StoRES

Promotion of higher penetration of Distributed PV through storage for all

Priority Axis 2: Fostering low-carbon strategies and energy efficiency in specific MED territories: cities, islands and remote areas

2.2: To increase the share of renewable local energy sources in energy mix strategies and plans in specific MED territories

Deliverable n°: **3.4.4 & 3.5.3**

Deliverable Name: **Data collection, monitoring and validation**

Project co-financed by the European Regional Development Fund

1. Document Info

Project Name	Promotion of higher penetration of Distributed PV through storage for all (StoRES)
Funding Scheme	ERDF
Work Package Number	WP3
Name of Work Package	Testing
Number	D3.4.4 & 3.5.3
Title	Data collection, monitoring and validation
Dissemination Level	PU
Date	19.09.2019
Authors	UCY
Contributors / Reviewers	All partners have reviewed the document before finalised
Status	Final

2. Document History

Date	Author	Action	Status
19.08.2019	UCY	Document creation	Draft
27.08.2019	UCY	Document update	Draft
05.09.2019	UCY	Document update	Draft
17.09.2019	AUTH	Document review	Draft
18.09.2019	UNICA	Document review	Draft
19.09.2019	UCY	Document finalisation	Final

3. Contents

4. List of Figures

5. List of Tables

6. Project Summary

The project addresses the development of an optimal policy for the effective integration of Renewable Energy Sources (RES) and Energy Storage Systems (ESS). The primary challenge is to achieve increased penetration of RES and predominantly Photovoltaics (PV), in the energy mix of islands and rural areas in the Mediterranean (MED) region without compromising grid stability. The main objective of StoRES is to boost self-consumption in the MED region with the integration of optimal storage solutions. Testing coupled PV-ESS solutions in different pilot sites and taking into account local particularities for optimization, current barriers concerning grid reliability with higher RES deployment will be eliminated. In addition to this, the development and integration of the proposed solution at both residential and community levels and the application of different policy scenarios will lift the barriers related to the grid integration of ESS and will extend the practical knowledge about this technology. It is expected that all the shortcomings regarding the intermittent nature of PV energy for increased penetration into the energy mix will be addressed whilst maintaining smooth operation of the grid.

The project started on the 1st of November 2016 and is expected to be completed within 36 months.

7. Introduction to Deliverable 3.4.4 & 3.5.3

In this report, the project's data collection process is briefly described. Moreover, the central database hosting the data collected from each pilot installation is presented. The University of Cyprus (UCY) created a central database, where all the measurements collected by each partner country are stored. The process of transferring the data to the abovementioned database is described. The pilots are regularly checked remotely to ensure their correct and optimal operation. This regards the Residential pilots in all participating countries and the Community pilot in Cyprus (Deliverables 3.4.4 and 3.5.3 respectively). Finally, the collected data are validated to ensure their high quality and usability for the research purposes of the project.

8. Data collection

8.1 Data acquisition

8.1.1 Residential pilots (Deliverable 3.4.4)

As thoroughly explained in the technical report of Deliverable 3.4.3 (Preparation of infrastructure, server & database), a common data collection format was agreed to be followed throughout the monitoring period, ensuring that appropriate data would be measured and collected by all project partners with implemented PV-ESS (Cyprus, Greece, Italy, Portugal, Slovenia and Spain) in each of their pilot installations (Deliverable 3.4.4). These data are further categorized to "Mandatory" and "Optional" data. The latter relies totally on the monitoring capability of each project partner, based on their available equipment and data collection process. Yet, similarly to the Mandatory data, the Optional data collection must comply with the commonly agreed format.

