Grant Agreement N° 872592

PLATOON

Digital platform and analytic tools for energy

Deliverable D6.5-Evaluation and Validation Report

Contractual delivery date:
M30
Actual delivery date:
30/06/2022
Responsible partner:
TECN, Spain

Project Title	PLATOON – Digital Platform and Analytic Tools for Energy
Deliverable number	D6.5
Deliverable title	Evaluation and Validation Report
Author(s):	Erik Maqueda
Responsible Partner:	TECN
Date:	30/06/2022
Nature	R
Distribution level (CO, PU):	PU
Work package number	WP6 – Large Scale Pilot Implementation and Validation

Work Package Leader	VUB
Abstract:	This report summarises the preliminary validation results of the different components developed in WP2, WP3 and WP4 for all the pilots and corresponding use cases. The validation is performed based on the KPIs identified in the validation plan as per explained in deliverable D6.6. This report is the first version of the validation report and might not be fully complete. The main objective of this first version is to be an initial touchpoint to understand the current status of the validation of the different pilots, identify potential risk and put a contingency plan towards the final complete version that will be submitted in the end of the project (M36).
Keyword List:	Validation, KPIs, Impact assessment, Large Scale Pilots, Use cases.

The research leading to these results has received funding from the European Community's Horizon 2020 Work Programme (H2020) under grant agreement no 872592.

Editor(s):	TECN
Contributor(s):	VUB, IMP, CS, SAM, IND, ENGIE, PI, ENG, ROM, SIS, PDM and IAIS.
Reviewer(s):	Jan Helsen (VUB) Philippe Calvez (ENGIE)
Approved by:	Jan Helsen (VUB) Philippe Calvez (ENGIE)
Recommended/mandatory readers:	WP6 partners and WP8 and WP9 task leaders.

Document Description

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
V.01	31/03/2022	TOC	Erik Maqueda
V1.0	02/06/2022	Executive summary, introduction and general information for pilots. Added validation results from TECN for Pilot 1A, 2B and 3C	Erik Maqueda
V1.1	20/06/2022	Added missing information from rest of the pilots.	VUB, IMP, CS, SAM, IND, ENGIE, PI, ENG, ROM, SIS, PDM and IAIS.
V1.2	30/06/2022	Amendments after internal review	Erik Maqueda

Executive Summary

This report summarises the evidence and analysis of the preliminary validation results of the different components developed in WP2, WP3 and WP4 collected upon the execution of the large-scale pilots for all sites and use cases. The validation is performed based on the KPIs identified in the validation plan comparing them to the defined targets as per explained in deliverable D6.6. The validation report collates the conclusions from all stakeholders involved in the different pilots. Also, it contains the validation results of the PLATOON Common components which are cross-pilot. Finally, in the conclusions section an overall evaluation is done including the identification of the main pending aspects and risks.

This report is the first version of the validation report and is not fully complete. The main objective of this first version is to be an initial touchpoint to understand the current status of the validation for the different pilots, identify potential risks and put a contingency plan towards the final complete version that will be submitted in the end of the project (M36).

Table of Contents

List of Figures

List of Tables

List of Terms and Abbreviations

CA	Consortium Agreement
CO	Confidential
DM	Dissemination Manager
EC	European Commission
ES	Energy System
EM	Exploitation Manager
GA	Grant Agreement
GAM	General Assembly Meeting
IDS	International Data Spaces
MSLE	Mean Squared Logarithmic Error
PM	Project Manager
PML	Process Mastery Level
POD	Point of Delivery
PDR	Point of Re-delivery
PU	Public
QA	Quality Assurance
RE	Restricted
ReLU	Rectified Linear activation fUnction
SC	Steering Committee
TM	Technical Manager
WP	Work package
WPL	Work package Leader

1. Introduction

This report summarises the evidence and analysis of the preliminary validation results of the different components developed in WP2, WP3 and WP4 collected upon the execution of all large scale pilot sites and corresponding use cases. The validation is performed based on the KPIs identified in the validation plan defined in deliverable D6.6. This deliverable is structured in the following sections:
Sections 2-9 summarise the validation results and conclusions for all the large scale pilot sites and corresponding use cases.
Furthermore, section 10 contains the validation results of the PLATOON Common components which are cross-pilot.
Finally, there is a conclusions section where an overall evaluation is done.
Besides, Annex I explains the KPIs for the different pilots and common components which were initially defined for deliverable D6.6 and some of which have been updated.

2. Pilot 1A Evaluation & Validation Report

2.1 Introduction

This pilot focusses on wind farms both onshore and offshore with a specific focus on wind turbines in the range of 1.5-3MW owned by ENGIE in different locations across Europe. There is a single use case focused on predictive maintenance of wind turbine electrical drivetrain components which aims to:

1. Develop, implement and validate accurate physical and data-driven models of the wind turbine electrical drivetrain components: generator and power converter.
2. Develop anomaly detection methods for identification of unhealthy behaviour of the components in scope.
3. Develop an approach to convert the identified anomalies towards health indicators to create a diagnostic tool.
4. Extract the relevant events that the electrical drivetrain components are exposed to and have a potential negative effect on the lifetime of the electrical components.

2.2 LLUC-1A-01-Predictive maintenance of wind turbine electrical drivetrain components

This use case focuses on data analytics tools to accurately detect failures in the electrical components of wind turbines, limited specifically to the generator (doubly fed induction generator) and the power converter. In this use case two different approaches are used:

1. Hybrid-digital twin approach developed by TECN
2. Data driven approach developed by VUB

2.2.1 Evaluation and Validation

2.2.1.1 Hybrid-digital twin approach

Table 1: LLUC-1A-01-KPIs evaluation- Hybrid-digital twin approach

KPI #	Description	Target Value	Actual Value	Comments
1	Modelling quality	3%	Active Power – MAPE=2.33% Current – MAPE=2.66%	The error for the active power and current parameters is below the threshold value of 3%.

			Stator Winding Temperature – MAPE=4.33%	However, the error for the stator winding temperature is above the target threshold.
2	Integration	1	0.8	All PLATOON components have been implemented except the IDS scenario where ENGIE is a data provider. Also, the semantic adaptation of results from TECN and VUB is missing.
3	Fault Detection	Compared to the current failure detection the speed should improve with at least 25%, while keeping false positives below 10%	The new algorithms have not improved the current detection speed. However, the false positives have stayed the same.	Although the fault detection target value has not been achieved, the developped algorithm has proven to help the troubleshooting of the failed components. Besides, for V2 of this document we are currently working on new algorithms that include more features to try to improve the fault detection KPI.
4	Processing Capability	Full processing chain for a farm should be able to run on a standard server.	Full processing chain is able to run on a standard server.	The training part and dockerisation of the Digital Twin in Matlab showed a high consumption of RAM and CPU. However, once it is dockerised the execution phase is significantly less computationally expensive and can be run in any standard server with 4CPUs and 16GB RAM.
5	Maintenance costs reduction	10-20%	N/A	Cannot be calculated given the actual results and available maintenance data.
6	Availability increase	2-5%	N/A	Cannot be calculated given the actual results and available maintenance data.

In order to validate the data analytic tools for predictive maintenance of wind turbines developed in WP4, the tools have been trained and tested with data from several onshore wind farms from ENGIE all with data from Senvion units of 2MW.
Any data regarding turbine, identifier, location and date have been removed due to confidentiality issues.

Regarding the modelling quality KPI, for the Normality Digital Twin of the Electric Generator it can be noted that the results have significantly improved compared to the ones obtained for the 1.5MW GE units used for model development in WP4. This is due to the fact that the 2MW Senvion turbines have a torque sensor and, thus, we can use the measured torque directly as an input to the electromechanical digital twin. This reduces significantly the uncertainty linked to the aero-mechanical model due to different parameters (direction, density…) that affect on the effective wind speed. Figure

1**¡Error! No se encuentra el origen de la referencia.** shows the comparison of real data (orange), simulated data using torque as input (blue) and simulated data using wind speed as input (yellow).

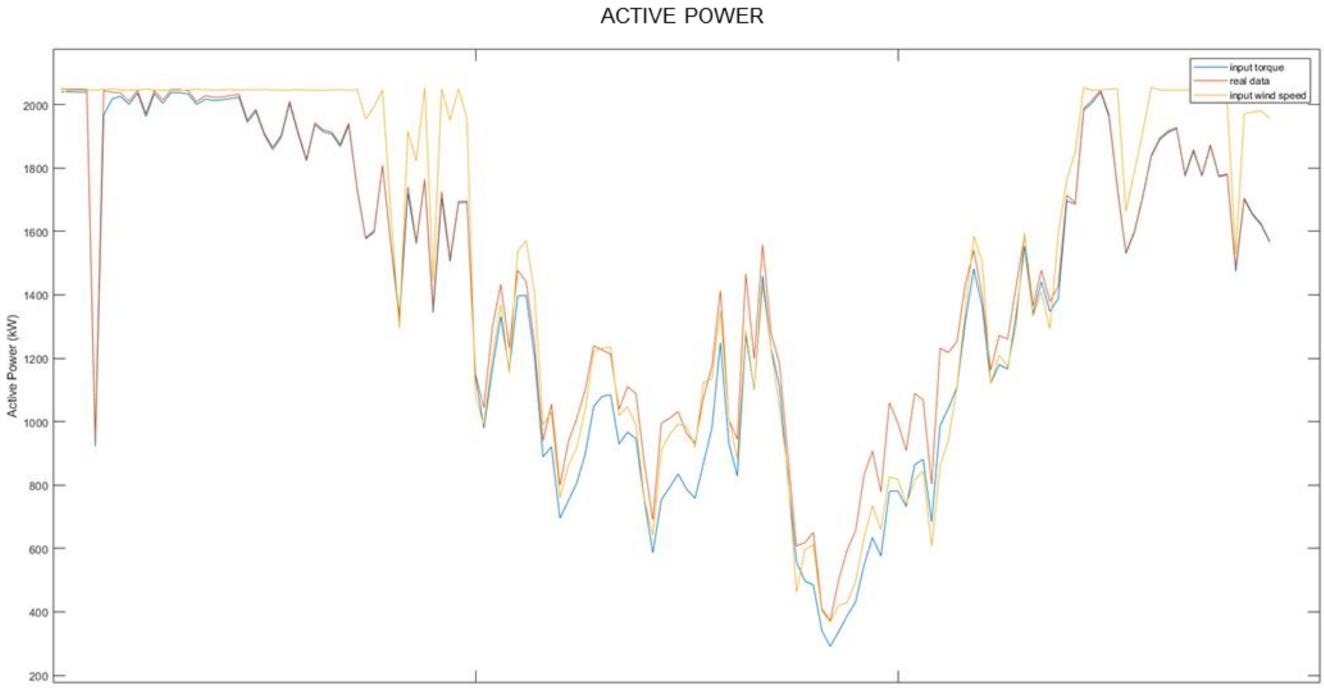


Figure 1: LLUC-1A-01-Normality Hybrid Digital Twin - Validation Results- Modelling quality - Active Power

Figure 2 and Figure 3 show the simulated (blue) and real data (orange) for current and stator winding temperature parameters, respectively. As it can be noted the error for the stator winding temperature parameter is larger compared to the active power and current. In fact, the error for the active power and current parameters is below the threshold value of 3%, whereas, the error for the stator winding temperature is above the target threshold.

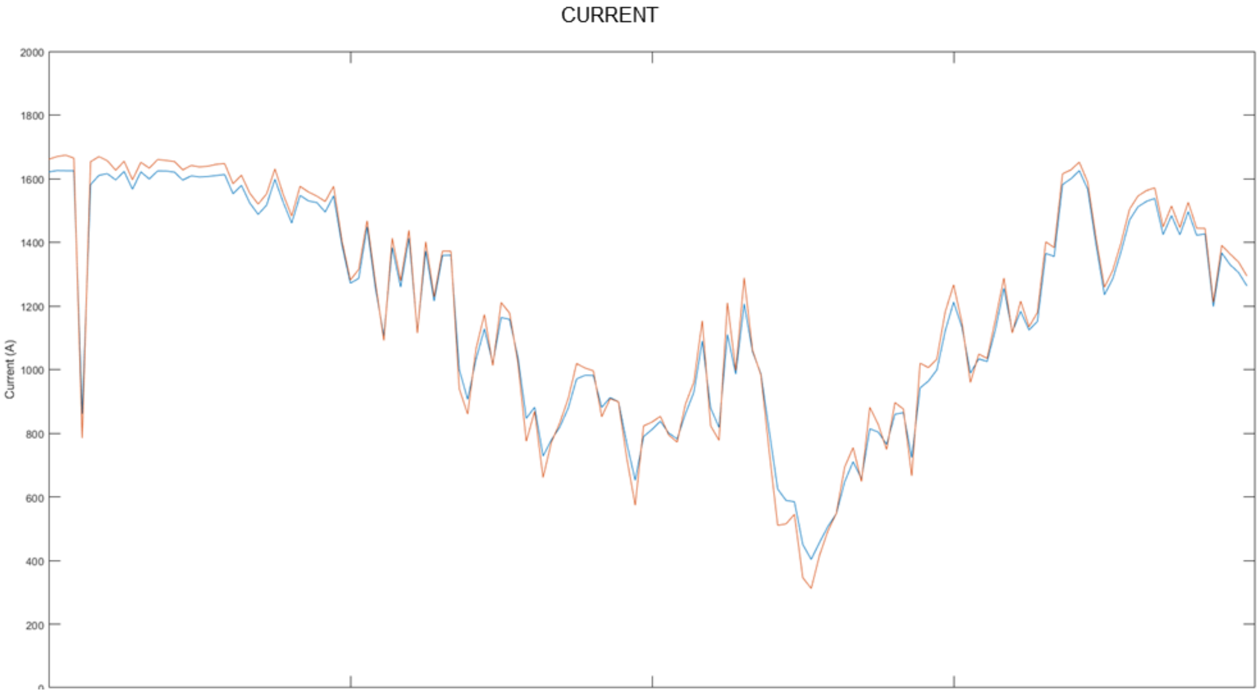


Figure 2: LLUC-1A-01-Normality Hybrid Digital Twin - Validation Results- Modelling quality- Current

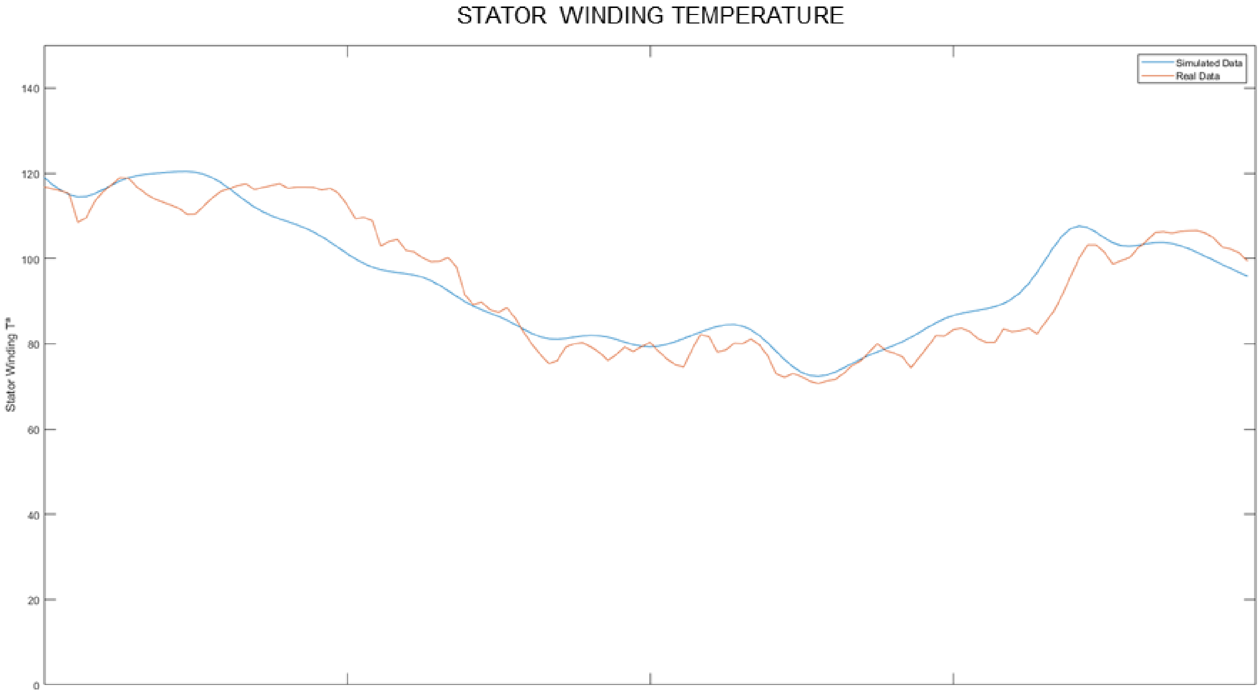


Figure 3: LLUC-1A-01-Normality Hybrid Digital Twin - Validation Results- Modelling quality- Stator Winding Temperature

Regarding the integration KPI, all the pipeline has been validated except IDS scenario where ENGIE is a data provider. This is due to a problem with Kubernetes Firewall that is being investigated (see more details in D6.2). Also, the semantic adaptation of results from TECN and VUB is missing. All this information will be included in V2 of this deliverable due in M36.

Regarding the fault detection KPI, the initial classifier developed in WP4 was validated and the results showed that it was only detecting anomalies but not failures. In fact, the selected parameters were selected to identify an over temperature. However, an overtemperature is not necessarily a symptom of a failure and could happen due to high ambient temperature conditions along with high wind speeds. The actual failure that ENGIE is interested in detecting is a “Generator Fan Failure”. In order to be able to detect the failure the classifier has been improved using as the condition indicator the difference between the real and simulated stator winding temperature. The modified classifier was tested for one of the wind turbines and showed that was capable of detecting an issue over one month prior to the failure.

