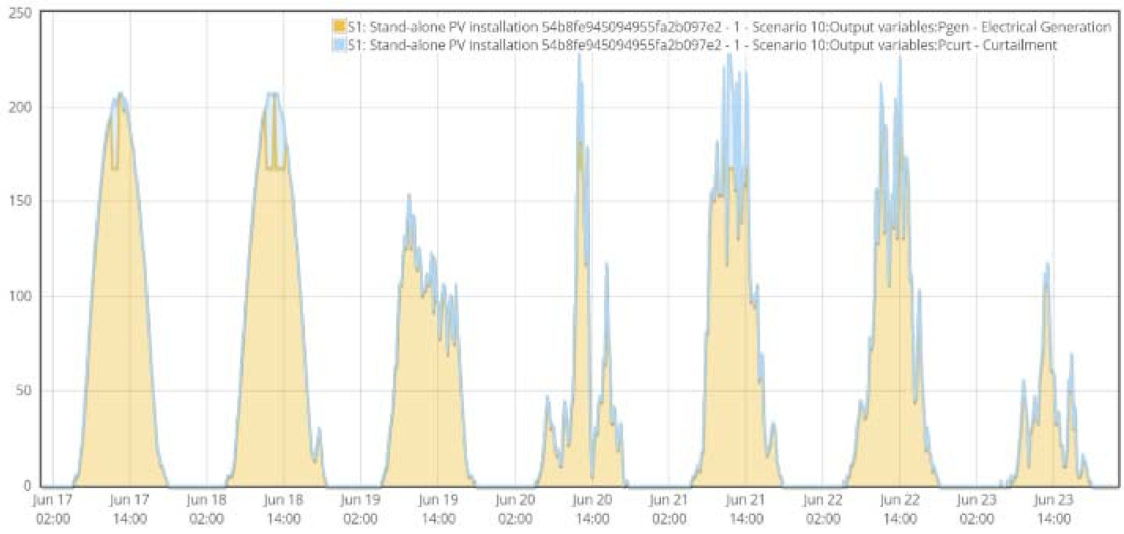


Figure 45 – *PV Curtailment Case – Grid Situation (one summer week in June) PV Production (shown in yellow) and PV Curtailment (shown in blue)*

Again, the voltage rise does not exceed the maximum acceptable voltage level of 1.10 pu (Figure 46).

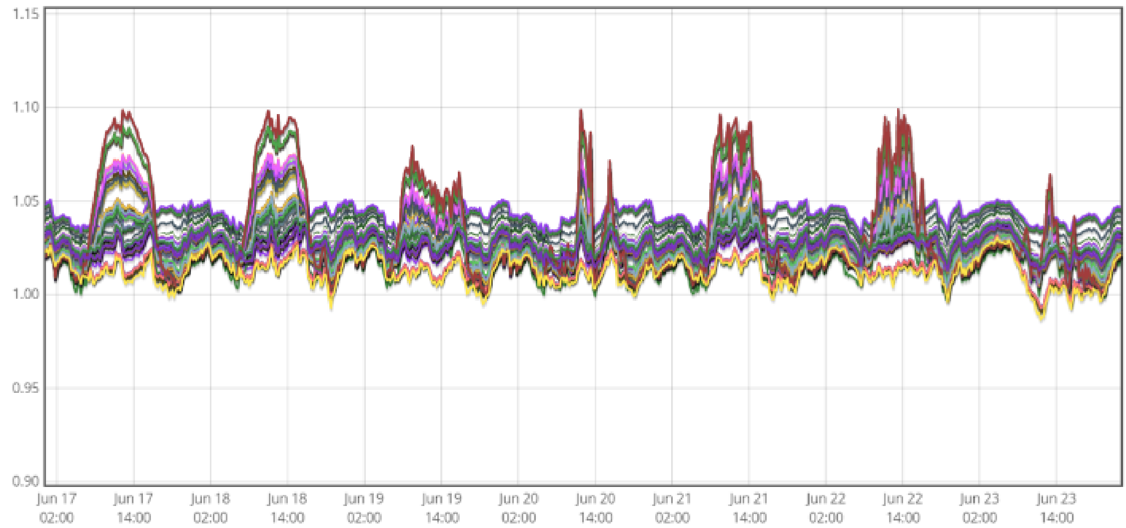


Figure 46 – *PV Curtailment Case – Grid Situation (one summer week in June)*
Voltage band evolution (no visible overvoltage violations)