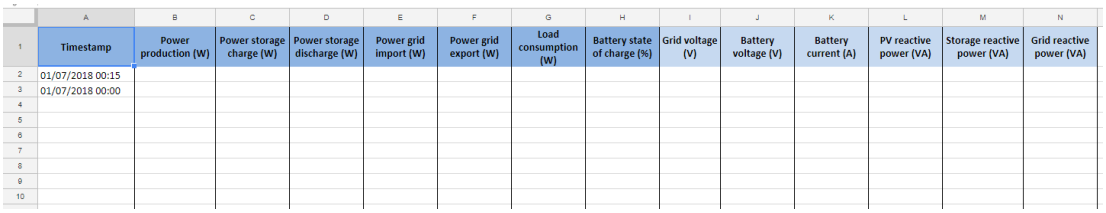
All data are acquired in 15-minute averaged values following the abovementioned data collection template. In Table 1, the residential pilots' Mandatory and Optional data are presented. The data collection frequency is depended on the limitation of the pilot equipment, i.e. selected metering apparatus, of most project partners. Specifically, data are acquired in a 15-minute time interval by four out of six project partners. Regarding the rest two partners, data are acquired in more frequent time intervals, i.e. per 1-minute and 5-minutes. These measurements are then readjusted in 15-minute values, to comply with the data collection template. It should be mentioned that the selected data collection frequency (15-minutes) is considered by all project partners as adequate for the research activities of the project.

Table 1: Residential pilots' monitored data.

Status	Parameter	Measurement unit
Mandatory	Timestamp	Date & time
	Power Generation	Watts (W)
	Power Storage Charge	Watts (W)
	Power Storage Discharge	Watts (W)
	Power Grid Import	Watts (W)
	Power Grid Export	Watts (W)
	Load Consumption	Watts (W)
	Battery State of Charge	Percentage (%)
Optional	Grid Voltage	Volts (V)
	Battery Voltage	Volts (V)

	Battery Current	Amperes (A)
	PV Reactive Power	Volt-Amperes reactive (VAr)
	Storage Reactive Power	Volt-Amperes reactive (VAr)
	Grid Reactive Power	Volt-Amperes reactive (VAr)

Figure 1 below illustrates the spreadsheet used as the data collection template. This must be filled for each pilot installation during the data collection period.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Timestamp	Power production (W)	Power storage charge (W)	Power storage discharge (W)	Power grid import (W)	Power grid export (W)	Load consumption (W)	Battery state of charge (%)	Grid voltage (V)	Battery voltage (V)	Battery current (A)	PV reactive power (VA)	Storage reactive power (VA)	Grid reactive power (VA)
2	01/07/2018 00:15													
3	01/07/2018 00:00													
4														
5														
6														
7														
8														
9														
10														

Figure 1: Data collection template (Mandatory and Optional data in dark blue and light blue respectively).

8.1.2 Community pilot (Deliverable 3.5.3)

Subsequent to the data collection regarding the Residential pilots, a similar procedure is followed for the data collection of the project's sole Community pilot, which is installed in Cyprus (Deliverable 3.5.3). The monitored data from the Community pilot are listed in Table 2.

Table 2: Community pilot's monitored data.

Parameter	**Measurement unit**
Measurement time	Date & time
Battery State of Charge	Percentage (%)
EMS-CM Frequency Control	Watts (W)
EMS-CM Power Balancing	Watts (W)
EMS-CM Target SoC	Watts (W)
EMS-CM Voltage Control	Volt-Amperes reactive (VAr)
External Meter Active Power Total	Watts (W)
Grid Current L1	Amperes (A)
Grid Current L2	Amperes (A)
Grid Current L3	Amperes (A)
Grid Frequency	Hertz (Hz)
Grid Total Active Power	Watts (W)

Grid Total Reactive Power	Volt-Amperes reactive (VAr)
Grid Voltage L1-N	Volts (V)
Grid Voltage L2-N	Volts (V)
Grid Voltage L3-N	Volts (V)
Inverter Active Power	Watts (W)
Inverter Reactive Power Total	Volt-Amperes reactive (VAr)

The above data are collected remotely by the manufacturer and provided daily to the UCY by an automated email. The daily files are then combined by the UCY to result to a monthly file, which is then stored directly in the central UCY database.