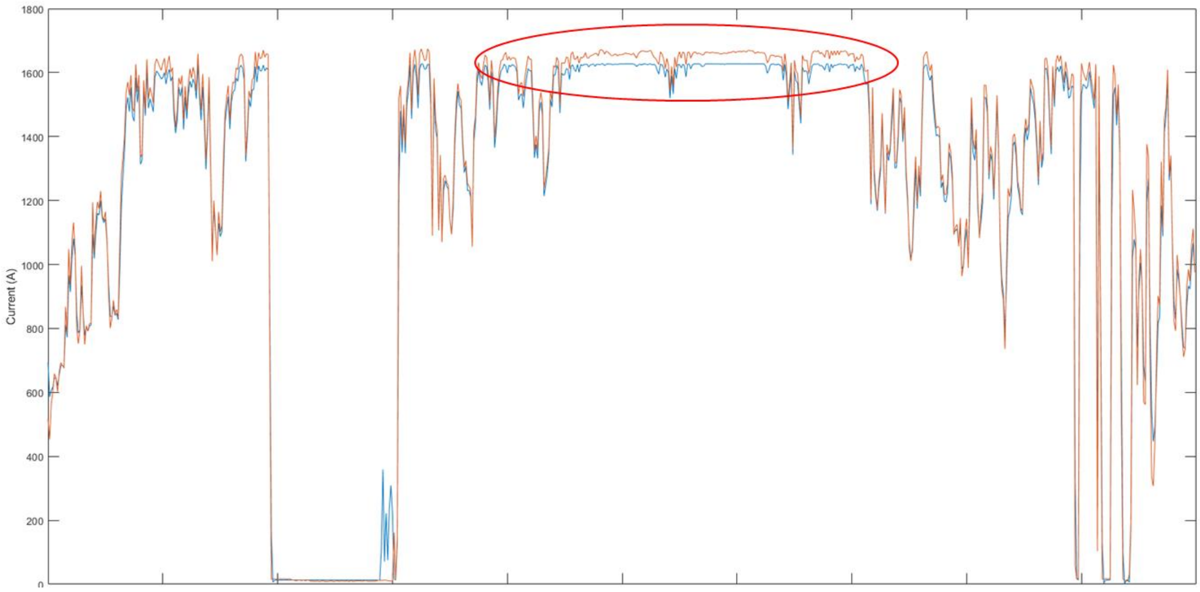


Figure 4: LLUC-1A-01-Failure Detection Classifier - Validation Results

However, from the failure identification date until the date when the fan was replaced the tool did not trigger anymore. This issue was validated with the operator who confirmed that there was a stator winding temperature sensor failure and the sensor was replaced. However, there were no further maintenance actions until the fan replacement. In addition, the developed data analytics tool was further validated with other labelled failures on other similar wind turbines and a similar pattern was observed. As a conclusion it was thought that the cause of the fan replacement could be due to a different reason of overtemperature (e.g. noise, vibration, etc.). ENGIE is currently checking this internally. On the other hand, TECN is currently improving the classifier to include other features which might help to identify a failure pattern. All this information will be included in V2 of this deliverable due in M36.

Regarding the processing capability, the full processing pipeline has been implemented (see D6.2) and is able to run on a standard server. The training part and dockerisation of the Digital Twin in Matlab showed a high consumption of RAM and CPU. However, once it is dockerised the execution phase is significantly less computationally expensive and can be run in any standard server with 4CPUs and 16GB RAM.

Regarding the pending aspects towards the end of the project is to validate with real data synthetic data generation and power converter RUL estimation tool.

2.2.1.2 Data driven approach

Table 2: LLUC-1A-01-KPIs evaluation- Data driven approach

KPI #	Description	Target Value	Actual Value	Comments
1	Modelling quality	3%	See Table 3	This goal can be considered accomplished given that the performance on all signals is close to or surpasses the target value and the fact that steady-state and transient are modelled together.
2	Integration	1	0.8	The different apps of the pipeline are well integrated.

				The integration has been tested thoroughly. There is still some work on the IDS. By the end of June the work on VUB side will be finalized. The full integration of the IDS depends on the progress speed of the relevant partners.
3	Fault Detection	Compared to the current failure detection the speed should improve with at least 25%, while keeping false positives below 10%	Accomplished	False positives can be held bellow the false positive threshold of 10% (see part KPI 3).
4	Processing Capability	Full processing chain for a farm should be able to run on a standard server.	Accomplished	The pipeline was validated on a standard server (see part KPI 4).
5	Maintenance costs reduction	10-20%	N/A	Cannot be calculated given the actual results and available maintenance data.
6	Availability increase	2-5%	N/A	Cannot be calculated given the actual results and available maintenance data.

To assess whether the KPIs were reached, five datasets from the ENGIE Senvion historical batch were used. These are linked to the following wind farms: FRCVE, FRPHA, FRHBA, FRSMV_FRKER and FRBRT. The historical batch also contained data of several other wind farms. However, due to large amounts of missing values, they could not be used for the validation.

KPI 1

The accuracy results presented here are for healthy steady-state and transient data combined. Improvements in the methodology have made it possible to model the steady-state and the transient behaviour of the wind turbines accurately with a single model. This means that the distinction between the two is no longer relevant. This is a major improvement over the original KPI.

Table 3: LLUC-1A-01-Model fit of the Anomaly_Detection app on healthy data from the FRCVE, FRPHA, FRHBA, FRSMV_FRKER and FRBRT wind farms

Turbine	Signal		
	TempGenBearing_1 (avg) (%)	TempGenBearing_2 (avg) (%)	TempStatorWind (avg) (%)
CV1	2.47	2.27	3.38
CV2	3.16	2.45	3.49
CV3	2.58	2.45	3.13
CV4	2.71	2.51	3.44
CV5	2.53	2.48	3.02
PH1	2.49	2.61	3.7
PH2	1.81	1.86	3.22
PH3	2.6	2.64	3.7
PH4	2.71	3.52	3.21
PH5	1.72	1.95	3.61
PH6	2.26	2.5	3.29
HB1	1.79	1.73	2.78
HB2	2.16	2.18	2.85
HB3	2.27	2.26	2.93
HB4	1.93	1.96	3.33
HB5	2.16	2.16	2.9
HB6	1.69	1.7	2.99
KER1	4.23	2.94	5.12
KER2	3.6	4.09	4.68
SMV1	5.82	6.15	3.94
SMV2	4.38	4.89	4.51
SMV3			
SMV4	5.36	6.43	4.36
SMV5	4.63	5.2	4.51
BRE1	2.56	2.1	3.84
BRE2	1.81	1.58	3.74
BRE3	2.05	1.74	3.93
BRE4	2.63	2.16	5.39
BRE5	2.23	2.02	4.85
BRE6	2.51	2.03	4.45
BRE7	2.67	2.33	4.82
BRE8	2.21	2.39	4.74
BRE9	2.74	2.55	4.5
BRE10	2.13	1.79	4.99
BRE11	2.6	2.19	4.41
BRE12	2.46	2.11	4.88
BRE13	1.82	1.41	4.54
BRE14	2.76	2.19	3.91
BRE15	1.93	1.61	4.43

Table 3 indicates that in general the 3% threshold is met for the modelling of the TempGenBearing_1 (avg) and the TempGenBearing_2 (avg) signals. For the TempStatorWind (avg) signal the accuracy is close to the 3% threshold. The fit for the FRSMV_FRKER wind farm has been also compared to the other wind farms showing worse results. The reason for this is under investigation and the conclusion will be reported as part of the V2 document due by M36.

KPI 2

All apps or tools that are part of the pipeline are fully integrated. They are all part of one package. The compatibility of the different apps has been tested thoroughly.

Implementation of the IDS Connector VUB app responsible for consuming and providing data keeps pace with the development of the IDS Connector. Currently, the pipeline can use the app to request data from our IDS Connector. By the end of June 2022, it will also be able to provide result data to the Connector. However, communication between connector of ENGIE can't currently be validated due to a problem the Connector has with a Kubernetes firewall which is being investigated by ENGIE. Also, the integration with other IDS components e.g., vocabulary manager or IDS Metadata registry is still pending.

KPI 3

Due to efficient training on the whole wind farm at once, the potential of transfer learning and the use of linear models, the training time could be reduced drastically. See KPI4 for an in-depth discussion of the speed of the applications.

The accuracy of the anomaly detection methodology is assessed using a confusion matrix. Several assumptions had to be made to be able to create this matrix:

- True Positive: if during the 6 months preceding a failure the health score increases substantially.
- True Negative: if during the 6 months following a failure the health score does not substantially increase.
- False Positive: If during the 6 months following a failure the health score substantially increases.
- False Negative: If during the 6 months preceding a failure no substantial increase in the health score is seen.

Table 4: LLUC-1A-01-Confusion matrix for failure identifications on FRCVE, FRPHA, FRHBA and FRBRT

		Predicted	
		Failure	No failure
Obs.	Failure	4	4
	No failure	0	8

Table 5: LLUC-1A-01-Performance metrics for failure identifications

Metric	Score
Balanced accuracy	0.83
F1	0.67

The confusion matrix gives us the results for all failure types combined. It indicates that the false positive requirement is obtained. This is because the algorithm has been deliberately made conservative. However, this results in an elevated false negative rate. The accuracy is 0.75, the balanced accuracy is 0.83, the F1 score is 0.67.
Figure 5 an Figure 6 show examples of correct detections of a rotor brush high temperature failure and generator bearing failure, respectively.

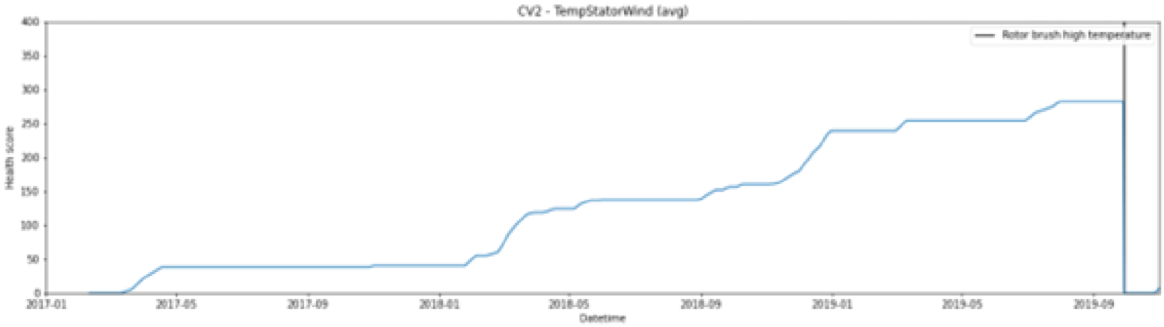


Figure 5: LLUC-1A-01-Example of a correct detection of a rotor brush high temperature failure

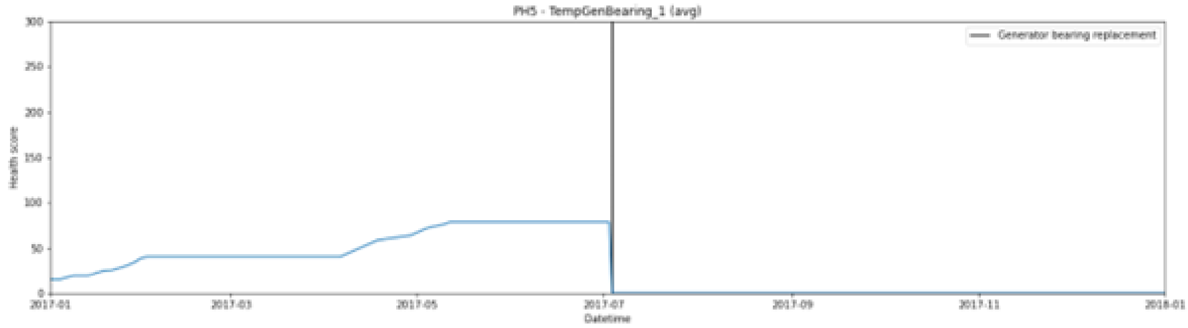


Figure 6: LLUC-1A-01-Example of a correct detection of a generator bearing failure.

KPI 4

The full data analytics pipeline has been tested using a VM that was assigned 20 cores on an server with an intel Xeon CPU (CPU type). The most time-consuming parts of the pipeline are the Anomaly_Detection and the SCADA_Data_Cleaner applications. The tables and figures below show the execution time and RAM consumption of the different applications. They show that the pipeline can easily be run on a standard server.

Table 6: LLUC-1A-01-Processing time and RAM consumption for the SCADA_Data_Cleaner app

SCADA_Data_Cleaner		
Wind farm	Time (s)	Max. RAM (GB)
FRCVE	132.27	7.8
FRPHA	155.41	8.2
FRHBA	148.23	7.2
FRSMV_FRKER	168.31	4.4
FRBRT	117.21	2.5

Table 7: LLUC-1A-01-Processing time and RAM consumption for the Anomaly_Detection app

Anomaly_Detector		
Wind farm	Time (s)	Max. RAM (GB)
FRCVE	908.73	see Figure 1
FRPHA	750.65	see Figure 2
FRHBA	953.76	see Figure 3
FRSMV_FRKER	799.15	see Figure 4
FRBRT	286.51	see Figure 5

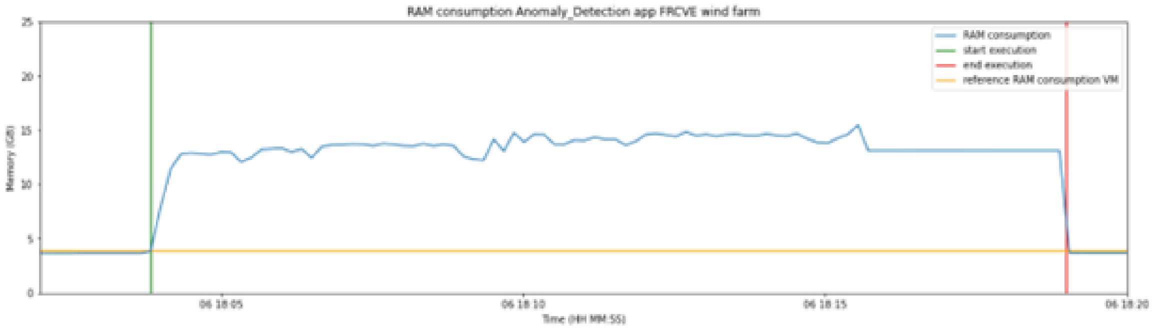


Figure 7: LLUC-1A-01-RAM consumption through time for the Anomaly_Detection app on the FRCVE wind farm

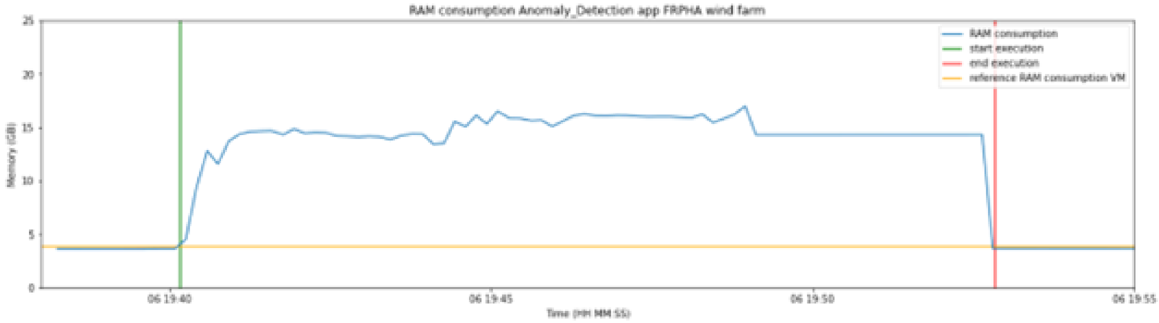


Figure 8: LLUC-1A-01- RAM consumption through time for the Anomaly_Detection app on the FRPHA wind farm

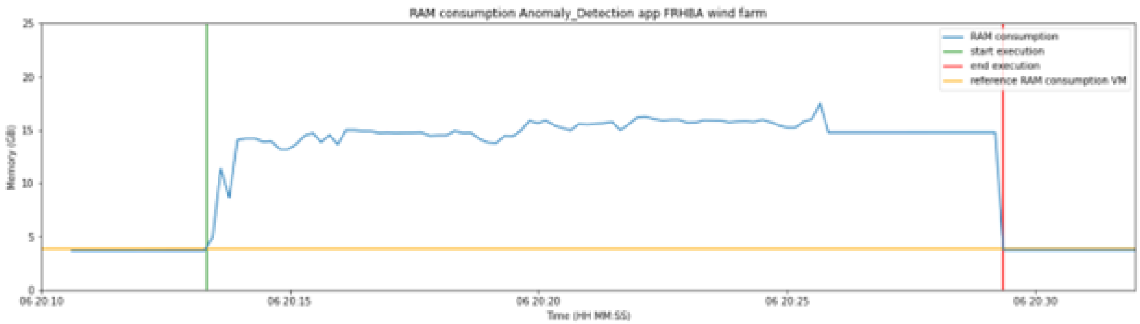


Figure 9: LLUC-1A-01- RAM consumption through time for the Anomaly_Detection app on FRHBA wind farm

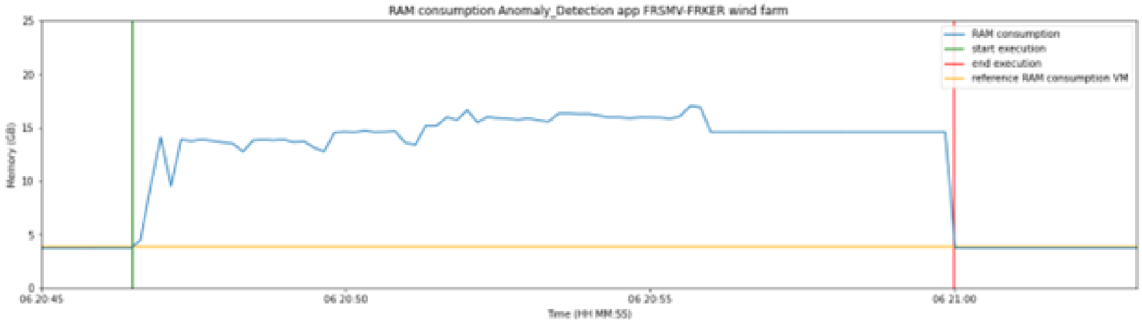


Figure 10: LLUC-1A-01- RAM consumption through time for the Anomaly_Detection app on FRSMV_FRKER wind farm

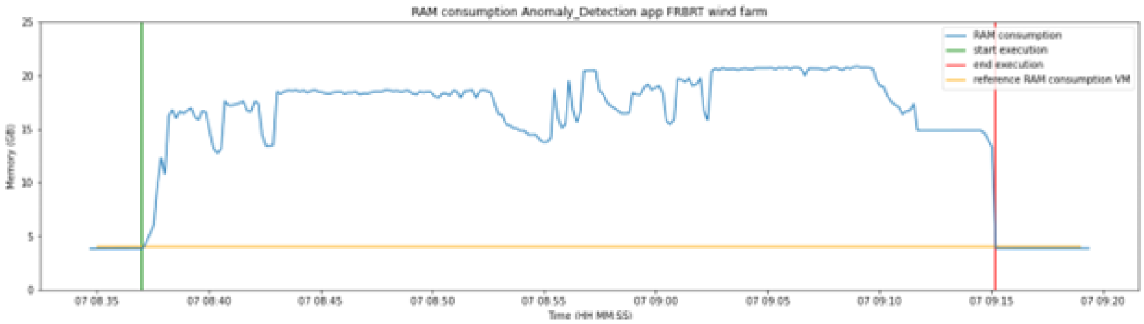


Figure 11: LLUC-1A-01- RAM consumption through time for the Anomaly_Detection app on FRBRT wind farm

Table 8: LLUC-1A-01-Processing time and RAM consumption for the Failure_Diagnosis app

Failure_Diagnosis		
Wind farm	Time (s)	Max. RAM (GB)
FRCVE	0.77	<1.0
FRPHA	1.14	<1.0
FRHBA	1	<1.0
FRSMV_FRKER	1.08	<1.0
FRBRT	1.64	<1.0

The execution time of the Root_Cause_Identifier app is 23.7 s on the FRCVE dataset. The RAM consumption is less than 1GB.