Battery Storage Utilization

Absorbing surplus PV power in-feed by a battery energy storage system (BESS) is another effective means for grid integration. Here, a BESS system with a power rating of 100 kW, an energy rating of 600 kWh, and charging and discharging efficiencies of η = 0.9 are employed. The battery control is a simple hysteresis, i.e., charging whenever the bus voltage level rises above 1.09 pu and discharging when the bus voltage falls below 1.04 pu. The BESS charging/discharging patterns are shown in Figure 47. The BESS system effectively limits the voltage levels in the LV grid within allowable ranges (Figure 48).

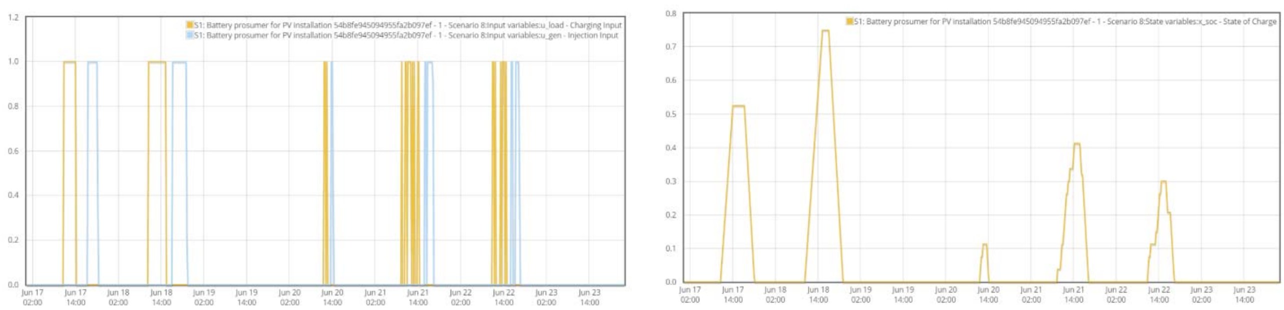


Figure 47 – *Battery Case – Grid Situation (one summer week in June)*
BESS Charging/Discharging (left) and State-of-Charge evolution [pu] (right)