In March 2019, a mode change regarding the operation of the Community pilot was performed. Specifically, the manufacturer re-evaluated the system functions (mostly the power balancing) and concluded that some improvements could be performed. Thus, the manufacturer inserted two independent power limiters, one responsible for import limitation and the other for export limitation. As a result, a slight change on the exported data was observed and subsequently the Community data collection file was modified. The data collection period after this change was named "Phase B".

The slightly modified data regarding the Community pilot in Cyprus are briefly described below:

- **Battery SoC [%]**
  State of Charge (SoC) of the Energy Management System (EMS).
- **Frequency Control [W]**
  Setpoint of the Grid Frequency Support (GFS) feature (active power according to configured curve).
- **EMS Limiter 1 [W]**
  Setpoint of power limiter 1 (configured as import limiter).
- **EMS Limiter 2 [W]**
  Setpoint of power limiter 2 (configured as export limiter).
- **EMS Preserve SoC [W]**
  Setpoint of the SoC-preservation feature (tries to stabilise the SoC if the battery is at the low end).
- **EMS Target SoC [W]**
  Setpoint of the SoC function (tries to get the SoC to the configured targeted SoC - currently 50%).

- **EMS Voltage Control [VAr]**
  Setpoint of the Grid Voltage Support (GVS) feature (reactive power according to configured curve).
- **Grid Current L1 [A]**
  Current phase 1 at the grid-connection point.
- **Grid Current L2 [A]**
  Current phase 2 at the grid-connection point.
- **Grid Current L3 [A]**
  Current phase 3 at the grid-connection point.
- **Grid Frequency [Hz]**
  Frequency at the grid-connection point.
- **Grid Total Active Power [W]**
  Active power at the grid-connection point (measured as consumer).
- **Grid Total Reactive Power [VAr]**
  Reactive power at the grid-connection point (measured as consumer).
- **Grid Voltage L1-N [V]**
  Voltage phase 1 at the grid-connection point.
- **Grid Voltage L2-N [V]**
  Voltage phase 2 at the grid-connection point.
- **Grid Voltage L3-N [V]**
  Voltage phase 3 at the grid-connection point.
- **Inverter Active Power [W]**
  Active power at the ESS-output inverter.
- **Reactive Power Total [VAr]**
  Reactive power at the ESS-output.

8.2 Post-processing phase

In order to comply with the desired residential data measurement requirements described in subsection 8.1.1, an additional step is required at a regional level to ensure the compatibility of the collected data. This step requires each pilot country to make additional data calculations or even convert data to the requested format if required, such as converting energy values to power values, adjusting the data measurement units (i.e. timestamp format, data collection frequency, power values in Watts, voltage values in Volts, etc.) and if needed, calculating any additional parameters such as direct PV consumption of the pilot based on the other collected data. Finally, this phase regards the replacement of any missing values (which mostly occur due to communication problems of the equipment) with "NaN" and the observation and correction of any possible issues with faulty numbers.

The so-called post-processing phase is considered of significant importance to ensure that the requirements of the common data collection format are met, as generally partners have installed different equipment and use various data collection processes. This phase takes place by each project partner before the transferring of the data to the central UCY database. In order to achieve that, an additional step is required to convert the data to a format that is compatible for all pilot countries during the post-processing phase. The files sent to the central UCY database are in a commonly agreed format (.CSV spreadsheet file).

It has to be noted that during the data collection period all pilots are operated in a common operation mode (i.e. increase of self-consumed energy generated by the PV system on-site).

Each pilot partner was responsible to implement the corresponding local database and more importantly to decide about its structure in line with the project needs. It is important to note that storing the data to a local database is also considered as a back-up of the project's central database.

Figure 2 illustrates the graphic representation of the post-processing phase and the storing of the data at the regional level.

Figure 2: Graphic representation of the post-processing phase and the data storing at the local database.

8.3 Data transferring

As thoroughly explained in Deliverable 3.4.3, the UCY stores the data from its pilot installations (5 Residential pilots and the Community pilot) directly to the central UCY database once per month. The Aristotle University of Thessaloniki (AUTH) provided access to the UCY team to a folder in its own cloud storage database. There, the data for the 5 pilots in Greece are uploaded by the AUTH team once per month and can be retrieved by the UCY team at any time. Regarding the rest of the project partners (Italy – 13 pilots, Portugal – 5 pilots, Slovenia – 1 pilot and Spain – 5 pilots), the UCY created a shared online folder for each partner, following a similar data acquisition process.