2.3 Conclusion

As a conclusion of the first validation, it can be drawn that the hybrid digital twin has reached the target KPIs regarding the modelling quality and processing capabilities. However, regarding the integration, the IDS scenario where ENGIE is a data provider and the semantic adaptation of results from TECN and VUB still need to be solved for V2. In addition, the classifier needs to be improved in order to try to meet the fault detection KPI. Furthermore, some of the KPIs related to maintenance costs and availability cannot be calculated given the actual results and available maintenance data. Finally, the validation of synthetic data and power converter needs to be completed.

For the data-driven approach developed by the VUB, KPI 1 is achieved as the performance of the models approaches or surpasses the target. On the other hand, KPI 2 is partially obtained as the different tools of the pipeline are integrated, however there is still some work on the IDS part. Besides, KPI 3 is obtained as the failure detection performance surpasses the KPI target. Finally, KPI 4 is obtained as the pipeline can easily be used on standard server equipment. Future work will focus on further improving the performance of the pipeline and completing the integration.

3. Pilot 2A Evaluation & Validation Report

3.1 Introduction

Pilot 2A focuses on electricity balancing at TSO level to ensure that total electricity withdrawals (including losses) equal total injections in a control area at any given moment (e.g. electricity production from solar and wind plants). This pilot is formed of 4 low level use cases:

- LLUC-03-Load Forecasting
- LLUC-04-RES Production Forecasting
- LLUC-05-RES effect calculation
- LLUC-07-PV Predictive maintenance

3.2 LLUC-2A-03-Load Forecasting

The objective of this use case is to provide a day-ahead forecast with and hourly resolution based on the previous 24-hour long hourly consumption data intended to be utilized by energy dispatch optimization engine.

3.2.1 Evaluation and Validation

Within 2a LLUC-03, day-ahead hourly load forecaster on national level have been developed. It was designed as innovative hybrid model, a combination of kNN and convolutional neural networks (CNN). The model obtains load from the previous day and current time-related parameters, and provides forecasted Serbian national load.

During the training phase, highly precise national load was obtained directly from Serbian TSO. Similar data could be found on ENTSO-E Transparency platform. Nevertheless, data that is being sent to the platform is not equally precise. Therefore, in order to validate the model with the same data that was used during training process, so that performance of the model is not jeopardized, similar data is gathered in batches by IMP. Hence, validation is carried out continuously, but on the delayed data.

Completely the same procedure with continuous validation could be applied to ENTSO-E or any other data source in the future.

Since forecasted and real national load are time series by their natures, common performance measurements have been selected as the most representative for the validation purposes. As given in Annex of this document, the list of the relevant KPIs for this service is following:

- Mean Absolute Error (MAE)
- Mean Absolute Percentage Error (MPAE)
- Root Mean Square Error (RMSE)
- Root Mean Square Error Percentage (RNSEP)

Example of the forecasting model output is given in Figure 12 left, whilst in the right hand of the figure the errors through forecasting samples are given. It could be noticed that maximal absolute error is a bit more than 200kW which is quite small, taking into consideration that total load is approximately between 2400kW and 4500kW. Additionally, the current values of the KPIs are given in the table below. Similarly to what was concluded in deliverable D4.4 Analytical Toolbox for Smart Grids, the model is quite precise and it could be utilized for the load balancing on the national level.

Table 9: LLUC-2A-03- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1a	Mean Absolute Error (MAE) [MW]	260	154	Calculated using old data that was available for the validation. Until the next report, new data will be obtained. Services for KPI evaluation developed and deployed.
1b	Mean Absolute Percentage Error (MPAE)	10	6	
2a	Root Mean Square Error (RMSE) [MW]	260	185	
2b	Root Mean Square Error Percentage (RNSEP)	10	6	

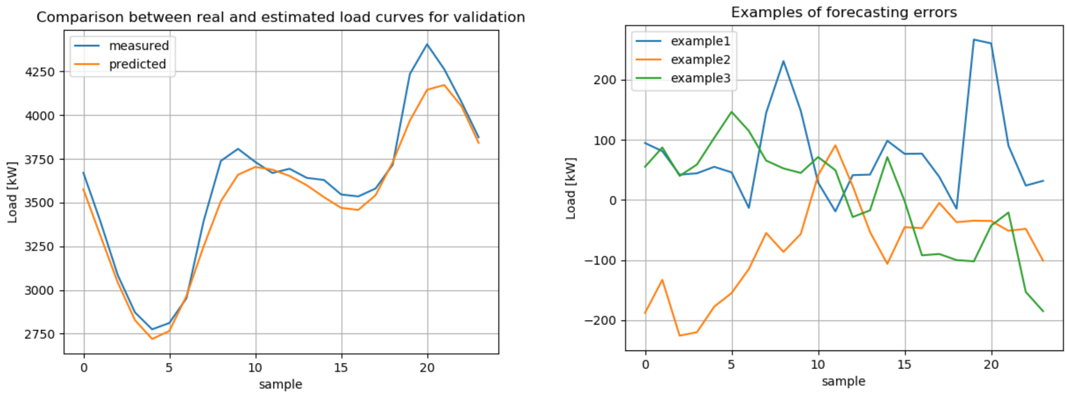


Figure 12 - Comparison between real and estimated load curves (left) and examples of forecasting errors (right)

3.3 LLUC-2A-04-RES Production Forecasting

The objective of this use case is to provide a day-ahead wind power production forecast with an hourly resolution based on forecasted weather conditions intended to be utilized by energy dispatch optimization engine.

3.3.1 *Evaluation and Validation*

As explained in deliverable D4.4 Analytical Toolbox for Smart Grids, a production forecaster based on LSTM neural networks has been developed, integrated and deployed. The example of different outputs of the production forecaster extracted from PLATOON MySQL is given in Figure 13. During the operation time, it was noticed that forecasting for some time was invalid, due to the fact that historical wind speed measurements corresponded to different height than the once from the WeatherBit service. Nevertheless, estimation of the wind speed at the considered height was evaluated as:

$$speed = speed_0/k * \ln{(h/h_0)}$$

and service was updated accordingly.

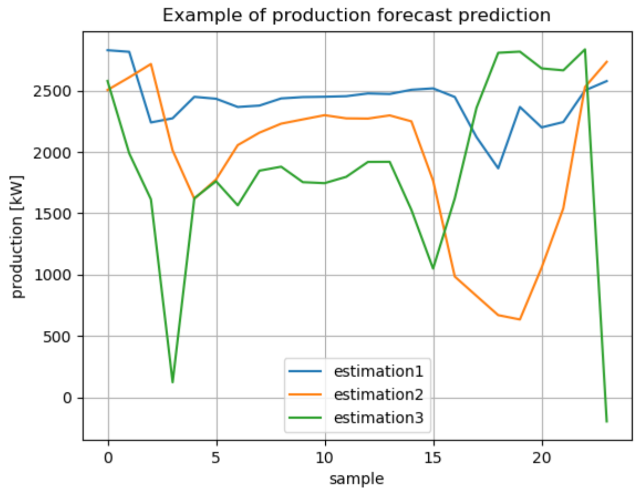


Figure 13 - LLUC-2A-04-Example of production forecast estimations

Since, forecasted and real production are time series by their natures, common performance measurements have been selected as the most representative for the validation purposes. As given in Annex of this document, the list of the relevant KPIs for this service is following:

- Mean Absolute Error (MAE)
- Mean Absolute Percentage Error (MPAE)
- Root Mean Square Error (RMSE)
- Root Mean Square Error Percentage (RNSEP)

Similar to the previous load forecasting service, production data is obtained in batches from Krnovo SCADA. Current KPIs could be seen in the table below and it could be observed that all KPIs are satisfactory. Namely, the model precision is relatively high, especially having in mind that the main input, wind speed, a highly fluctuating quantity, is only considered on an hourly basis.

Table 10: LLUC-2A-04- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1a	Mean Absolute Error (MAE) [MW]	260	139	Calculated using old data that was available for the validation. Until the next report, new data will be obtained. Services for KPI evaluation developed and deployed. Calculation continuous, but with delay, due to the fact that data will not be available in real time.
1b	Mean Absolute Percentage Error (MPAE)	10	8	
2a	Root Mean Square Error (RMSE) [MW]	260	159	
2b	Root Mean Square Error Percentage (RNSEP)	10	10	

As it could be noticed from Figure 14, the service is successfully working and storing outputs in PLATOON MySQL DB. As it could be noticed, last MySQL DB update was done on June 15th, when this deliverable was prepared.

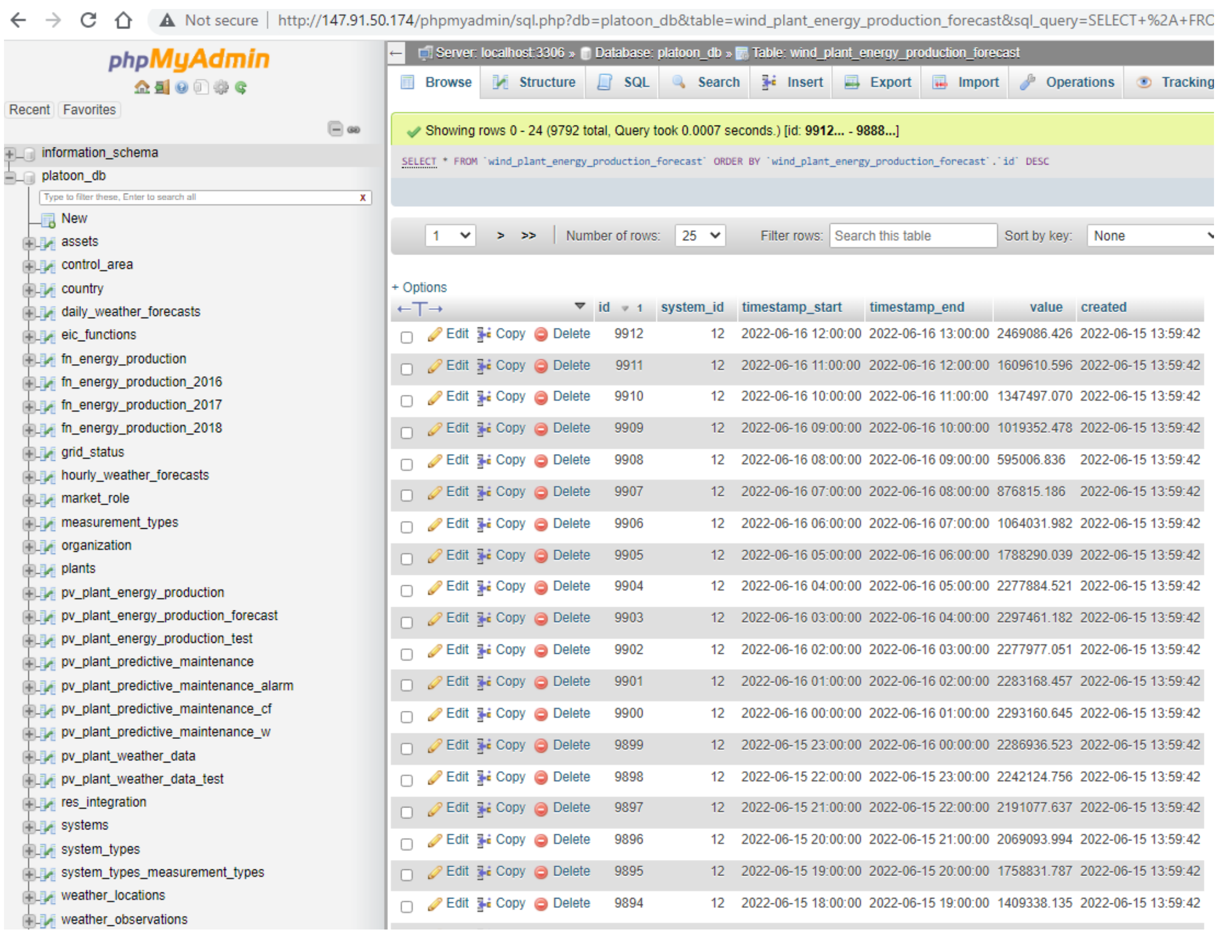


Figure 14 - LUC-2A-04- Illustration of production service filling in MySQL table

3.4 LLUC-2A-05-RES effect calculation

The objective of this use case is to analyse unexpected variations (voltage profile of the power system) before and after RES integration to the power system. Since the services need real-time data with high reporting rates of the grid status, a PMU is deployed at the Edge. In addition, analytics tools are also deployed at the edge.

3.4.1 Evaluation and Validation

Table 11: LLUC-2A-05- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Increase in PV insertion capacity	> 100 %	150 %	KPI is calculated daily, therefore the minimal value is only reported in this table.

The KPI was calculated in a few steps utilizing the data from the PMU and actual production of the installed PV (P_n = 50 kWp). The service first measures the grid with PV and estimates the state of the grid without the PV. The main goal is to estimate the impedance of the line towards the substation from the measurements. According to the calculated impedance and maximally allowed voltage on the LV grid defined by standard EN-50160, the maximum PV power can be estimated (see Figure 15). Then, for each day, the insertion capacity is calculated. However, in the case of bad weather, the results are not reported due to higher uncertainty, hence some missing values can be seen. Finally, the insertion capacity is normalized to the installed PV plant, which is 50 kWp resulting in a KPI higher than 150 %, with an average value of more than 200 % (see Figure 16).

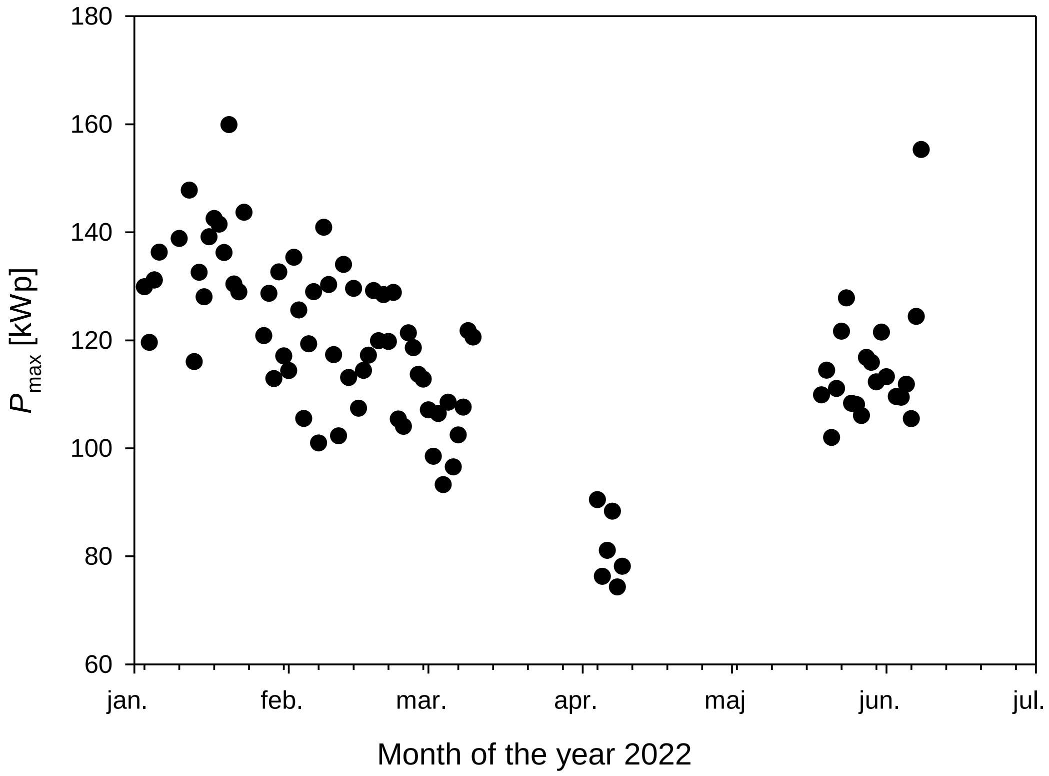


Figure 15: LLUC-2A-05- Maximal PV insertion capacity for the LV Grid.

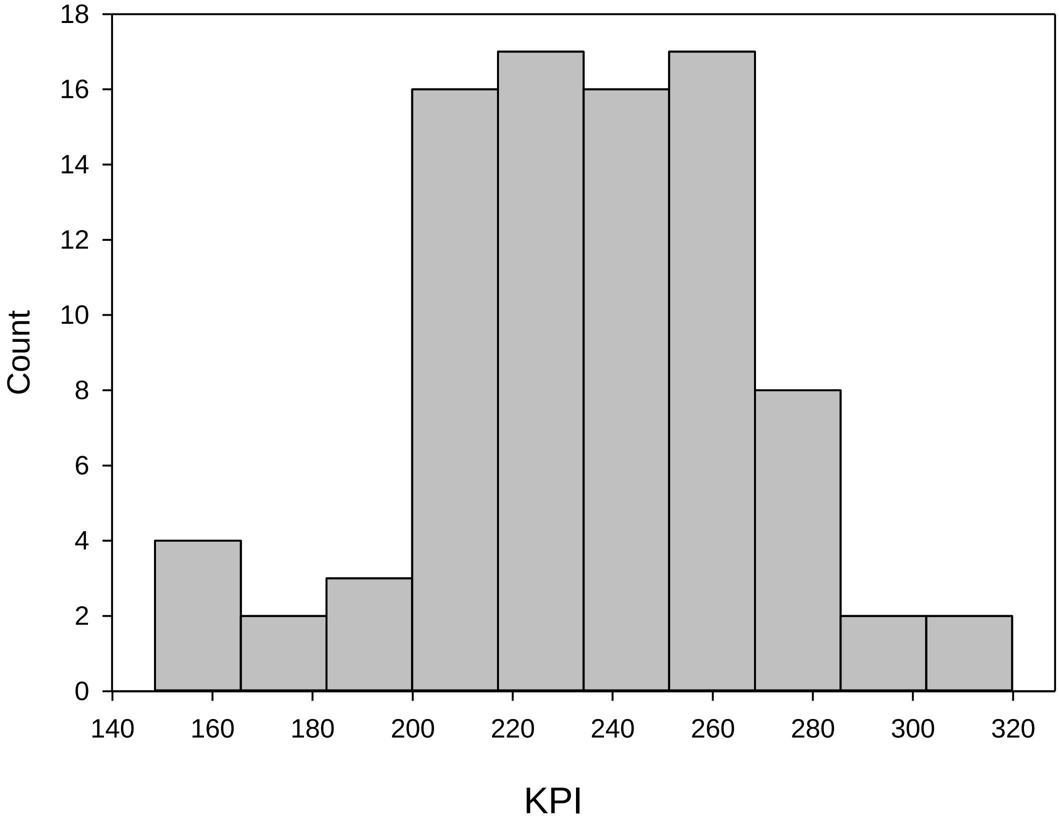


Figure 16: LLUC-2A-05- Histogram of calculated KPIs for the current year.