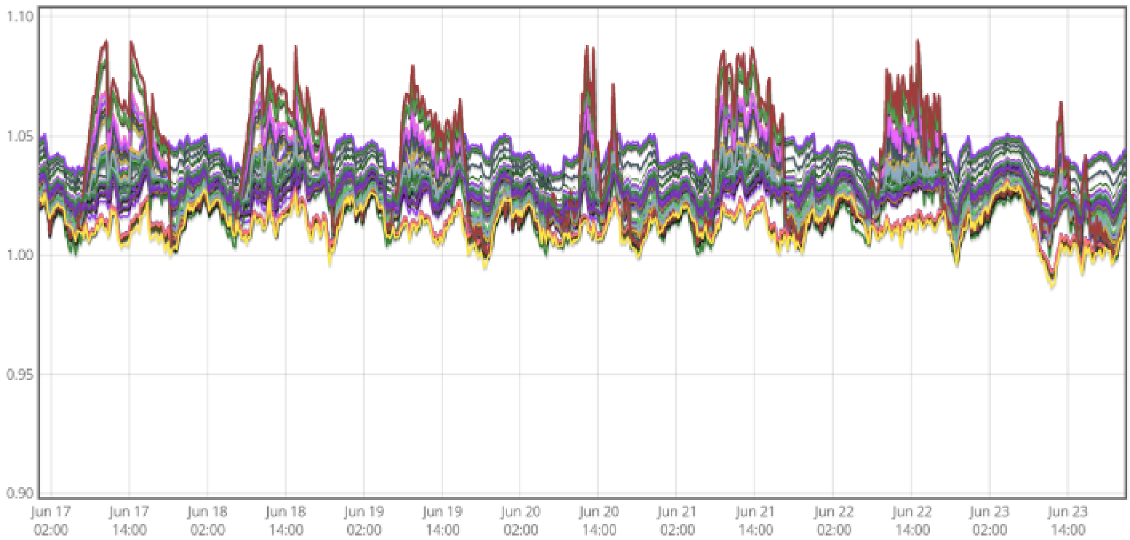


Figure 48 – *Battery Storage Utilization Case – Grid Situation (one summer week in June) Voltage band evolution (no visible overvoltage violations)*

Low-Voltage On-Load Tap Changer (OLTC)

Also a LV transformer station is considered that can dynamically adjust voltage levels via on-load tap changes (OLTC). The voltage measurement takes place at the location of the PV installation with a dead-band of 1.6 pu and a tap size of 1.5 pu. As a result, voltage levels are kept below 1.06 pu. However, tap change actions can, at times, also lead to severe under-voltages at other buses, close to the formally acceptable minimum of 0.90 pu (Figure 49). Considering the significant under-voltages as such, as well as the rapid voltage changes during the day, an OLTC alone does not appear to be a favorable means to mitigate the voltage problems caused by the PV installation.

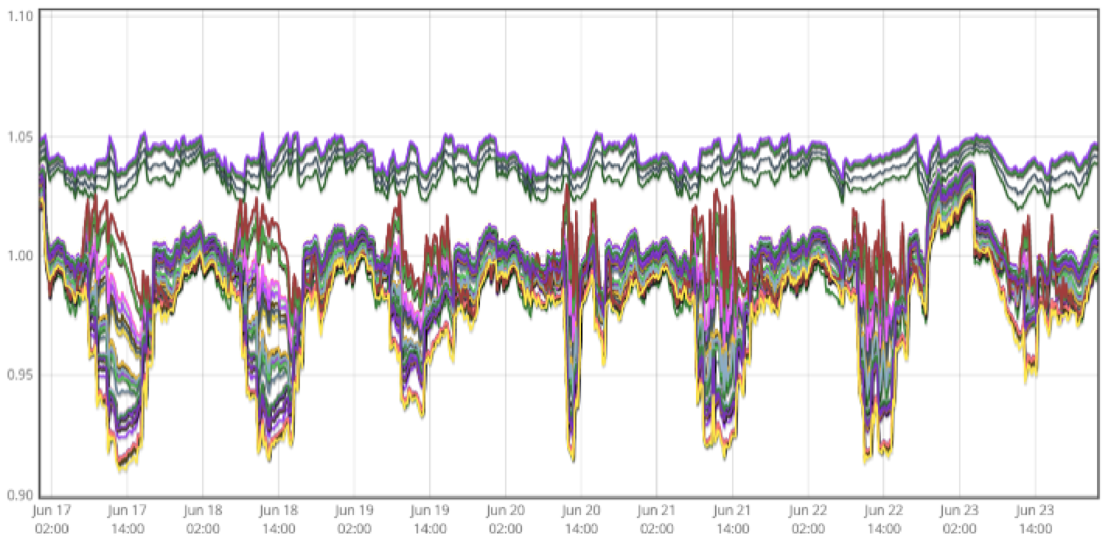


Figure 49 – *Low-Voltage OLTC Case – Grid Situation (one summer week in June) Voltage band evolution (no visible overvoltage violations)*

Technical Evaluation of Conventional and Smart Grid Upgrade Options

The presented study shows the nowadays much larger and more complex solution space for distribution grid planning strategies for accommodating large PV shares. It also showcases the unintended side-effects that the mitigation of over-voltage events can have, namely under-voltage events at other buses due to OLTCs and increased line loading due to reactive power control.