The UCY uploads and archives the retrieved files to the central UCY database once per month, specifically during the first week of every new month. Tables have been created in the database for each pilot installation specifically (35 pilots in total and an additional for the second phase of the data collection of the Community pilot) under a specific name format, as it can be seen in Table 3.

Table 3: List of pilots in the Central UCY Database.

PROJECT PATNER	COUNTRY	TABLE NAME
1.	Cyprus	CY_Pilot_1; CY_Pilot_2; CY_Pilot_3; CY_Pilot_4; CY_Pilot_5; CY_Pilot_6Ca; CY_Pilot_6Cb
2.	Greece	GR_Pilot_1; GR_Pilot_2; GR_Pilot_3; GR_Pilot_4; GR_Pilot_5
3.	Italy	IT_Pilot_1; IT_Pilot_2; IT_Pilot_3; IT_Pilot_4; IT_Pilot_5; IT_Pilot_6; IT_Pilot_7; IT_Pilot_8; IT_Pilot_9; IT_Pilot_10; IT_Pilot_11; IT_Pilot_12; IT_Pilot_13
4.	Portugal	PO_Pilot_1; PO_Pilot_2; PO_Pilot_3; PO_Pilot_4; PO_Pilot_5
5.	Slovenia	SL_Pilot_1
6.	Spain	SP_Pilot_1; SP_Pilot_2; SP_Pilot_3; SP_Pilot_4; SP_Pilot_5

8.4 Central UCY database

The UCY created a database named "storesdb" in the MySQL group, specifically for the purposes of the StoRES project as it can be seen in Figures 3 and 4.

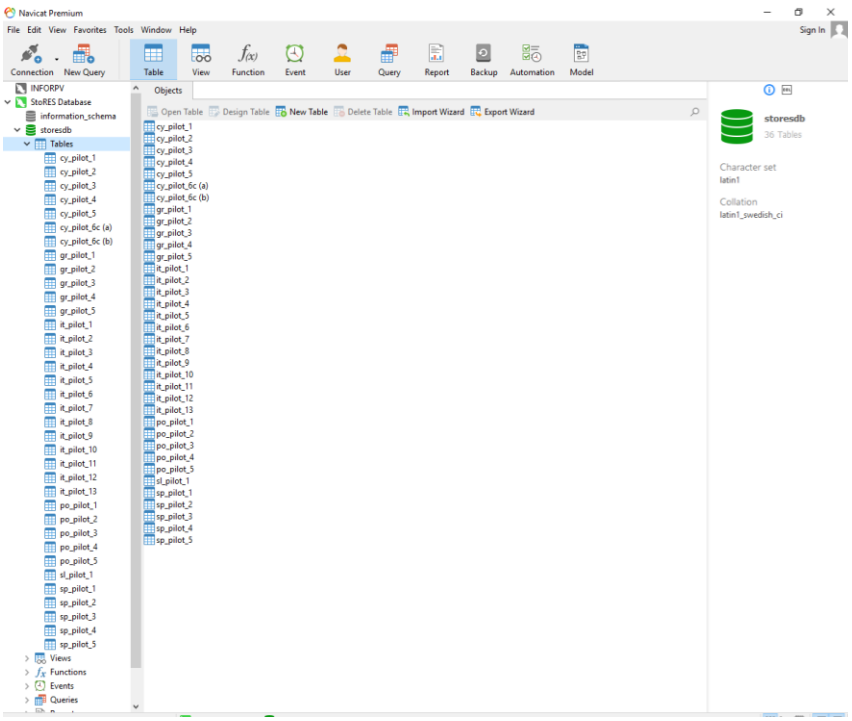


Figure 3: StoRES database interface illustrating the 36 tables (for each project pilot).