The service runs locally on the edge computer next to the existing PV plant using the edge-cloud framework as described in deliverable D4.2. The service is dockerized and executed once a day. The results are saved in the IMP SQL database on the central computer. In the future, we will monitor the execution of the service and analyze the reported results. This will be further evaluated to remove the results with higher uncertainty to get an even better insertion capacity estimation.

3.5 LLUC-2A-07-PV Predictive maintenance

The objective of this use case is to develop a set of data analytics tools that use existing data from sensors and whether to predict and monitor the degradation of the modules of PV plants.

3.5.1 Evaluation and Validation

Table 12: LLUC-2A-07- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Saving costs	> 0 €	10.95 € (estimation)	Cannot be calculated given the actual results and available maintenance data. However, an estimation is given as per explained below.

According to the KPI template, all the steps to calculate the KPIs are already done. However, since the modules are still in good condition with an estimated degradation of 1.1% (see Figure 17), the calculation has not yet triggered an alarm. The calculation is done for two types of failures: on the one hand, the estimation of the failure of inverters that are constantly monitored and for which the alarm is triggered when an anomaly is detected, and on the other hand, the performance of the modules which is evaluated once per day. During the deployment period, none of the alarms was triggered, so the KPI for cost savings cannot be calculated. However, an estimate can be made when the fault is detected assuming the following parameters:

- $N_{\text{days\_estimate}}$ = 3, the typical value if the manual inspection is performed periodically.
- $N_{\text{days\_after detectingfailure}}$ = 0
- E_{daily} = 73 kWh (one inverter, month May)
- Price = 0.05 €/kWh[i]
- KPI = 3*73 kWh*0.05kWh = 10.95 €

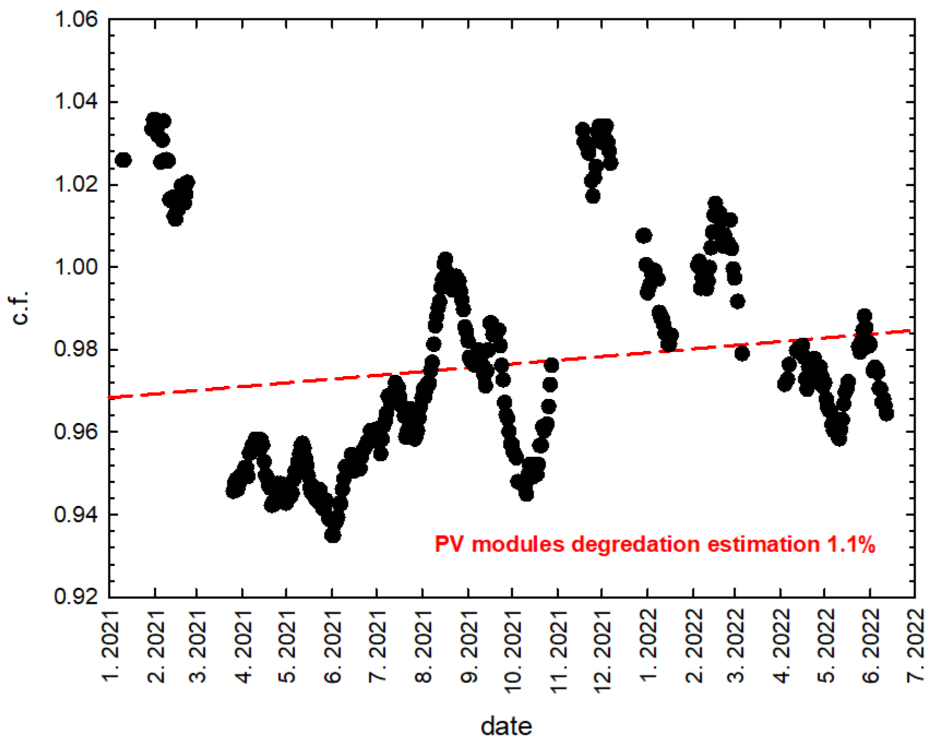


Figure 17: LLUC-2A-07- Daily calculated c.f. for PV plant installed at IMP and estimated PV module degradation.

The grid is monitored every 5 minutes, so the estimate of the number of days after the estimate of failure $N_{\text{days\_after detectingfailure}}$ = 0 is valid. The service is dockerized and deployed to the edge computer next to the PV system. The service checks the voltages and the symmetry of the inverter-related powers for all three phases (see Figure 18), and can immediately report an alarm about the failure into IMP MySQL.

```
if np.nanmean(P) > 3000:
    if np.nanmean(V1) < 180:
        write_alarm(now, "V1 issue")
    if np.nanmean(V2) < 180:
        write_alarm(now, "V2 issue")
    if np.nanmean(V3) < 180:
        write_alarm(now, "V3 issue")
    if np.nanmean(P1) < np.nanmean(P) / 5:
        write_alarm(now, "P1 inverter issue")
    if np.nanmean(P2) < np.nanmean(P) / 5:
        write_alarm(now, "P2 inverter issue")
    if np.nanmean(P3) < np.nanmean(P) / 5:
        write_alarm(now, "P3 inverter issue")
```

Figure 18: LLUC-2A-07- Part of code that constantly monitors the inverters and reports alarms to IMP MySQL

3.6 Conclusion

Taking into consideration previously presented results, all the services whose performances could be evaluated (LLUC 3, 4 and 5) are performing satisfactory, whilst similar behaviour is expected for LLUC 7, as well. The only underperformance was noticed in the beginning with production forecaster, due to difference between the historical and online input data. Nevertheless, the data was accordingly preprocessed and now the performance is satisfactory. During the next 6 months, until the final validation report, all the methods will be validated during different seasons (summer, winter…) and in case of performance degradation, they will be accordingly updated.

4. Pilot 2B Evaluation & Validation Report

4.1 Introduction

This pilot consists in two Use Cases related with the electricity grid stability, connectivity and life extension of the components in a smart grid in ParcBit, Majorca (Spain). The ese cases defined within this pilot are the following:

- LLUC-2B-01 Predictive Maintenance for MV/LV Transformers.
- LLUC-2B-02 Detection of NTL in electrical grids.

4.2 LLUC-2B-01 Predictive Maintenance for MV/LV Transformers

This use case focuses on transformer predictive maintenance, estimating transformer components health and its maintenance costs, planning maintenance actions, monitoring transformer alarms and studying different grid scenarios in case of replacing old transformers or adding complementary transformers.

4.2.1 Evaluation and Validation

Table 13: LLUC-2B-01- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Temperature estimation accuracy (%)	5%	0.23%	A validation of different virtual sensor algorithms with different features has been done and the one with the best results has been reported (see results below).
2	True positives (TP)	N/A	N/A	As no failures have happened this KPI is N/A.
3	False Positives (FP)	N/A	N/A	As no failures have happened this KPI is N/A.
4	False Negatives (FN)	N/A	3	As no failures have happened this KPI is N/A.
5	True Negatives (TN)	N/A	N/A	As no failures have happened this KPI is N/A.
6	Specificity (%)	N/A	100%	As no failures have happened this KPI is N/A.
7	Sensitivity (%)	N/A	N/A	As no failures have happened this KPI is N/A.
8	Cohen's Kappa (%)	N/A	1	As no failures have happened this KPI is N/A.
9	Savings (€)	N/A	N/A	As no failures have happened this KPI is N/A.
10	Additional Costs (€)	N/A	N/A	As no failures have happened this KPI is N/A.
11	Anticipation time (days)	N/A	N/A N/A	As there are no TP, FP or FN, the value of this metric is N/A.
12	Risk decrease (€)	N/A	N/A	As there are no TP, FP or FN, the value of this metric is N/A.
13	Maintenance costs savings (€)	N/A	N/A	As no failures have happened this KPI is N/A.
14	Useful Life Extension (years)	N/A	N/A	Not calculated yet.

Regarding the temperature estimation accuracy (%) of the top oil temperature virtual sensor, different algorithms with different features have been validated and a benchmarking analysis has been performed. The models of top oil temperature have been developed using distinct sensor configurations, going from low amount of necessary installed sensors to a configuration where all sensors need to be installed. This comparison allows future installations to decide the amount of investment on sensors depending on the required accuracy.

Table 14 shows the results of the models developed by SAM. The accuracy for each model is registered with %MAE and the test data comprises the 25% of the data that has been selected randomly.

Table 14: LLUC-2B-01- Sampol - Top Oli Temperature model results

Model	Train set MAE	Test set MAE
LV_load+Room_temp	2.07%	2.16%
LV_load+nearest_public_temp	3.55%	3.65%
LV_load+Room_temp+1Case_Sensor	1.30%	1.32%
Room_temp+1Case_Sensor	2.12%	2.42%

LV_load+Room_temp+1Case_Sensor+MV_load	3.55%	3.54%
LV_load+Room_temp +MV_load	2.12%	2.42%

The expected results are that the error decreases when using more sensors. However, the problem of the models that include MV load data is that the sensors have been installed and configured later than the rest. This implies that these models are trained with less data and obtains worse results than the rest.

Without installing the MV sensors, the best results are obtained when one of the case temperature sensors is installed. But in fact, the cost of this sensors must be considered due to the low difference of accuracy between models.

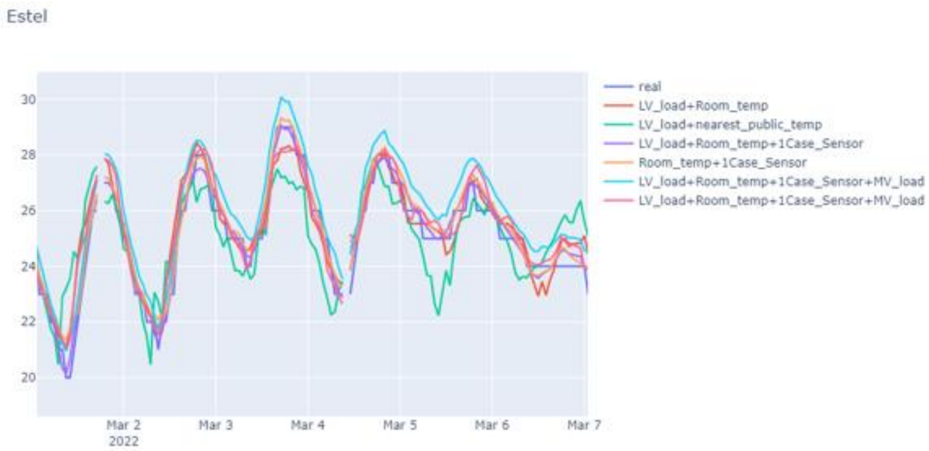


Figure 19: 2b-01 KPI 1 - Comparison of the temperature estimation between all trained models

On the one hand, regarding the models developed by Tecnalia the one that provided the best results provided a MAPE of 0.23%. The figure below shows the validation results of the best performing top oil temperature virtual sensor model over a period of 2 months (15/01/22 – 15/03/22). As it can be seen the predicted value (orange) is very close to the real value (blue). Also, it can be seen that there is a gap from 17/01/22 to 14/02/22 where there is no validation data. This was due to a problem with the integration of the current analyser in the primary winding of the transformer. For the final version of the deliverable (V2) the model will be validated with more recent data to confirm if the results are still comparable.

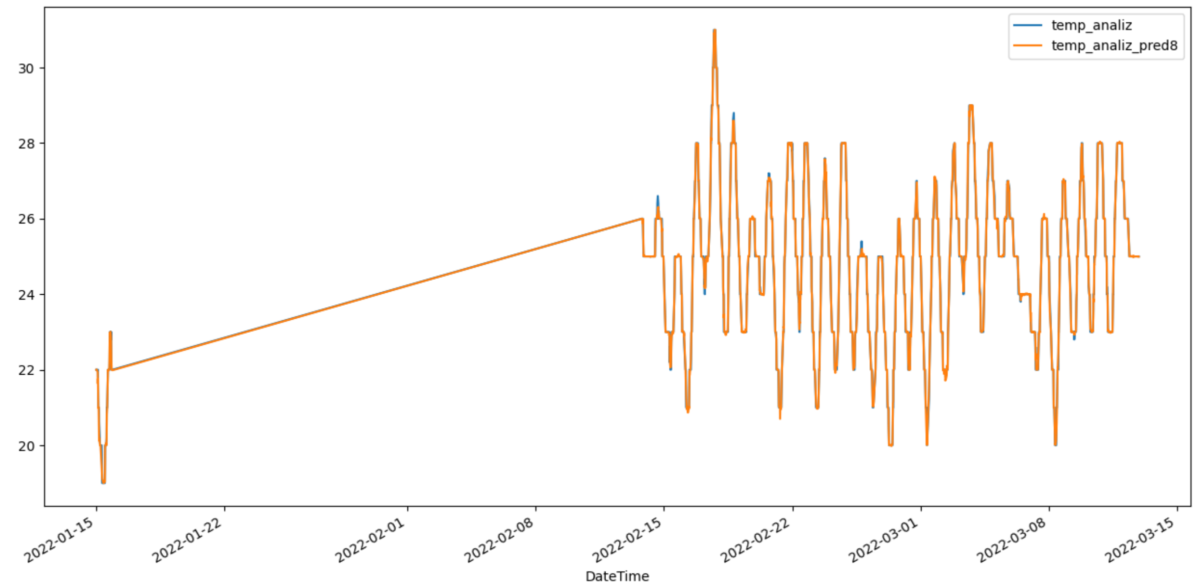


Figure 20 LLUC-2B-01- Top oil temperature virtual sensor best performing model validation results

Besides, regarding the predictive monitoring tools for electrical transformers, several functionalities regarding the and health-related issues that have been implemented so far. However, no real-time processing and validation has been done yet, so the application of the previously defined KPIs has not been possible until now. This document summarizes the evaluation and validation actions hitherto accomplished.

There is a fast model, (executing every 10 mins) and a slow model (executing every hour); the second one includes the last three signals from the table above. The models are trained using historical data (the training sample). Ideally, the dataset used as the training sample should cover the range of variation of those signals representing boundary conditions (i.e., ambient temperature or transformer power load).
As a result of the training, the tool displays a series of statistical indicators showing the accuracy of the estimation model in the case of every variable. These indicators include :

- ME (Mean Error)
- MAE (Mean Absolute Error)
- MPE (Mean Percentage Error)
- MAPE (Mean Percentage Absolute Error)
- MSE (Mean Squared Error)
- RMSE (Root Mean Squared Error)
- R2 (Correlation coefficient)

The following figure shows a sample of the results obtained for the previous accuracy indicators for some of the signals under analysis.

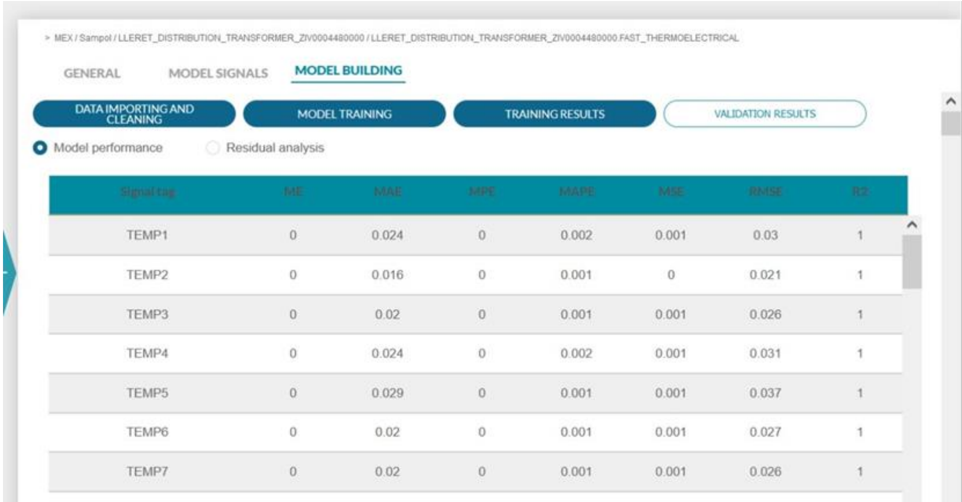


Figure 3 LLUC-2B-01- Training results for the predictive model training

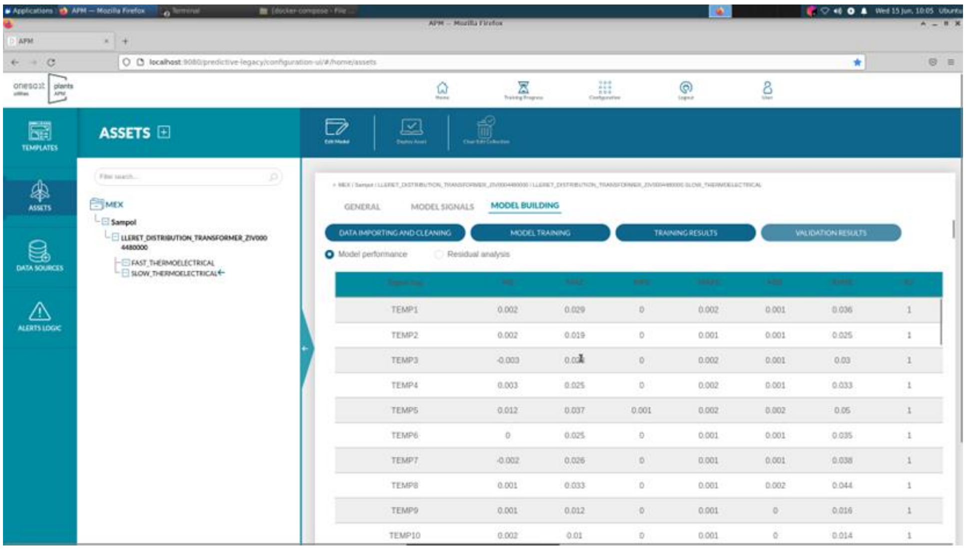


Figure 4: LLUC-2B-01- Validation results for the predictive model training

The results obtained in the training of the predictive model are quite good, thus anticipating a suitable fit in the future operation with real-time data. In addition, 30% of training selected points are randomly excluded from model training and used in an automatic validation. Results are similar to those in model training.

As far as the health-index related modules are concerned (replacement and overload calculations) the situation is similar since no real-time operation has been implemented yet. The modules have been designed and programmed and currently have undergone unitary tests, covering:

- Data model and conversion to different data formats (serialization and deserialization to protocol buffers or json)
- Data base interactions (get, insert, delete and update)
- General service (protocol buffer queries and serialization)
- Calculation functions

These tests have been run with actual data and compared with manual calculations according to each standard (CNAIM and IEC60076-7). Manual calculations have been previously validated with examples contained on standard definitions. For results testing purposes, calculations of each standard have been separated in several functions:

- **IEC60076-7:**
    - Theta0
    - ThetaHs
    - DeltaH1
    - DeltaH2
    - Aging factor
    - Actual Relative S to nominal apparent Power
    - Overaging
    - Whole combined calculation
- **CNAIM:**
    - Location factor
    - Duty factor
    - Observed condition
    - Oil Condition
    - Dissolved Gases Condition
    - Furfurladehyde oil condition
    - Initial Health score
    - Estimated health score along time
    - Financial consequences factor
    - Financial estimation of time to change
    - Whole combined calculation

The accuracy of calculation software tests is over 99%.