This underpins the need for novel grid simulation and planning tools that allow the evaluation of advantages and disadvantages of conventional and, in particular, novel Smart Grid reinforcement strategies. In the end this also includes a stringent assessment of the technical performance of each upgrade option, as shown below for the here presented large-scale PV deployment case (Figure 50).

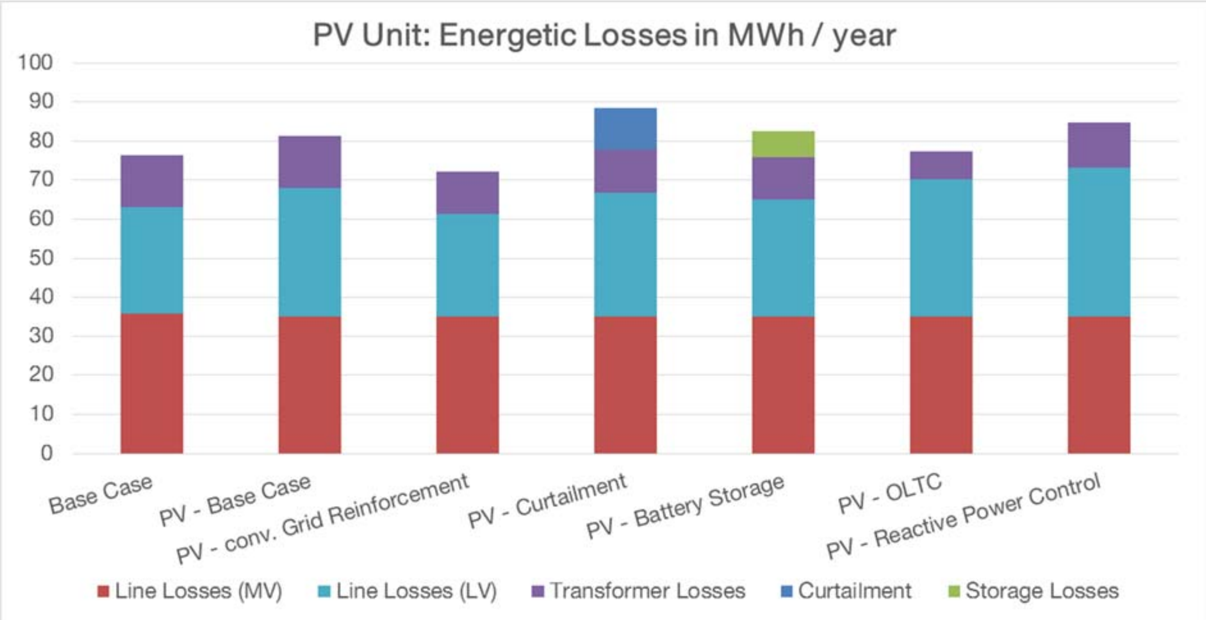


Figure 50 – *Technical Assessment of various Distribution Grid Upgrade Options*

3. Project Achievements

The main achievements of the project are, divided by activity field, the following:

Work Package ETH-G1

State-Estimation using Smart Metering data
The project results show that the developed concept for Distribution System State Estimation (DSSE) for Smart Meter measurements can accurately calculate the grid system states. This allows, for instance, to robustly estimate the grid situation for time-steps in between the typical 15 min. sampling intervals of Smart Meter systems today.

- Further efforts are needed to better understand the value of additional Smart Meter data collection beyond simple active energy measurements for state-estimation. Specifically relevant is here the assessment of the value of reactive power and local voltage measurements for state-estimation.

Grid Simulations using Smart Metering and other operational grid data
For the *Kleinhüningeranlage* pilot grid region, the project results showcased how – in the case of IWB – already today available operational grid data, ranging from Smart Meter data sets to PV production profiles can be employed to simulate and analyze low-voltage distribution grids with good temporal and special accuracy. The increasing availability of such power production and consumption time-series data sets allows an improved insight into the operation and system state of distribution grids, especially into the low-voltage grid parts inside settlements. This creates distribution grid transparency in grid parts which are today in most cases still black boxes for day-to-day grid operation.