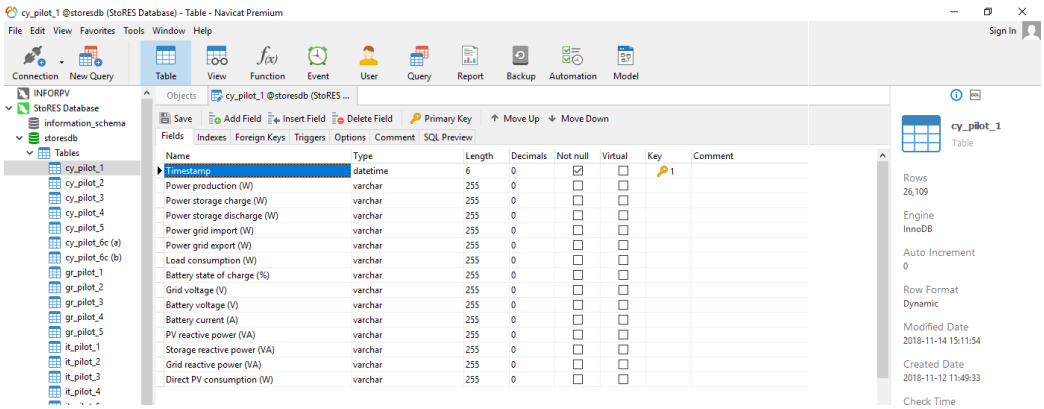


Figure 4: An example of the measured parameters of a pilot.

8.5 Store to central UCY database

Once the data are collected to the local databases, the final step of the data collection phase is to store the data to the central UCY database, as mentioned above. The data from each pilot are monitored frequently by the project partners to ensure the proper operation of the installations, whereas the central database is updated once per month. The official initial date of data collection in all pilot countries was the 1st July 2018, however most of the project partners were already collecting data, whilst others faced minor issues regarding their pilot installations and an extension in data collection was required.

Specifically, Table 4 presents the initiation of the data collection phase in each country, which was mainly depended on the commissioning of the pilot systems. However, it can be seen that for most pilot countries, data of more than a year have already been collected. It has to be noted that, to this point, the central UCY database includes data until the end of August 2019. The data collection is planned to continue until the end of the project (October 2019).

Table 4: Data collection initiation in each country.

Country	Data collection initiation
Cyprus	February/July/November 2018 (Residential pilots)
	January 2019 (Community pilot)
Greece	June 2018
Italy	July/October 2018
Portugal	April 2018
Slovenia	October 2018
Spain	June/July 2018

Figure 5 illustrates the whole data collection process: from data acquisition in each country's pilot installation to the storage of the agreed data to the central UCY database.