4.3 LLUC-2B-02 Detection of NTL in electrical grids

The main objective of this use case is to develop a tool for the quantification of losses in the distribution grid of a DSO and the detection of non-technical losses (NTL), using the available smart meter data.

4.3.1 Evaluation and Validation

KPI #	Description	Target Value	Actual Value	Comments
1	Global Losses Energy Percentage	<15%	33.03%	Calculated with the synthetic data, average between all loops.
2	NTL Energy Percentage	5%	21%	Calculated with the synthetic data, average between all loops.
3	TL Energy Percentage	<10%	12.03%	Calculated with the original data using periods where all smart meters were registered.
4	Customer NTL Energy Percentage	<10%	100%	Calculated with the synthetic data, average between all loops.

5	Non-customer NTL Energy Percentage	-	0%	As there are no anomalies detected and the synthetic data does not introduce this type of NTL its value is 0.
6	True positives (TP)	-	27	Number of anomalies detected correctly.
7	False Positives (FP)	-	306	Number of anomalies detected not generated at synthetic data
8	False Negatives (FN)	-	50	Anomalies not detected by the algorithms.
9	True Negatives (TN)	-	1276	Normal behaviour data with no anomalies detected.
10	Specificity (%)	-	0.3421	The algorithm does detect only a 34% of the anomalies.
11	Sensitivity (%)	-	0.8065	The algorithm classifies 80% non-anomalous smart meters correctly.
12	Cohen's Kappa (%)	-	0.0571	It is a low value of kappa. But can still be valid until there is no stablished limit.
13	Economic Savings (€)	-		

As there are no real anomalies classified in the past, to validate this use case, a synthetic anomaly generator has been developed. It is based on the work done at [ii].

The synthetic data has been generated in a loop until 71 anomalies are generated (this number is calculated using Cochran technique [iii] to assure that the results are statistically significant). This strategy avoids the introduction of too much simultaneous anomalies. It is not expected that a big number of prosumers starts developing fraud during the same period. Each loop follows the next steps:

1. The percentage of anomalous smart meters is selected to be between [5%,10%] of the total number of smart meters registered.
2. Each anomaly starts in a date randomly selected from the range [2022-01-01, 2022-06-01].
3. The anomaly type is selected randomly between shunt and Interrupt shunt.
4. The effect of the shunt is randomly selected from the range [25%,85%].
5. The anomalies created using the interrupt shunt technique, are generated with an interrupt coefficient selected randomly from the range [50%,90%].
6. Detect the anomalies using the synthetic data.
7. Clear anomalous noise and go back to step 1.

The improvement of energy losses evaluates the reduction of technical energy losses in the distribution network. The Platoon project is not addressing the actions that can be implemented in order to minimize them, but the objectives are aimed at the implementation of new or improvements in existing algorithms for their detection and identification with the evident purpose of designing subsequent actions for their reduction (outside the project scope).

Another objective is to deploy these algorithms at the node level so that the detection of these possible losses is carried out at the local level, thus minimizing the volume of data that must be sent to the central systems for its calculation, the capacity of processes in these central systems therefore the necessary calculation times.

4.4 Conclusion

Preliminary tests have been conducted both in the case of the predictive module training and unit tests for the health-related modules. The results of the top oil temperature virtual sensors are successful. However, the model has been validated only with 1 month worth of data. Thus, for V2 of this deliverable it must be validated with more recent data to confirm that the results are valid. Nevertheless, the unavailability of real-time results for predictive and health-related calculations does not allow for final tests and KPIs to be properly calculated and displayed. This needs to be completed for V2 of the deliverable.
Regarding the validation of LLUC-2B-02 NTL detection use case, due to the lack of fraud data the developed models have been validated using synthetic data.

5. Pilot 3A Evaluation & Validation Report

5.1 Introduction

Pilot 3a is related to the ENGIE Lab CRIGEN building office located in the Paris region. The office has a Building Management System (BMS) controlling the HVAC and comfort in different zones of the building. Two low level use cases have been developed within the scope of this pilot:

- LLUC-3A-01-Optimizing HVAC control regarding occupancy.
- LLUC-3A-02-Provide demand response services through building inertia and HVAC controls.

5.2 LLUC-3A-01-Optimizing HVAC control regarding occupancy

This use case aims to provide an optimized operation schedule for each day of the week for the office building and its different zones based on the occupancy in the building and the comfort level required by the occupants. The HVAC optimization and control aims to:

- Optimize the building energy consumption.
- Maximize the comfort of occupants with the best energy efficiency.
- Automate HVAC system control and reduce manual intervention on system controls.

5.2.1 Evaluation and Validation

An initial implementation has been conducted to collect the data required on the platform, implement the pipelines and a first version of the tools to produce the optimized controls that should be sent to the Building Management System for optimization.
Some update of the tools and the data pipeline are still ongoing to have the different bricks required to run efficiently the use case.

Different challenges were encountered in the implementation of the use case and required update and extra work on the use case:

- Quality of the input data and extra treatment required to assess the occupancy of the different zones in function of IT data connexion. Some extra data treatments were required and still must be implemented.
- Heating and cooling of the building proved to operate not properly / exactly as expected. For heating and cooling, the local regulation was not often used in certain rooms (heating/cooling through ventilation and internal heat gain for example). Some of the controllers have

schedule/setpoint problem that didn't really fit with a normal operation of the building. The data collected for the Data analytics model is then not so relevant.
- Challenges of managing heating and cooling (differences in the data), building operating both in heating and cooling at the same time with a regulation that is not optimal.

In addition, it is not yet possible to send orders to the BMS as some protocol and technical difficulties need to be solved to send setpoints plannings to the different controllers. There is currently an important focus to tackle this subject quickly and to be able to fully test the use case on the building.

The following KPI calculation were implemented on the platform for calculation. It is still not possible to properly assess the results of the KPI since the tools training models had to be updated with new datasets, but we were able to check the validity of most of the calculation (except KPI2).

Table 15: LLUC-3A-01- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Deviation to target comfort during occupancy time	0.5°C to comfort range	0,45	Calculation validated but update needed for meaningful results
2	Unnecessary HVAC heating emission	<10%	-	Calculation validated but update needed for meaningful results
3	Unnecessary HVAC cooling emission	<10%	-	Calculation validated but update needed for meaningful results
4	Gain on heating consumption	>10%	-	Calculation validated but update needed for meaningful results
5	Gain on cooling consumption	>10%	-	Calculation validated but update needed for meaningful results

<u>KPI1 UC1: Deviation to target comfort during occupancy time</u>

{"result":0.4541737476634559,"completionTime":"20211117 10:36:22","startTime":"20211117 10:35:36"}

Execution Schedule	00 19 * * *
Notes	The result was verified to be within acceptable ranges for the following dates: Nov9 Execution : 24hrs Data of 8.11.2021 Result: 1.158 Nov10 Execution : 24hrs Data of 9.11.2021 Result: 1.158 Nov11 Execution : 24hrs Data of 10.11.2021 Result: 1.158 Nov12 Execution : 24hrs Data of 11.11.2021 Result: 1.158 Nov13 Execution : 24hrs Data of 12.11.2021 Result: 1.158 Nov14 Execution : 24hrs Data of 13.11.2021 Result: 1.158 Nov15 Execution : 24hrs Data of 14.11.2021 Result: 1.158 Nov16 Execution : 24hrs Data of 15.11.2021 Result: 1.158 On laptop executed for smaller data set (1hr) to be verifiable manually; (data can be extract for BMS and occupancy for this

	duration or access through Python if further verification is needed) Start time: 2021-09-20 15:00:00.000000" End time: 2021-09-20 16:00:00.000000" Result: 0.454

KPI2 UC1: Unnecessary HVAC heating emission

Result	Not verified
Notes	KPI is deployed and producing results on ES. Results are not in acceptable range

KPI3 UC1: Unnecessary HVAC cooling emission

Result	Verified
Notes	KPI is deployed and producing results on ES. Results are in acceptable range provided no optimization.

KPI4 UC1Gain on heating consumption

Result	Verified
Notes	Result cannot be verified but we see that scheduled on platform and results are produced. Later with more data result will be verified.

KPI5 UC1: Gain on cooling consumption

Execution Schedule (current)	30 04 * * *
See Daily Execution Results	<not available yet>
Result	Verified
Notes	Results are now available on dashboard, to be verified.

The figure below shows a screenshot of the dashboard with the different KPIs:

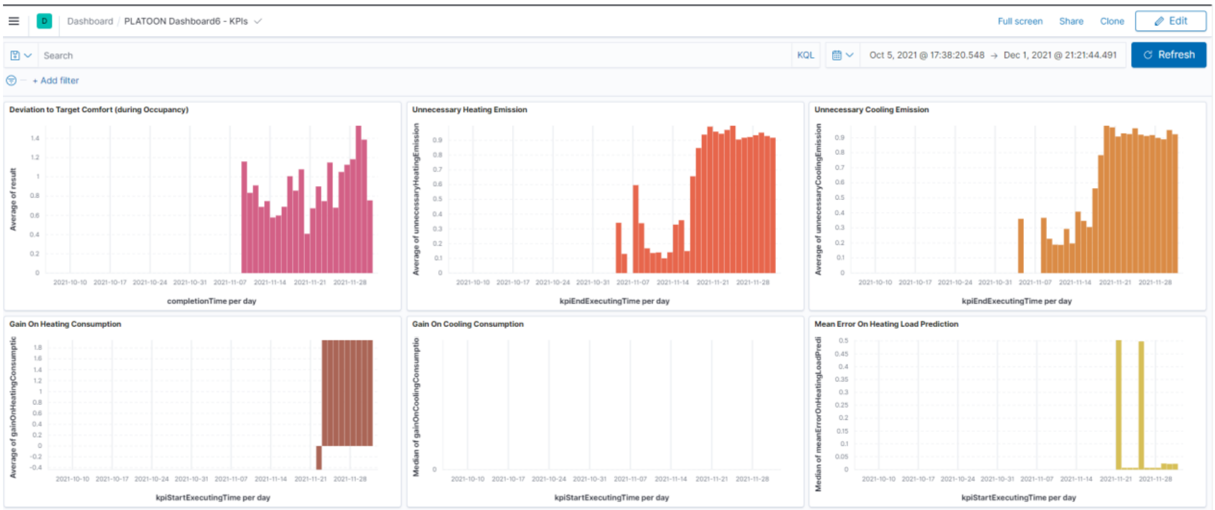


Figure 21 LLUC-3A-01- KPIS dashboard

5.3 LLUC-3A-02-Provide demand response services through building inertia and HVAC controls

The use case intends to provide a smart module to supervise the implementation of Demand Response services in an office building using HVAC control and building inertia. This use case aims to:

- Provide flexibility services to contribute to the grid balance (helping to reduce peak demand on the grid)
- Provide accurate predictions of the flexibility available for the next day to help the aggregators to evaluate the Demand Response services provided on the market
- Generate income by contracting with an aggregator

5.3.1 Evaluation and Validation

A first implementation of the tools has been realized with the cooling data of 2021 but it was not really possible to test its implementation on cooling without being in the summer season (that is just starting). Some developments are still ongoing for the implementation of the results on the heating part.

Some challenges were encountered on two levels:

- The operation of the cooling system feeding the cooling network of the building present a lot of on/off cycle probably due to oversizing. In this condition it is quite difficult to precisely model the electricity consumption with the start and stops of the system.
- The data regarding heating consumption from the gas counter is updated every hour which limits the resolution of the output data of the model. In fact, it is only possible to predict the energy consumption every hour.

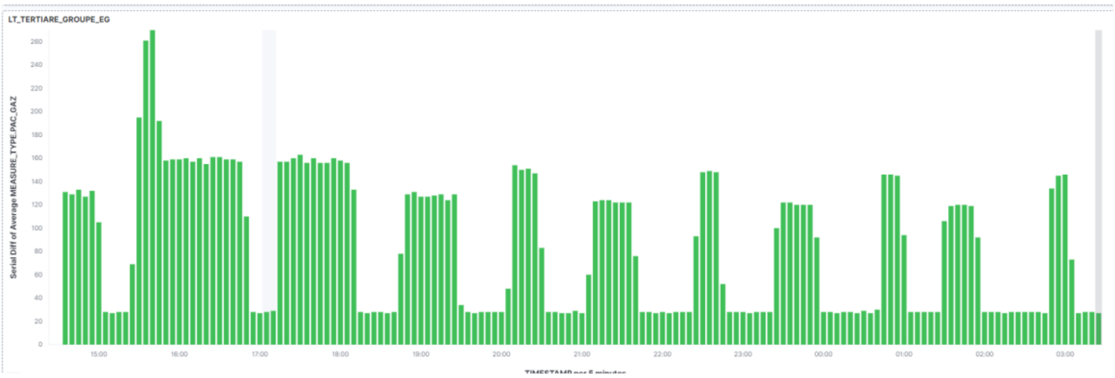


Figure 22 LLUC-3A-02- Energy consumption of the cooling system

As a consequence, the prediction on energy consumption is difficult to do at the 30min time step as initially planned due to the 2 problems mentioned above.

Furthermore, some of the tools developed for the energy consumption prediction are based as well on the tools of the first low level use case, especially the one regarding the occupancy prediction. Updates on these tools are needed to run more accurately the predictions.

The different KPIs for this low-level use case have been implemented on the platform and their output calculation were verified. However, it is still needed to wait for the right period (summer for cooling) and the new implementations (for heating) to be able to assess the business relevance of the KPI.

Table 16: LLUC-3A-02- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Mean error on heating load prediction	Error <10%	-	Tools still to be implemented on heating
2	Mean error on cooling load prediction	Error <10%	-	Not really tested yet, waiting for the 2022 cooling period and updates on the tools of the LLUC-3A-01.
3	95-percentile error on heating load prediction	<20%	-	Tools still to be implemented on heating
4	95-percentile error on cooling load prediction	<20%	-	Not really tested yet, waiting for the 2022 cooling period and updates on the tools of the LLUC-3A-01.
5	Error on the flexibility prediction	Error <10%	-	Tools to be implemented
6	Mean error on HVAC load prediction for days with load shifting programs	Error <10%	-	Tools to be implemented

KPI1 UC2: Mean error on heating load prediction

Result	Validation in progress (tool 3 output is now integrated)

Notes	KPI are calculated and look OK but further business validation is needed. Output of tool3/4/5 are not very relevant to reach a conclusion yet. Tool3 output is now integrated instead of simulated data. Reassessment in the heating period is still needed.

KPI2 UC2: Mean error on cooling load prediction

Result	Validation in progress (as tool 3 output is now integrated)
Notes	Tool3 output is now integrated instead of simulated data. Reassessment in the heating period is still needed.

KPI3 UC2: 95-percentile error on heating load prediction

Result	Validation in progress (as tool 3 output is now integrated)
Notes	Tool3 output is now integrated instead of simulated data. Reassessment in the heating period is still needed.

KPI4 UC2 : 95-percentile error on cooling load prediction

Result	Validation in progress (as tool 3 output is now integrated)
Notes	Tool3 output is now integrated instead of simulated data. Reassessment in the heating period is still needed.

KPI5 UC1 Error on the flexibility prediction

Result	Verified
Notes	Results are now available on dashboard.

The figure below show a screenshot of the dashboard with the different KPIs:

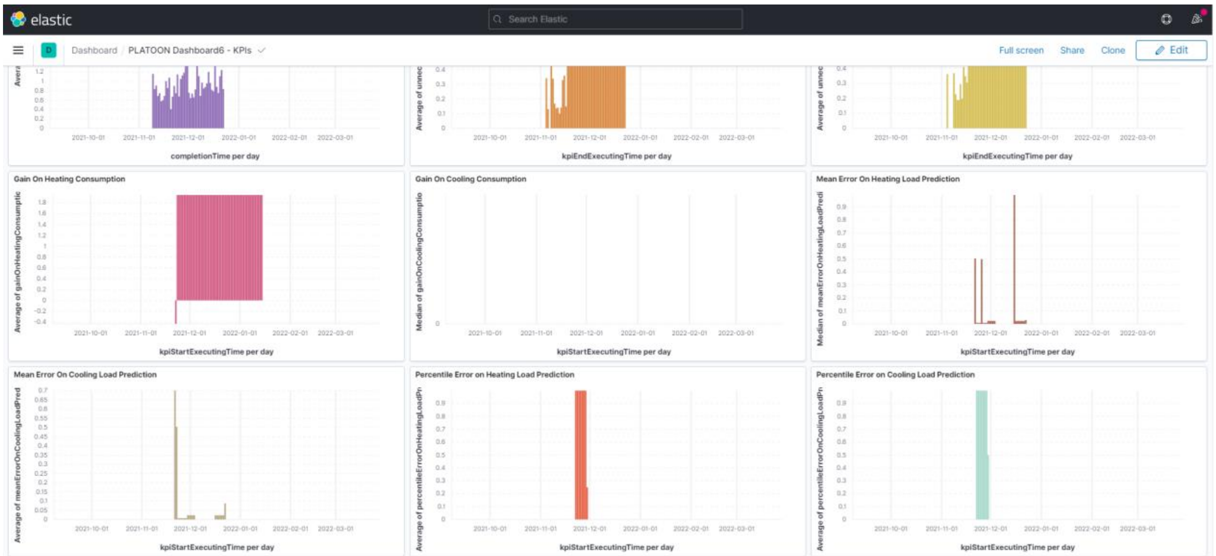


Figure 23 LLUC-3A-02-KPIS dashboard

5.4 Conclusion

With different challenges encountered on the implementation of the use case, there is still ongoing work to get some relevant results and to be able to assess the KPIs. The business validation of the KPI result will be realized for the last validation report that will be delivered on M36.

6. Pilot 3B-PI Evaluation & Validation Report

6.1 Introduction

The scope of the Pilot is to create a new way to work in order not only to optimize energy usage and identify behaviours to be changed, but also as an opportunity to reduce maintenance and service interruptions through a better usage of cooling / heating and lighting systems and use Augmented Intelligence algorithms for anomaly detection in HVAC plants. The type of data used in the pilot span from internal consumption data and plants performances to comfort targets managed by user together with external information related to weather forecasts and real time conditions.
The Pilot 3B-PI includes the following use cases:

- LLUC-3B-PI-01- Building Heating & Cooling consumption Analysis and Forecast
- LLUC-3B-PI-02 – Anomaly Detection of cooling & heating plants
- LLUC-3B-PI-03 - Lighting Consumption Estimation & Benchmarking

6.2 LLUC-3B-PI-01- Building Heating & Cooling consumption Analysis and Forecast

The use case focuses on efficiently forecasting and benchmarking of energy consumption to reduce costs and emissions and improve the comfort of the working environment. For optimization of both cooling and heating systems, it is important to correlate the energy consumption with the occupancy (based on number of employees and clients), as well as to benchmark with similar buildings.