Work Package ETH-G2

Data analytics and Day-Ahead Predictions of Smart Metering and other operational data
In the scope of this project, comprehensive data analytics have been performed with a large data set, allowing the usage of so-called *Big Data* technologies. Promising results were obtained for both Smart Meter clustering and forecasting; especially thanks to a good preprocessing work and an appropriate choice and use of machine learning algorithms in the analytics platform H_2O whose flexibility and scalability has been of great advantage. First of all, although the raw smart metering data from IWB is of good quality with few missing values, it still had to be prepared like any original dataset to facilitate the application of learning methods. The most challenging part was the missing data imputation carried out by means of the efficient “K-Nearest Neighbors” method. In a second phase, the “K-Means” clustering algorithm could detect various groups amongst small consumers even if the majority exhibits household properties. In addition, an interactive tool helps to visualize clustering outcomes as well as individual features since each smart meter is associated with a post address. Finally, neural networks have been chosen to perform short-term predictions. While an individual household is hardly predictable, the fact of aggregating multiple profiles considerably improves the forecasting performance. Although it is a delicate matter to compare forecasting methods based on various datasets, the results obtained here are close to current research study outcomes on the subject. The project results show that predicted Smart Meter time series can serve as a good substitute to often not available real-time data when carrying out power flow calculations in distribution grids.

Work Package ETH-L1

Development and Evaluation of Grid-friendly Local Control Strategies
As part of the Smart Metering project, several concepts for grid-friendly local control strategies for improving the grid operation in *stressed* distribution grids were developed and also employed in realistic Swiss distribution grids.

- Unfortunately, the showcasing of local active distribution grid control schemes could not be performed on IWB pilot grids (grid data only became available towards the end of project timeline).

4. Outlook

The combined results of the project are providing several proof-of-concepts of how increasingly available operational grid data, such as Smart Metering and PV production data, can effectively be exploited for providing better insights into actual distribution grid system state and grid operation.

The results are motivating the transformation of the here developed concepts into practical industry-grade grid simulation and analysis tools that enable an improved grid operation by providing a better degree of distribution grid transparency.

The next step will be the development and implementation of an online distribution grid monitoring and grid analysis tool, incorporating all available operational grid data. This will enable a clear(er) view inside the 'black-box' that low-voltage distribution grid operation still is today.

Note from IWB regarding Smart Metering Data Usage

Original German Version

Der Einsatz von fernauslesbaren Zählern im Versorgungsgebiet von IWB erfolgt im Einklang mit dem geltenden Datenschutzrecht und ist mit dem Datenschutzbeauftragten des Kantons Basel-Stadt abgestimmt. Nach den Ausführungsbestimmungen der IWB betreffend die Abgabe von Elektrizität vom 28. November 2011 (SG BS 772.400) ist der Einsatz von Smart Metern gestattet. Mit diesen Zählern werden einmal die abrechnungsrelevanten Verbrauchsdaten erfasst und fernausgelesen. Darüber hinaus werden auch Lastgänge (15min Lastverläufe) erfasst. Diese Daten werden (im Unterschied zu den abrechnungsrelevanten Daten) nicht individualisiert, sondern pseudonymisiert erfasst und in aggregierter Form zur Optimierung der Versorgung und der Stromnetze gespeichert. Die pseudonymisierten und aggregierten Lastgänge machen tageszeitliche Schwankungen des Stromverbrauchs sichtbar. Die Kenntnis dieser Schwankungen ist für die IWB als Netzbetreiber wichtig, weil Strom nicht im Netz gespeichert werden kann – aber die bereitgestellte elektrische Energie trotzdem exakt den zeitlichen Schwankungen des Stromverbrauchs entsprechen muss. Das ist ein wesentlicher Nutzen, der allen Stromkunden zugutekommt.

Eine Zuordnung dieser Daten zu einzelnen Kunden ist grundsätzlich ausgeschlossen. Rückschlüsse auf das individuelle Verbrauchsverhalten einzelner Kunden lassen sich aus diesen Daten nicht ziehen.