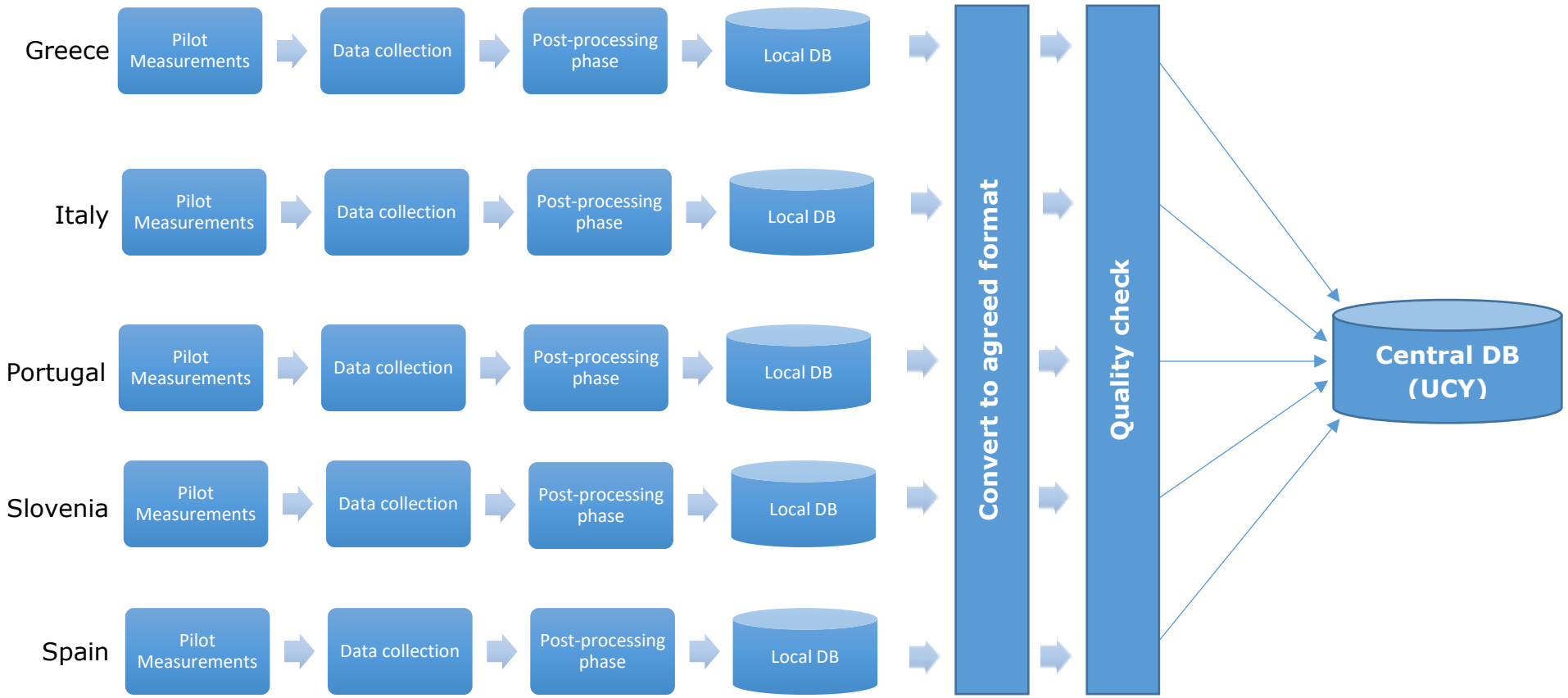


Figure 5: Representation of the data acquisition phase in each pilot country and the data transferring to the central database.

9. Monitoring

The pilots are regularly checked remotely to ensure their correct and optimal operation. Each partner is responsible for storing the measurements in their local database and mostly for providing them to the central database administrated by the UCY in the agreed time (within the first week of every new month). In case of any missing data, which generally occurs due to communication problems of the equipment, the value must be presented as "NaN".

For the case of the Community pilot, the manufacturer also performs regular remote monitoring of the equipment and provides any correction to the operation of the system and subsequently of the data if required, after any issues are observed during the validation.

Finally, the UCY remains responsible to regularly perform a data quality test and ensure that the data from all partners are stored on time and are of good quality.

10. Validation

Each partner is responsible for validating their data. One of the methods used by the UCY during the quality check of the data in the central database is the "Energy Balance observation", whereas the sum of the generated, charged and imported energy in a pilot installation must equal at all times the sum of the consumed, discharged and exported energy. Furthermore, the UCY uses another two equations during the quality check, namely the "Consumption observation" and the "Generation observation". In the former, the sum of the total energy consumed in the household must be equal to the total demand, while in the latter, the total energy generated by the household must be equal to the total PV production. Specifically, with regards to the parameters presented in Table 1, equations (1), (2) and (3) should be met for all timesteps:

$$A + B + C = D + E + F \quad (1)$$

$$A = C + E + G \quad (2)$$

$$D = B + F + G \quad (3)$$

where A: Load Consumption,
B: Power Storage Charge,
C: Power Grid Export,
D: Power Generation,
E: Power Storage Discharge,
F: Power Grid Import,
G: Direct PV Consumption.