6.2.1 Evaluation and Validation

Table 17: LLUC-3B-PI-01- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Deviation between actual and forecasted energy consumption	+/- 5%	[SB] 110% [MO] 18% [L102] 15 %	The value of the KPI given here is Weekly (YYYY-WW) for each building averaged by cluster (102, Multi-hourly, SB) in order to give a general vision. The provided results correspond to the last week before the deliverable writing.

2	Energy consumption gap of a building with itself during the time (year)	+/-10%	From –4% to +2%	The KPI meets the target value (see Figure 26).
3	Energy consumption gap of a building with itself during the time (short period)	+/-10%	+/- 24 % (calculated on a building taken as sample)	The results don't meet the target value but this doesn't mean that the corresponding data analytics tool is not correct. We are investigating the reasons that led to a spike in consumption in the last week
4	Benchmark of a building energy consumption with a cluster of similar buildings	+/-10%	7%	The current method of calculating the KPI doesn't take into account the size of the buildings, which, even if they are part of the same cluster, can be very different and thus provide a misleading KPI.
5	CO2 emission reduction	≥ 10%	--	The KPI is of business type and the value is calculated for each Building. The KPI is calculated annually, so the final value is still expected. The values represented by the KPI therefore indicate the trend of energy consumption during the year with the aim of monitoring its trend.

KPI-01 calculates the deviation (%) between the energy consumption forecast and the actual consumption in the building. The scope is to measure the effectiveness of the predictive model.
The result obtained by the KPI highlights that there is a significative difference in the average values calculated according to building category (Smart Building, Multiorarie, DL102). This depends on the granularity of available data for each cluster of buildings. In fact, the algorithm is based on learning mechanism that refine itself as the information is enriched.
In addition, the different clusters have very different characteristics in terms of energy infrastructure and configuration that may affect the prediction criteria. Based on these considerations the model for KPI-01 needs to be refined.
The following picture shows how the deviation between the blue (actual energy consumption value) and the orange lines (forecasted value) is very small. This indicates a good forecasting ability of the algorithm in case of building that has a data history from the year 2018 onwards. At points where the deviation is very significant, it is necessary to investigate whether this is a false negative result.

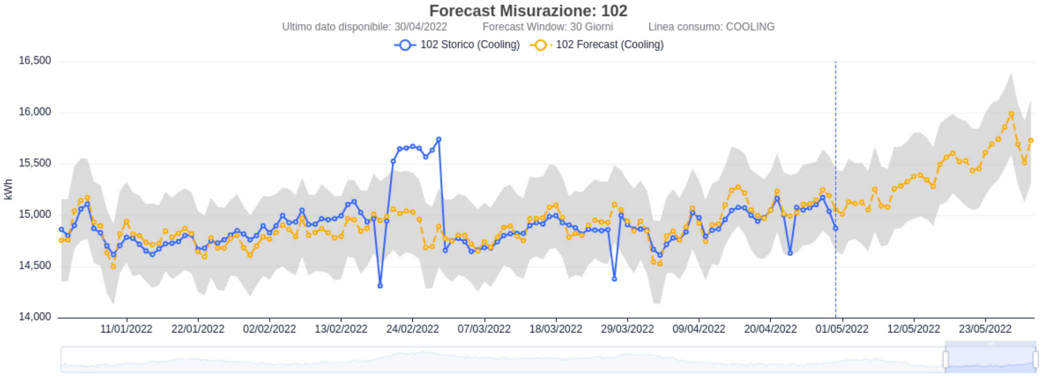


Figure 24: LLUC-3B-PI-01 KPI-01 Energy Consumption forecasting

Figure 25: LLUC-3B-PI-01 KPI-01 Energy Consumption forecasting Data Log

Figure 26 shows the result of the KPI 02 on consumption trends of a building over time. The business KPI is validated because the results are consistent with the data analysed and allows consumption to be kept under control and provides useful information for decision support to those involved in defining efficiency strategies or managing buildings.

The green lines indicate the limits of the target value. In the example shown, it is evident that the building's consumption, although having a non-constant trend, is within acceptable consumption limits. The information that is provided shows that, in the first 3 periods of the year, heating consumption consumes less than the calculated budget.

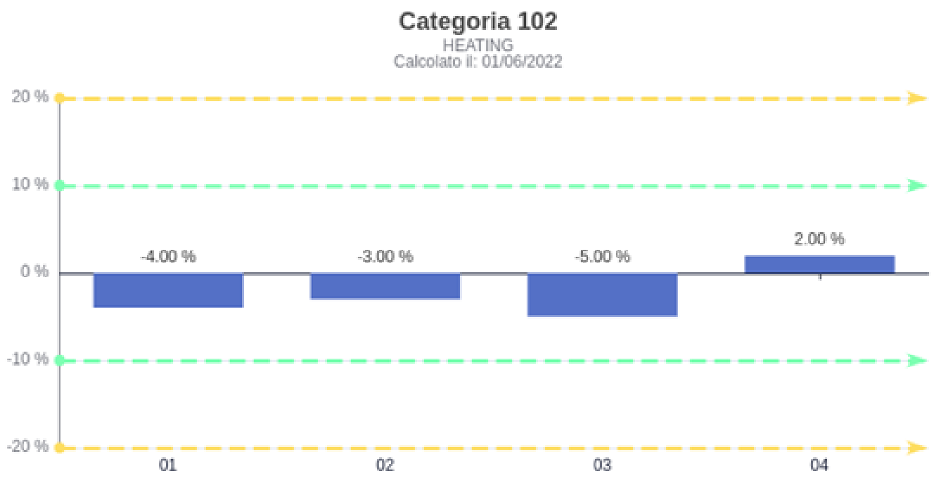


Figure 26: LLUC-3B_PI-01 KPI-02 Energy Consumption Gap of a building with itself during the time

The KPI 03 is based on the benchmark of a building compared to itself on a very short period. The result of the KPI must be investigated. In fact, there could be events (internal or external) in the building that could have changed the consumption request. Also, the availability of data could impact on the effectiveness of results. So, we are analyzing a possible solution to enforce the analysis.

Regarding the KPI 04, even if, in the example the actual value matches with the target value (this value is not technical but a business one), analyzing the graphics also on more buildings it can be noted that the current method of calculating the KPI doesn't take into account the size of the buildings, which, even if they are part of the same cluster, can be very different and thus provide misleading information. In this case some changes will be made in order to refine the KPI.

Finally, the KPI 05 represents the CO2 emission reduction in a year for each building. Also, this is a business KPI so that the target value depends on the action taken by the energy manager to match with the company strategies and objectives. Hence, the validation of the KPI is not based on the achievement of the KPI's target value but on its effectiveness in providing information for the business. In the graphic below, taken a building as a sample, it shows that in the period from January 2022 to May 2022, there has been a reduction in C02 emissions between 6% (01/2022) to 41% (05/2022).

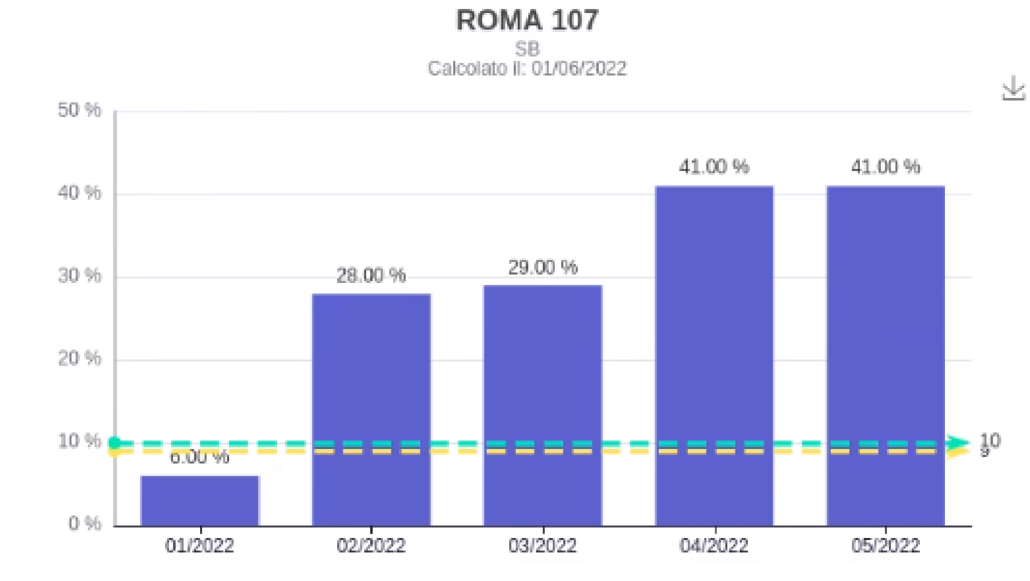


Figure 27: LLUC-3B-PI-01 KPI-05 CO2 Emission reduction monitoring

6.3 LLUC02-3B-PI-02 Anomaly detection of cooling & heating plants

The objective of this use case is to optimise maintenance efforts through monitoring techniques that can track equipment performance during normal operation and identify anomalies before they result in actual failures.

Based on information collected through meters and sensors installed in the buildings (such as systems energy consumption, internal temperature, number of sensors, …) the app detects possible anomalies in the sensor values, which might indicate a problem on the heating or cooling system.

6.3.1 Evaluation and Validation

Table 18: LLUC-3B-PI-02- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Recall – True positive anomalies identification	90%	n/a	The anomalies detected by the system must be compared to the actual number of anomalies (true positive) occurred. This number is not yet available.
2	Precision -	90%	n/a	The anomalies detected by the system must be compared to the actual number of anomalies (true positive) occurred. This number is not yet available.

3	F1-Score	90%	n/a	This KPI cannot be calculated at the moment because it depends on the above-mentioned KPIs
4	Performances Analysis	5%	n/a	KPI not yet calculated

During the period of the project the app detected multiple violations of the specified thresholds especially on the temperature sensor measurement on multiple buildings. The tables below show an example of detected event. However, to calculate the KPIs it is necessary to have false positive or false negative values. These last values are not yet collected. To calculate the violations, it is possible to select three different methodologies: rule-based detection, Spikes-based detection and Trend-based detection. A sample of anomaly result based on Rule-based detection is in the following tables were the tool reports all the events occurred when the temperature is lower (<17°C) or higher (>25°C) than the defined threshold.

Table 19: LLUC-3B-PI-02-Validation results- Roma Corviale (RML61900) for the KET-THL-200 D2 sensor

#	Date	Event	Measurement
1	04/06/2022 13:08	Upper Threshold Violation	27.1
2	06/06/2022 07:08	Upper Threshold Violation	27.1
3	06/06/2022 11:23	Upper Threshold Violation	27.1
4	06/06/2022 11:38	Upper Threshold Violation	27.3
5	06/06/2022 11:53	Upper Threshold Violation	27.2
6	06/06/2022 12:08	Upper Threshold Violation	27.6
7	06/06/2022 12:23	Upper Threshold Violation	27.6
8	06/06/2022 14:53	Upper Threshold Violation	27.2
9	07/06/2022 07:08	Upper Threshold Violation	27.4

Table 20: LLUC-3B-PI-02-Validation results- Roma Corviale (RML61900) for the KET-THL-200 D1 sensor

#	Date	Event	Measurement
1	10/01/2022 08:15	Lower Threshold Violation	16.9
2	10/01/2022 08:45	Lower Threshold Violation	16.8
3	10/01/2022 09:00	Lower Threshold Violation	16.5
4	10/01/2022 09:15	Lower Threshold Violation	16.6
5	10/01/2022 09:15	Lower Threshold Violation	16.9
6	06/06/2022 07:08	Upper Threshold Violation	27.5
7	06/06/2022 07:23	Upper Threshold Violation	27.1
8	07/06/2022 07:08	Upper Threshold Violation	28.1
9	13/06/2022 07:08	Upper Threshold Violation	27.2

Table 21: LLUC-3B-PI-02-Validation results- RML83630 for the KET-THL-200 D2 sensor

#	Date	Event	Measurement
1	17/01/2022 07:30	Lower Threshold Violation	16.7
2	28/05/2022 16:10	Upper Threshold Violation	27.1
8	02/06/2022 18:55	Upper Threshold Violation	29.9
17	02/06/2022 21:10	Upper Threshold Violation	27.9

18	02/06/2022 21:25	Upper Threshold Violation	27.8
58	07/06/2022 08:10	Upper Threshold Violation	27.1

Table 22: LLUC-3B-PI-02-Validation results- RML83630 for the KET-THL-200 D1 sensor

#	Date	Event	Measurement
1	02/06/2022 18:55	Upper Threshold Violation	29.6
2	02/06/2022 19:10	Upper Threshold Violation	29.4
3	02/06/2022 19:25	Upper Threshold Violation	29.2
4	02/06/2022 19:40	Upper Threshold Violation	29
5	02/06/2022 19:55	Upper Threshold Violation	28.9
6	02/06/2022 20:10	Upper Threshold Violation	28.8
7	02/06/2022 20:25	Upper Threshold Violation	28.8
8	02/06/2022 20:40	Upper Threshold Violation	28.7

Figure 28 shows the anomalies represented through bullets. However, there are still some inaccuracies that need to be solved.

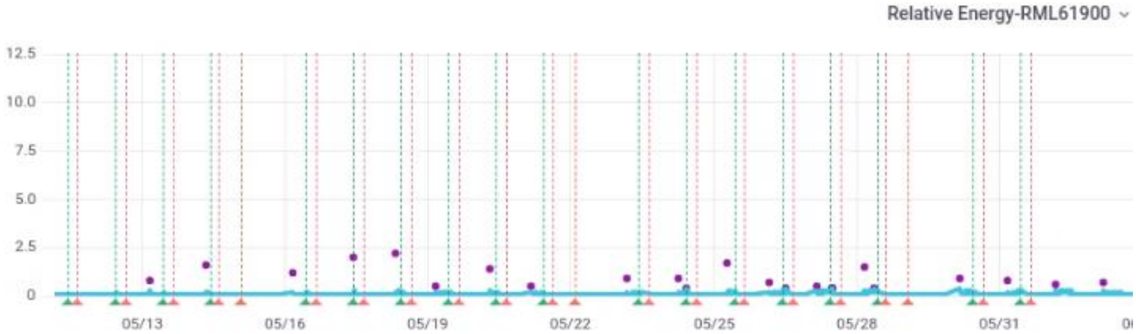


Figure 28: LLUC-3B_PI-03 Violation Detected in RML61900 Building

At a broader view of the tool, energy consumption analysis and testing found that in one specific office, energy consumption data had abnormal and inconsistent consumption peaks compared with expectations.

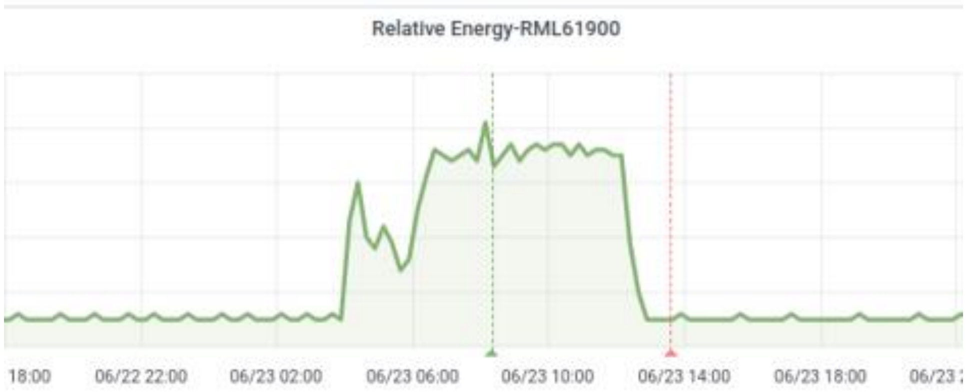


Figure 29: LLUC-3B-ROM-02- Anomalous energy consumption peaks

The detection of this peaks led to the discovery of an error in the installation of power lines (line inversion) not intercepted by other systems currently in operation. In fact, as the following figure shows the lighting and air conditioning consumption (blue and orange lines) are greater than the value of the total building analysed (green line).

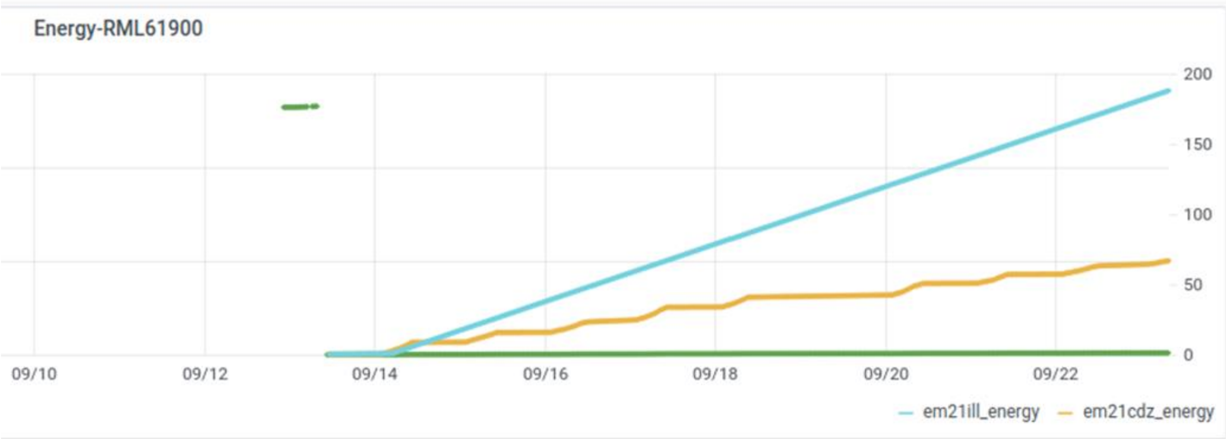


Figure 30: LLUC-3B-ROM-02- Anomalous energy consumption in a building

6.4 LLUC03-3B-PI-03 Lighting Consumption Estimation & Benchmarking

The objective of this use case is to estimate the specific building lighting consumption, in order to benchmark, detect anomalies and plan optimization actions to reduce lighting consumption of the building and the corresponding Green House Gases (GHG) emissions.

6.4.1 Evaluation and Validation

Table 23: LLUC-3B-PI-03- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Lighting Estimation	+/-5%	+7%	The KPI is calculated on the last week and for each building. The value reported (as sample) is related to the ROMA 107 building in the last week.

The information on building lighting consumption is almost never available. The knowledge of the total energy consumption of a building and that of some systems in it is not sufficient to have an estimate of lighting consumption. The app calculates this value starting from information on consumption but analyzing also other information and produces the value required. To test the validity of the algorithm, an analysis on buildings with the lighting consumption information (classified as Smart Buildings) was performed and instructed the algorithm to make an estimate. The estimate is compared with the actual data. As shown in Figure 31 the KPI has a value of 7% with a deviation of +2% from the target value.