English Version

The deployment of remotely readable meters in the area supplied by IWB is carried out in accordance with the applicable data protection law and has been agreed upon with the Data Protection Supervisor of the canton of Basel-Stadt. Based on the implementation provisions of IWB concerning the electricity supply of 28 November 2011 (SG BS v 772.400), the deployment of smart meters is permitted. First, the consumption data relevant for accounting can be gathered and remotely read thanks to these meters. Furthermore, 15-minute load profiles are collected. This data (unlike the data relevant for accounting) is not individualized, but gathered and stored in strictly pseudonymous and aggregated form in order to optimize the supply and the power grids. The pseudonymized and aggregated load profiles make visible the fluctuations in the power consumption over the course of the day. The knowledge of these fluctuations is important for IWB as a grid operator since power cannot be stored in the grid – but the supplied electrical energy must nevertheless correspond exactly to the temporal fluctuations in the power consumption. This is an essential benefit that serves all electricity customers. An assignment of this data to single customers is generally excluded.

Conclusions about the individual consumption pattern of single customers cannot be drawn from this data.

References

[dena2012] German Energy Agency (dena), 2012, "dena Distribution Grid Study – German electricity distribution grids in need of significant expansion by 2030", [URL]: http://www.dena.de/en/projects/energy-systems/dena-distribution-grid-study.html.

[CIGRE2014] Cigré Working Group C6.19, 2014, "Planning and Optimization Methods for Active Distribution Systems ", Cigré Report.

[Heussen2012] K. Heussen, S. Koch, A. Ulbig, G. Andersson: Unified system-level modeling of intermittent renewable energy sources and energy storage for power system operation, IEEE Systems Journal, März 2012.

[Pflugrad 2014] Load Profile Generator. URL: http://loadprofilegenerator.de

[Alimardani et al. 2015] A. Alimardani, F. Therrien, D. Atanackovic, J. Jatskevich and E. Vaahedi, "Distribution System State Estimation Based on Nonsynchronized Smart Meters," in IEEE Transactions on Smart Grid, vol. 6, no. 6, pp. 2919-2928, Nov. 2015

[Angioni et al. 2016] A. Angioni, T. Schlösser, F. Ponci and A. Monti, "Impact of Pseudo-Measurements From New Power Profiles on State Estimation in Low-Voltage Grids," IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 1, pp. 70-77, Jan. 2016

[Manitsas et al. 2012] Manitsas, E.; Singh, R.; Pal, B.C.; Strbac, G., "Distribution System State Estimation Using an Artificial Neural Network Approach for Pseudo Measurement Modeling," in Power Systems, IEEE Transactions on , vol.27, no.4, pp.1888-1896, Nov. 2012

[Singh et al. 2010] Singh, R.; Manitsas, E.; Pal, B.C.; Strbac, G., "A Recursive Bayesian Approach for Identification of Network Configuration Changes in Distribution System State Estimation," in Power Systems, IEEE Transactions on , vol.25, no.3, pp.1329-1336, Aug. 2010

[Singh et al. 2009] Singh, R.; Pal, B.C.; Jabr, R.A., "Choice of estimator for distribution system state estimation" in Generation, Transmission & Distribution, IET , vol.3, no.7, pp.666-678, July 2009

[Papathanassiou et al. 2005] Papathanassiou S., Hatziargyriou N., Strunz K., 2005, "A benchmark low voltage microgrid network", Proceedings of the CIGRE Symposium: Power Systems with Dispersed Generation, April 2005, Athens, Greece.

[Jain2010] Jain, Anil K. "Data clustering: 50 years beyond K-means." *Pattern recognition letters* 31.8 (2010): 651-666.

[Zufferey2016] Thierry Zufferey, Andreas Ulbig, Stephan Koch and Gabriela Hug, Forecasting of Smart Meter Time Series Based on Neural Networks, Workshop „Data Analytics for Renewable Energy Integration (DARE), European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Riva del Guarda, 19-23. Sept. 2016.

[Vogel2016] Overview and Learnings of SFOE Smart Metering Project, SEV Bulletin, 10/2016.