In addition, the data are frequently checked in order to identify any negative values (i.e. PV power production or SoC) or any other possible issues with faulty numbers (e.g. outliers). In case of any missing data, which generally occurs due to communication issues of the equipment as mentioned in Section 9, the "Energy Balance observation" is used to estimate and replace any missing values. In cases that more than one value is missing, the developed mining algorithm, as presented in Figure 6 observes the pre and post values and provides a suitable estimation of the missing data.

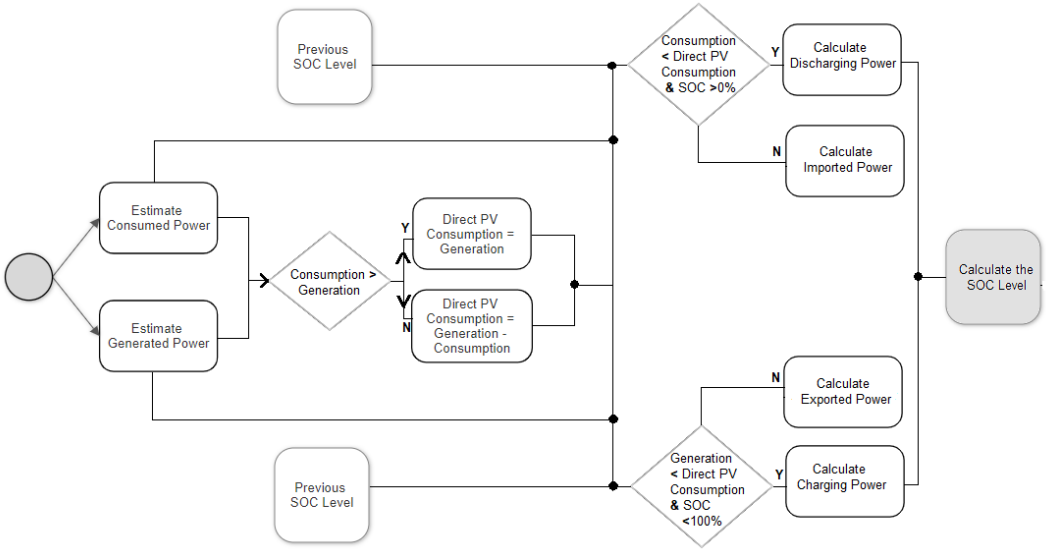


Figure 6: Developed mining algorithm for missing data.

It has to be noted that Slovenian and Spanish partners follow the same approach for data validation during their data post-processing.

The data validation during the post-processing phase of the Greek pilot installations can be summarised into three steps, given the particularities of the metering devices. Initially, any negative values that might be identified in the measurements of PV production are replaced with zero. Next, the load consumption is calculated using the measurements from the points where the metering devices are installed, i.e. grid, ESS and PV system. Finally, due to the different measurement resolution between the battery controller and the energy analyser installed, the calculated load consumption may acquire negative values. Thus, if measurements of the Load Consumption parameter are negative, their values are replaced with "NaN".

As far as the Portuguese pilots are concerned, the data validation process involves essentially the following steps:

- Observation through the manufacturer's platform of the various graphical representations, in order to detect any anomalous register.
- After data extraction, the number of 15-minute timesteps is counted according to the number of days of the month under analysis.
- If there is any fault in the previous step, the number of 15-minute timesteps per day (96) is still quantified in order to identify which day the data is missing.
- After identifying the missing data, this fault is "fixed" using the values of the previous 15-minute period.
- Finally, if the data missing exceed a consecutive period of 15-miniutes, the values are extrapolated assuming generic weekdays and weekends to fill in the gaps.

Regarding the Italian pilots, once the data are acquired from the cloud database provided by the battery manufacturer, before the post-processing phase, they are validated following the same approach based on the "Energy Balance observation", the "Consumption observation" and the "Generation observation used by the UCY. However, there may be situations in which time intervals are lost. Particularly, two events may happen:

- Lack of hours of information, mainly due to the lack of communication between the storage system and the cloud database provided by the battery manufacturer. When the storage system is able to communicate again with the database, it also provides the missing data allowing filling the gap. Such data are updated in the local database.
- Lack of few time intervals; in this case the missing data are calculated as an average between the pre and post available data.