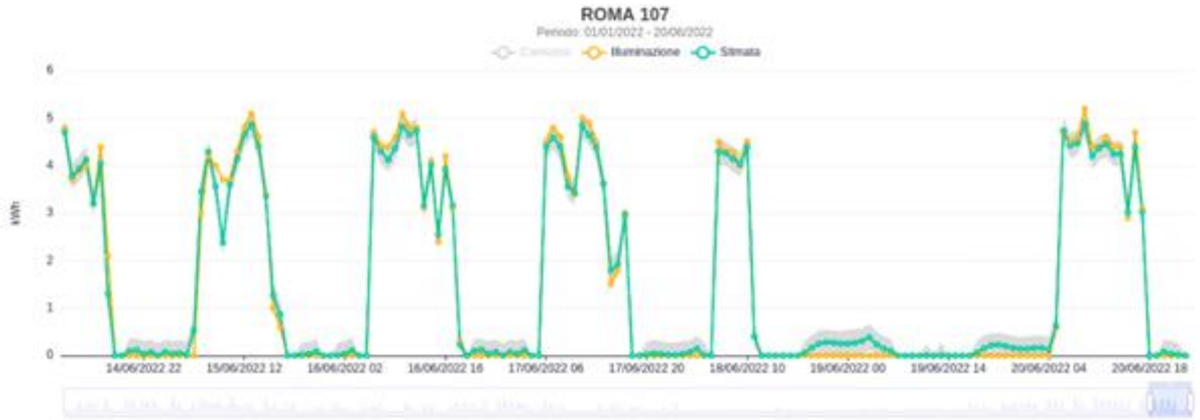


Figure 31: LLUC-3B-ROM-03- Lighting Estimation in a sample building

During the monitoring of lighting consumption, it was found that since 8t March 2022 the consumption has significantly reduced and stabilized at a new threshold as shown in Figure 32. This was due to an intervention to replace lighting technologies on that day.

Figure 32: LLUC03-3B-PI-03- Histogram of calculated lighting consumption during a specific period

The benchmarking service provided near real-time information to the energy manager giving immediate and objective feedback on the effectiveness of the chosen solution that led to a significant reduction in consumption.

6.5 Conclusion

Compared to the use cases identified for the 3B-PI pilot, the tools developed, although simple, are proving to be a valid support for end users in the energy domain (building and energy managers) to monitor, in some cases almost in real time, consumption trends and the behavior of HVAC and lighting systems as external or internal conditions change, even unexpectedly.

The analysis method developed for the prediction of energy consumption (LLUC01) and lighting consumption (LLUC03) when applied and measured showed that in the presence of scarce historical data, the margin of error can become significantly high. This needs to be investigated in order to

improve the performance of the algorithm. The remaining KPIs calculated for the LLUC01 use case measure business performance and thus provide useful elements for performance monitoring and decision support. A margin for improvement can be considered for the calculation of KPI 04 by introducing new elements of comparison into the analysis.
Since we are still working on acquiring significant information for the calculation of KPIs (e.g. reporting of actual malfunctions) and on the integration of the tool into the Digital Enabler platform, KPIs for the LLUC02 use case have not yet been produced, although interesting aggregations and comparisons of data can be obtained through the available dashboards, which can direct analyses on system behavior and identify relationships and interdependencies between the various factors analysed.
In this scenario, the final validation of the KPI result will be realized for the next validation report.

7. Pilot 3B-ROM Evaluation & Validation Report

7.1 Introduction

Pilot 3B-ROM is formed of more than 2000 building owned by the municipality of Rome and focuses on a single low-level use case: LLUC-3B-ROM-04 - Monitor and analysis system of Data coming from energy meters of ROME Municipality buildings asset.

7.2 LLUC-3B-ROM Monitor and analysis system of Data coming from energy meters of ROME Municipality buildings asset

This use case focuses on building an integrated monitoring and analytical system for data coming from the meters of different buildings of the Rome Municipality that can increase the awareness on the energy consumption profiles, anomalies, forecasting, PV plants potentialities on roofs and more in general on the efficiency measures potentialities. It can also increase the capacity of the Energy Management office to produce more frequent and accurate Energy Audits. This use case is formed of 4 services: 1) Spatial reporting, 2)Benchmarking, 3)Forecasting and 4) RES potentiality (PV plants).

7.2.1 Evaluation and Validation

The evaluation of the pilot can be run at three different levels: 1)energy planning and policy level; 2)information and data quality level and 3) energy efficiency technical level. Most of the KPI presented are focused on the energy efficiency technical level although some of them also concern the energy management level.

It is also important to notice that the dashboards corresponding to the 4 services are used by officers with specific competences and tasks to achieve, that will be logged and will also find in the Notification Area of the toolbox the functionalities to report and comment their work sessions, to propose the efficiency or maintenance measures to be implemented and to receive automatic notifications based on pre-set criteria. In order to analyse and assess these notifications, consisting in an adequate volume of items, it will be necessary to wait not less than 6 months after the final user test phase.

Table 24: LLUC-3B-ROM- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
01	Total Energy Savings TES (kWh / y) [% : kWh-saved / kWh-Yc] [Yc current year = past 12 months]	1 % =relevant; 2 % =good; 3 % =very good; Over 3% =excellent	the Total Energy Savings [%] calculated and limited to interventions resulting from the toolbox use (test phase) is between 1% and 2% (Good)	The analysis of the meters data (historical and current) produces a series of measures and interventions that should reduce the yearly total energy consumptions, such as dismission of un-useful meters, maintenance and intervention plans on buildings following consumptions anomalies detection. A derived KPI is the Saved Energy Cost (€/y) that depends on energy tariffs but could be also impacted by contractual redefinition resulting from the data analytics toolbox developed in the project. A list of the EVENTS (actions/interventions) impacting on TES, will be provided after the final user test phase.
02a	Saving Costs Personnel costs (Euro/y)	Up to 10k€ = relevant Up to 30 k€= good Up to 60 k€ = very good Over 100 k€ = excellent [Personnel Hourly cost X Total hours of work avoided] for activities	To be calculated This KPI calculation needs for the supply of specific reports by the SIMU offices, but in the first period KPI could be estimated through user interviews	The use of the toolbox and the automatization of some functionalities offered by the 4 services will decrease the amount of worked hours dedicated to the same tasks, freeing up time for other activites. The installation of a nRT monitoring systems (WP7) is going to further reduce the costs for the personnel.
02b	Saving Costs Energy Related Costs other than 2a (Euro/y)	Could be included in KPI02a targets; Other Cost Savings resulting from the use the 4 services offered by the toolbox	To be calculated (see 2a)	This component of cost saving refers to costs other than personnel. i.e. fixed fees paid for meters that have to be dismissed as a result of the toolbox services application. Note: this is NOT the Cost for Energy Saving that is derived from KPI_01
03	Nb of Meters with Energy Savings Results (Nb of Meters)	To be defined and then to be calculated	To be calculated on the basis of KPI-01 analysis	This indicator counts the number of energy meters for which PLATOON data analytics tools produce some action resulting in energy saving during the year. Derived KPI: KPI01/KPI03 represents the average energy saved per meters involved, and measures the average

				intensity of the single EE interventions
04	Nb of Anomalies detected (Nb of Recorded Notifications Anomalies)	10 =relevant; 20 =good; 30 =very good; Over 30 =excellent	More than 20 ; Good (1/2 year of observation) A list of detected anomalies identified through the toolbox will be provided	Not all alerts sent by Platoon tools produce Energy Savings therefore it is interesting to track separately the number of anomalies occurred during a period of observation. The definition of anomaly for a specific energy meter is based on the occurrence of the consumption divergence from the expected value (see benchmark analysis), in the same period. Typically, when the building itself or is usage is highly inefficient Platoon will send a series of alerts. This must be considered a good result of the project even if the beneficiary is unable to intervene producing energy savings.
05	% of CO2 emission reduction	To be defined on the basis of KPI-01 and then to be calculated	To be calculated	See KPI n.01 comments
06	RES suggested self-consumptions (kWh/years)	Over 130.000 =relevant; Up to 400.000 =good; Up to 800.000 =very good; over 1.200.000 =excellent Rom_04_Kpi_R06 is an additional component to ROM_04_Kpi_01 as it represents the potential further Energy saving (self-consumptions) and new local RES production	A list of new potential pV plants that can be installed on municipal roofs have been identified through the toolbox. The related calculated KPI will be provided in the next month	The calculation of the RES potentiality or more precisely of the energy from new PV plants that can be installed on municipal roofs, is based on the load curves, on the availability of irradiated surfaces to install RES plants, their tilt/orientation, etc. It includes new self-consumption energy quote that depends also on RES/Storage solutions that can be foreseen. Platoon output in terms of Total Potential RES kWh/Y calculation represents directly a positive impact to be measured.
07	Nb of Tools Outputs (Number of occurrences from Toolbox log)	Over 100 =relevant; Up to 200 =good; Up to 400 =very good; over 600 =excellent	A list of outputs coming from the Platton Toolbox for Pilot-3b-ROM have been identified through the toolbox log. **The related calculated KPI will**	Platoon results in terms of Queries with Output processed by the offered (tested) tools represents a positive impact to be measured. Measuring the usage, this KPI is referring about the effective engagement of the ROM personnel. Counting outputs for each distinct services and tools will help to address further development and exploitation strategies.

			be provided in the next month	
08	Vote assigned by the Test Users	Target: >3.5 Range [0 – 5]	**The related calculated KPI will be provided in M32**	To be calculated at the end of the final User Test phase (M32). Rating will be assigned for each services.

Regarding KPI_01 and KPI_02, they will be calculated after 6-12 months of full use of the toolbox, when each concrete action contributing to the Energy Savings will be recorded by the officers in their Notifications Area of the dashboard. For example, dismissing n.15 power meters on the basis of the toolbox analysis results on Energy Saving (KPI_01, KWh/y), on reduced Personnel cost (KPI_02, €/y) and on Money Saving (not only for reduced energy cost but also for other management costs and cut fees).

A list of actions and planned interventions, resulting from the pilot Toolbox use, will be delivered at the end of the final user test phase foreseen at M32. This list will allow to calculate KPI_01 (and some derived KPIs), KPI_03, KPI_05, KPI_07.

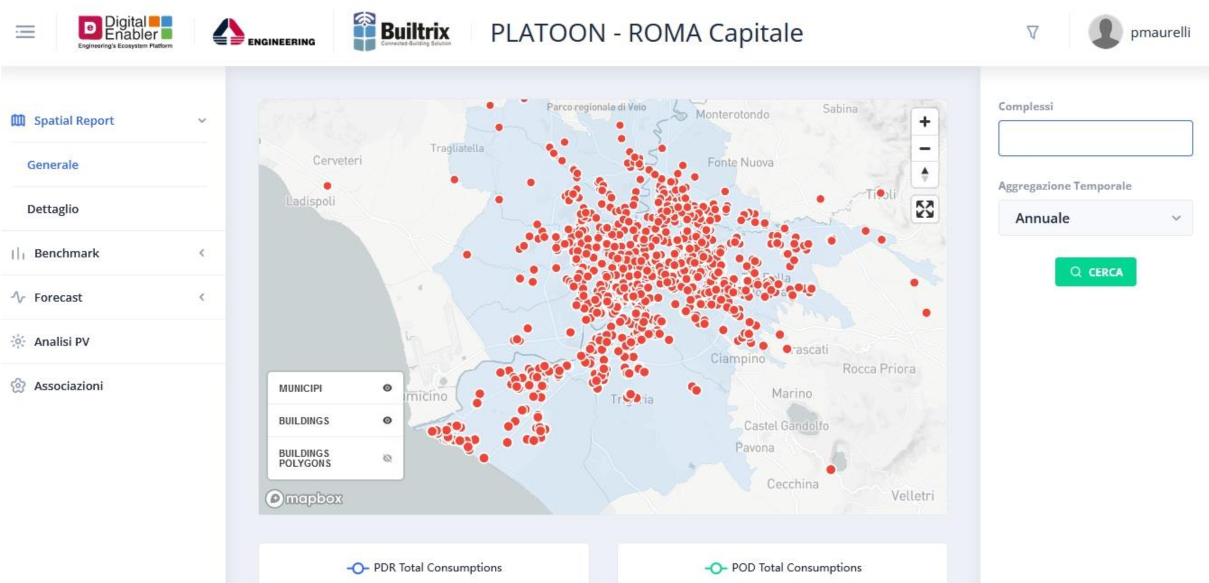


Figure 33: Pilot 3b-ROM-01 – Spatial Reporting dashboard: 1200 buildings with power and/or gas meters supplying data to the toolbox. The queries and spatial selections offered consent to obtain partial aggregated reports per Districts or per typologies of building

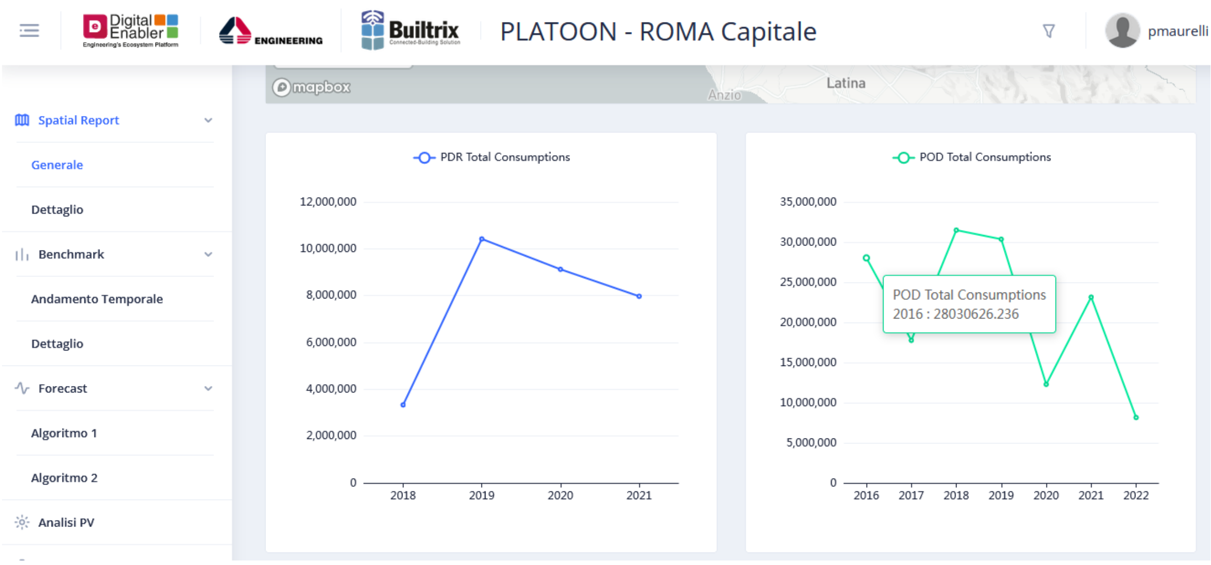


Figure 34: Pilot 3b-ROM-01 – Spatial Reporting dashboard: overall energy consumptions from Gas Meters annual data and Power Meters annual data, for the whole analyzed asset. The same can be done for clusters of buildings based on several selection criteria.

KPI_08 (Votes by test users) will be calculated in M32 specifically for this service. The automated reporting functionalities seems to be at present among the most appreciated outputs reducing significantly the time users have to dedicate to this task.

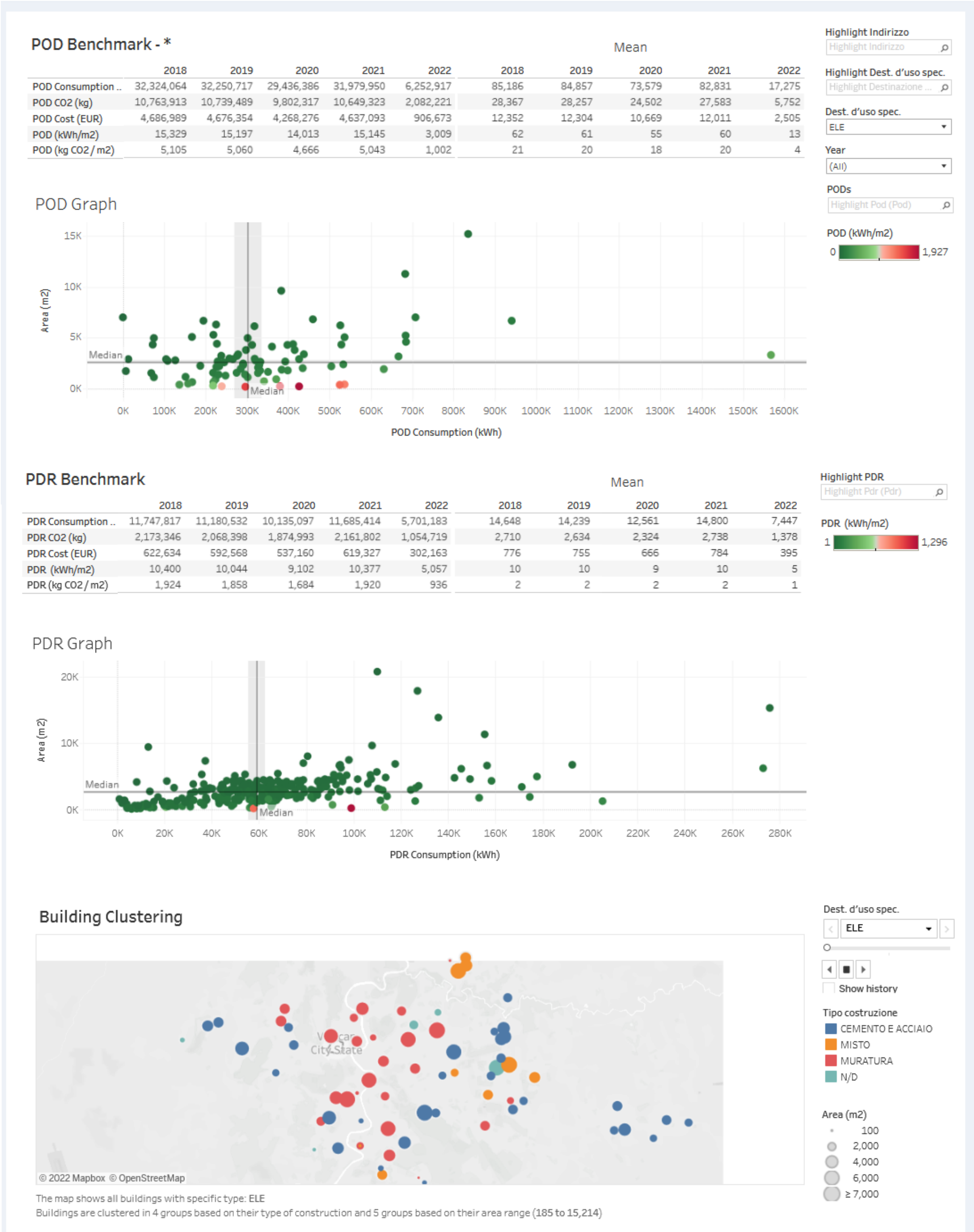


Figure 35: Pilot 3b-ROM-02 – Benchmarking dashboard: overall energy consumptions, costs, CO2 for both Gas and Power meters and clustering of buildings by type of construction

Besides, within the benchmarking service is included the High Level Anomaly Detection functionalities but these need to be further tested and used by the ROM officers in order to implement the specific thresholds that will define the anomalies conditions. Basically the automatic benchmarking analysis presents buildings that exceed their reference cluster parameters (Building Clustering in fig.). The main assessment can be done in terms of performance (KWh/m2) through a graphic interface that highlights buildings that exceed the average value, orienting and supporting the user in the search for technical causes and in the definition of response measures.POD (electricity) and PDR (gas) benchmarking can be aggregated for each building or complex of building.In this early stage of the pilot the KPI_04 and the KPI_07 will result in a picture of the usability and effectiveness of this service. A margin for improvement can be considered for the calculation of these KPIs by introducing new elements of comparison into the analysis, with the scope to automatically detect other kind and conditions of consumptions anomalies.