10.1 Quality check before Data Analysis

Auvergne-Rhône-Alpes Énergie Environnement (AURA-EE) is responsible for the data analysis purposes of the project. AURA-EE receives all the data stored at the central UCY database once per month directly by the UCY. When the first datasets where provided in July 2018, AURA-EE made a detailed check of all the files. Some errors were identified for a number of pilots at that point.

			discharge (W)		(W)			
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	complet
Ussaramanna	data collected with many errors	data collected with many errors	data collected with many errors	data collected with many errors	data collected with many errors	data collected with many errors	data collected with many errors	beaucoup de valeurs nulles, trop de séries manquaient pour p reconstituées pour le mois d'octobre.Depuis le mois de décem priori réparé
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	complet
Ussaramanna							missing for Sonnen System	mois de novembre reconstruit. Données manquantes en décen problèmes de connexion internet. Complet jusqu'à novembre
Ussaramanna	2 days missing (connection problems)	2 days missing (connection problems)	2 days missing (connection problems)	2 days missing (connection problems)	2 days missing (connection problems)	2 days missing (connection problems)	2 days missing (connection problems)	problèmes de connexion mais les données ont été stockées. E reconstruites
Kozani	* PV production sometimes negative : clarification required	* les batteries sont chargées depuis le réseau toutes les 2 semaines pour maintenir un SOC minimal - PV trop faible pour couvrir la charge qui est élevée (= consommation + remplissage batteries) - paramètrage en cours d'ajustement pour voir s'ils peuvent résoudre ce problème	battery sometimes discharges to the grid - clarification required			* load consumption sometimes negative : clarification required	some data missing	fichiers avec données surprenantes
Koilada	* PV production sometimes negative : clarification required						some data missing	fichiers avec données surprenantes
Mavrodendri	* PV production sometimes negative : clarification required					* load consumption sometimes negative : clarification required	some data missing	fichiers avec données surprenantes
Koilada						* load consumption sometimes negative : clarification required	some data missing	fichiers avec données surprenantes
Valtero	* PV production sometimes negative : clarification required					sometimes negative : clarification required * conso de type industriel - quasi en autoconso totale	some data missing	fichiers avec données surprenantes
								Il manquait des points 15min - tout est corrigé
								Il manquait des points 15min - tout est corrigé
								Il manquait des points 15min - tout est corrigé
								Il manquait des points 15min - tout est corrigé
Benalm								Il manquait des points 15min - tout est corrigé
								va corriger les données erronnées - méthodo en atten

Figure 7: Identification of data errors in July 2018.

Specifically, the following errors were observed:

- Missing timesteps
- Additional timesteps provided
- Missing values
- Negative values for load consumption
- Charge higher than production

To this context, it was decided that the data post-processing protocol had to be improved by all partners. As a result, the number of errors was greatly decreased.

In order to ensure that the collected data are always consistent, AURA-EE developed an algorithm for the automatic check of the energy balance, as seen in equation (1), for all pilot sites. This regards the timesteps containing only one "NaN" value, i.e. one missing value. If an "NaN" is detected, then the missing value is replaced by its calculated value using the abovementioned equation.

11. Conclusions

In this report, the data collection process in the context of the StoRES project is briefly described. Starting from the data acquisition at each pilot installation, the very important post-processing phase which ensures the suitability of the collected data is demonstrated. This phase occurs before transferring the data from the local databases to the central database administrated by the UCY. A simple, yet efficient, method for the UCY to acquire the data and store them in the central database was designed and implemented in full communication with the rest of the project partners.

The data from each pilot are monitored frequently by the project partners to ensure the proper operation of the installations, whereas the central database is updated once per month. For the vast majority of the pilot countries, data of more than a year have already been collected. To this point, the central UCY database includes data until the end of August 2019, which are used for the research purposes of the project. This regards data from the Residential pilots in all participating countries and the Community pilot in Cyprus (Deliverable 3.4.4 and Deliverable 3.5.3 respectively). The data collection is planned to continue until the end of the project (October 2019).