Figure 36: Pilot 3b-ROM-02 – Benchmarking dashboard: building energy consumptions compared year by year; note the anomaly Gas metering for building n.1803 on 2021

The Forecasting functionalities are periodically used by the officers of SIMU Department engaged in administrative tasks, including the forecast reports on expenditure. The Covid restrictions impacted on the energy Consumptions depending on relevant reduction of most of the municipal buildings, so different algorithms were developed, tested and then implemented in order to take into account the anomalies in the time-series induced by Covid emergency. The aggregation for districts or for buildings typologies is one of the main task the users are conducting in order to produce periodic reports. KPI_07 is also a measure of the benefit and frequency of use of functions within this service.

Figure 37: Pilot 3b-ROM-03 – Forecasting dashboard: whole asset energy consumptions for Gas (PDR) and for Power (POD) meters. First algorithm.

Besides, the KPI_3b_ROM_09 presented in the appendix and used also by the 3B_PI pilot, can calculate the % of deviation between the energy consumption forecast and the actual consumption in the building. This KPI checks how closely the predictive model adheres to reality, measuring the Effectiveness of the forecasting functions of the pilot toolbox.

In order to evaluate results for this KPI_09 it is necessary to wait some more months acquiring new data on energy consumptions and comparing with predictions calculated at beginning of 2021. The result for this procedure will be presented between M33 and M35.

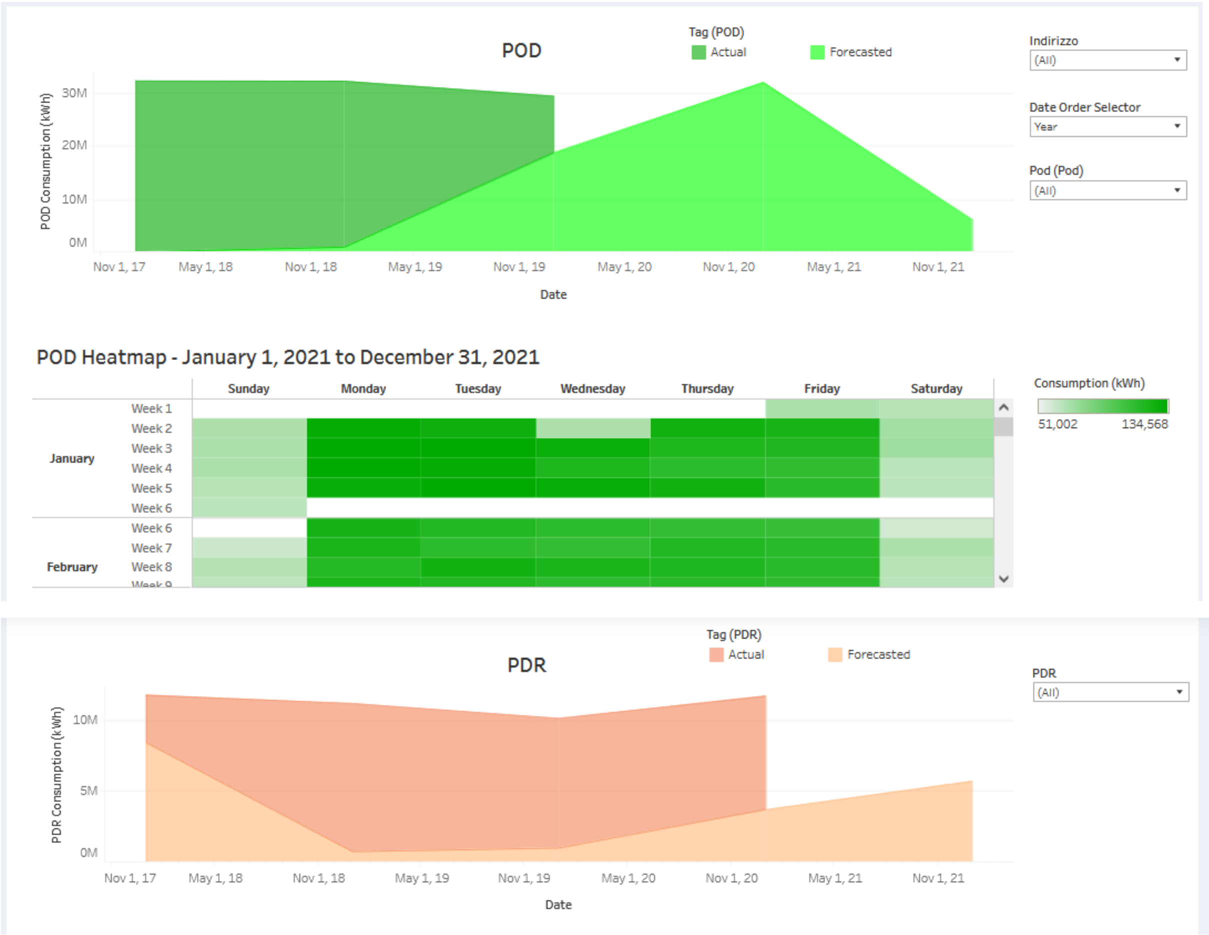


Figure 38: Pilot 3b-ROM-03 – Forecasting dashboard: whole asset energy consumptions for Gas and for Power meters. Second algorithm. A POD Heat map is also presented.

Finally, the RES Potentialities (PV plants on roofs) is highly appreciated in the SIMU Department as it supports directly the planning process for PV plants asset extension on the owned buildings roofs. During the project the introduction at national level of new public incentives and connection schemes related to Renewable Energy Communities (REC scheme, Sharing PV energy surplus with other proximity users) prompted the pilot project team to redefine the scope and implementation of this service 3b-ROM_04 in order to obtain the estimation for each roof of the maximum peak power and the maximum PV production surplus (over the self-consumption quote of the building).

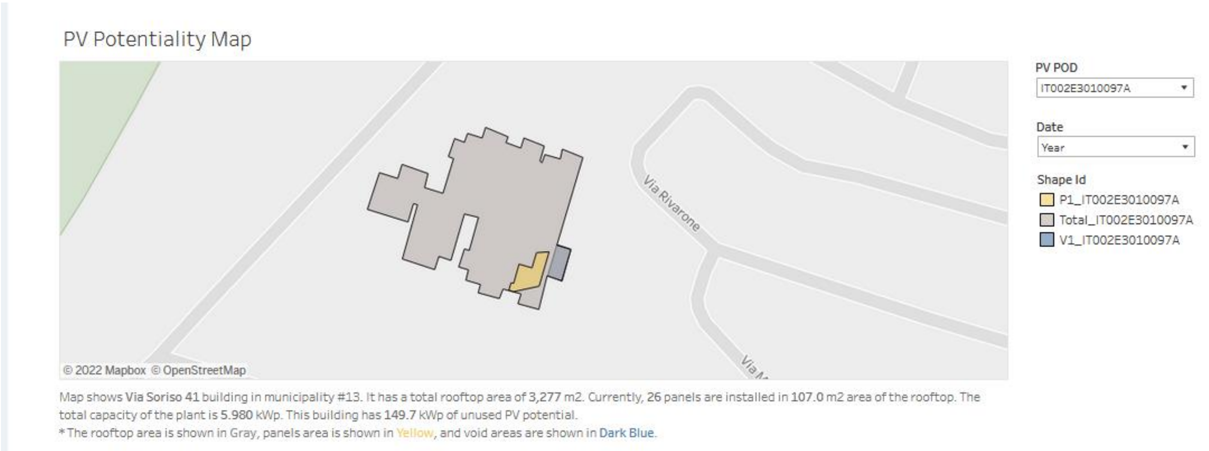


Figure 39: Pilot 3b-ROM-04 – RES Potentialites dashboard: for each building hosting a PV plant the map shows and calculates the free surface useful to expand the PV plant.

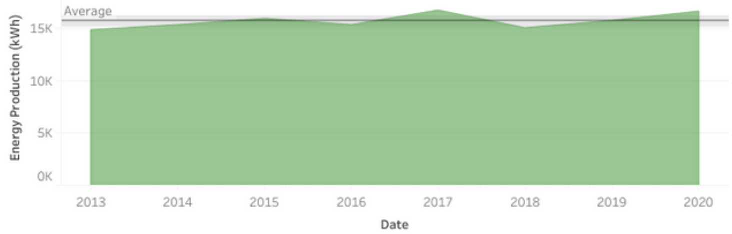
The KPI_06 (kWh/y of RES production that can be installed on the roofs) is now limited to the extension of 160 roofs already hosting existing PV plants, where the algorithm calculates the free surface available, applying custom parameters for PV technology to simulate for the new PV plant, and gives as outputs the Total RES Production (kWh/y) that can be realized, the investment and the ROI.

The result is excellent (more than 1.200.000 kWh/y estimated from new PV plants) and can directly influence the planning strategy of the Municipality accelerating the design and the realization of many PV plants within the REC scheme.
The RES stakeholders, civil society organizations and the municipalities are going to meet and discuss in the next period to define the business model and the operative strategy for RECs on public roof, probably giving priority to school roofs, creating synergies also with another EU funded H2020 project (SUN4ALL) involving the municipality.

The services could be extended in the next future also to the roofs not already hosting any PV plant to calculate the new plant installation potentialities and also to estimate the costs and the CO2 impacts for these eventual future investments.
Automatic calculations can be improved introducing more accurate data on free surfaces and ideal orientation/tilt on the roofs.

Map shows **Via Giuseppe Scalarini 27** building in municipality #4. It has a total rooftop area of **2,256** m2. Currently, **48** panels are installed in **134.0** m2 area of the rooftop. The total capacity of the plant is **11.47** kWp. This building has **152.7** kWp of unused PV potential.
* The rooftop area is shown in Gray, panels area is shown in Yellow, and void areas are shown in Dark Blue.

Estimated PV Production

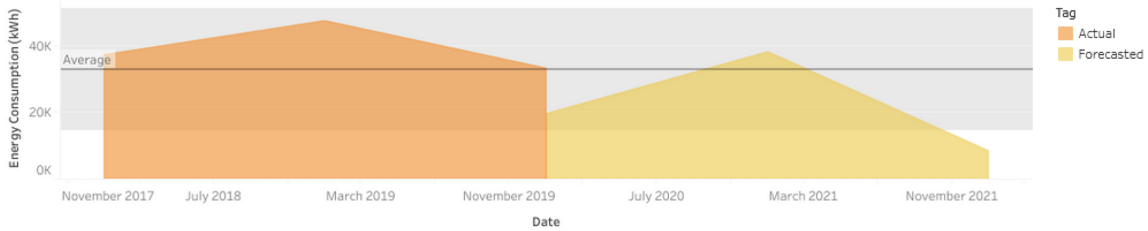


This graph shows the PV Generation of existing plant with **POD: IT002E3054878A** in **Via Giuseppe Scalarini 27** which is estimated using the Solar Satelite Data.The generation range is **14,842 to 16,668** (kWh) per **Year**.

Measured (Lovato), Efficiency: None %

For those panels that have data in Lovato Platform. The efficiency is estimated based on the hourly generation data.
* Efficiencies higher than 100% may be seen due to incorrect meta-data such as number of panels or their peak power capacity.
* Efficiencies lower than 50% may be seen becasue of huge data losses in the lovato system.

Consumption



The Energy Consumption (kWh) for **IT002E3054878A** in **Via Giuseppe Scalarini 27**. The consumption range is **8,332 to 47,323** (kWh) per **Year**

Monthly Analysis / Annual Analysis

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	2020
Total Consumption (kWh)	4,897	4,089	1,598	850	821	1,132	1,777	1,296	3,429	4,238	4,460	4,434	33,022
Total PV Generation (kWh)	2,704	3,019	3,347	3,865	4,034	4,090	4,344	3,982	3,115	2,687	2,345	1,717	39,247
Utility Consumption (kWh)	2,986	2,061	942	467	421	485	819	687	1,926	2,452	2,807	3,275	19,327
PV for Trade (kWh)	792	991	2,691	3,481	3,634	3,442	3,385	3,373	1,612	900	692	558	25,552
Utility Cost (EUR)	448	309	141	70	63	73	123	103	289	368	421	491	2,899
Trade Income (EUR)	95	119	323	418	436	413	406	405	193	108	83	67	3,066
Total Saving (EUR)	382	423	421	475	496	510	550	496	419	376	331	241	5,120

This table presents the **Total Consumption** and total **PV Generation** in **IT002E3054878A** in each month for the selected year in the filter. The **Utility Consumption** is the energy purchased from the grid and **PV for Trade** is the excess of the **PV Energy** available for the Trade. The utility cost is calculated based on the **Average Utility Tarif (0.15 EUR)**. The Trade income is calculated based on the **Trade Tarif** (0.12 EUR). The total Saving is estimated based on the self-consumption savings and trade incomes.
* The annual aggregations are presented in the last column.

Year: 2020
POD: IT002E3054878A
Panel Capacity (Wp/m2): 142
New Plant Area (m2): 111
Trade Tariff (EUR): 0.12
Average Utility Tariff (EUR): 0.15
Plant Cost (EUR/Wp): 1.7

Self-consumption Analysis

Is Possible: Yes
Required Plant Area (m2): 111.0

This analysis is based on the available rooftop area (**1,804** m2) and estimated Rquired Plant Area (**111.0** m2) to acheive the self-consumption. Self-consumption is not possible if the required area is bigger than available area on the rooftop.
User may put the **111.0** m2 in the **New Plant Area** (m2) field to see the changes in the Monthly Analysis table and Return on Investment Analysis widget.

Return On Investment Analysis

Estimated Capital Cost (EUR): 26,795
Period (Year): 8.1

The Estimated Capital cost is based on Plant Cost (**1.7 EUR/Wp**) and the New Plant Area (**111** m2). The ROI Period (**8.1** Years) is calculated based on the forecasted generation timeseries, their comparison with forecasted consumptions, and disaggregation of the self-consumption and trade scenarios.

Figure 40: Pilot 3b-ROM-04 – RES Potentialites dashboard: for each building roof where is possible to expand the PV plant, PV yearly production is estimated then the investment cost and the Pay-Back Time, the self-consumption is calculated and also some standard

7.3 Conclusion

An initial evaluation of the pilot has been be conducted at different levels. At present the impact that is verifiable and measurable is at the Energy Management Office level that benefit in terms of responsiveness, of completeness and depth of the cognitive picture, of full integration of the dashboard and the datasets on energy consumption and production (PV). The KPI_01 (KWh/y saved) and the correlated KPI_05 (reduced CO2) together with KPI_02 (Personnel cost reduction) are the main indicators of the impact of the Pilot-3b-ROM toolbox in terms of energy transition and sustainability and more specifically in terms of improved behaviours of the personnel (SIMU Department – Plants Operative Unit) engaged in the energy management of the Rome Municipality asset. The general idea is that each session or use of the Toolbox (KPI_07) can produce knowledge, information and indications on how to improve the energy efficiency of this large asset of buildings. This awareness can result directly or indirectly into actions. Direct actions on buildings plants, meters or management can be recorded in the notification area of the toolbox marking the date of each specific intervention and consenting to later calculate the reduction of the EC for the correlated meters. This means that only few types of direct actions consent to use the toolbox to quickly calculate the KPI_01 (i.e. dismission of meters) while the majority of the enabled interventions need an observation period to calculate the resulting savings. The toolbox services for forecasting can help to estimate the expected saving after one or two months from the intervention on the basis of the data flow frequency coming from the meters.Indirect actions to improve energy efficiency consist in planned interventions or scheduled maintenance; in this case the KPI_01 validation will proceed with the recording of the scheduled intervention in the notification area and with the estimation of the future impact in terms of yearly saved energy.

On the other hand, the Information & data quality level evaluation is an ongoing process that focuses on the different data sources and on their evolution in time. The power and gas meters large asset analysed is evolving quickly thanks to the installation of new generation meters, where the quality and frequency and accessibility of data is improved. At the same time the vendors and energy service providers change and offer different data connectors or web services. Furthermore, within WP7 the pilot will be enriched with near Real Time data coming from sensors (2400 Contatermie for Heating and Test sensors for Electricity) that could significantly increase the Data quality.
Furthermore, the Energy Efficiency Technical Level evaluation can be effectively conducted through most of the KPIs presented. This evaluation aims to describe and report the impact in terms of Energy Efficiency that the Pilot is producing or can produce in the next future.

Some of the KPIs could not be properly calculated until the Energy management organization of the municipality will be restructured (New Directorate on Climate) and once the CDP (City Data Platform) will be connected with the Pilot Data. The complete results for all KPIs will be presented on M36, as during the last semester the test users will intensify their application.

8. Pilot 3C Evaluation & Validation Report

8.1 Introduction

Pilot 3c focuses CIC Nanogune building which is a public research center located in San Sebastian (SPAIN) managed by GIROA-VEOLIA. The building has 7319 m2 distributed over six floors and it contains offices, 15 ultra-sensitive laboratories and a cleanroom of nearly 300 m2 where the air purity is under strict supervision. The building has a BMS system and PV panels installed on the roof. Pilot 3c focuses on two main low-level use cases:

- LLUC-3C-01-Advanced EMS
- LLUC-3C-02-Predictive Maintenance

8.2 LLUC-3C-01-Advanced EMS

The objective of this use case is to match the demand prediction and RES generation prediction and to optimize the operation of building HVAC in order to achieve two objectives: (1) reduce the grid dependency and (2) reduce the energy bill.

8.2.1 Evaluation and Validation

Table 25: LLUC-3C-01- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Integration	1	-	This use case has not been validated yet as we have focused on use case 2. All this information will be included in V2 of this deliverable due in M36.
2	Energy Bill reduction	20%	-	
3	RES utilisation ratio	30% increase	-	

8.3 LLUC-3C-02-Predictive Maintenance

The main objective of this use case is to have a centralised control of the health status of different equipment of the building HVAC system based on the readings from multiple sensors for each machine. Amongst all the machines that form the building HVAC system, this use case focuses on two types of machines:

1. Hydraulic Pumps
2. Chillers

8.3.1 Evaluation and Validation

8.3.1.1 Hydraulic Pumps (TECN)

Table 26: LLUC-3C-02-Hydraulic Pumps-KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Health Monitoring	100%	100%	The developed algorithm is able to distinguish well the healthy and non-healthy operation using test data from an open-source dataset. Now we are validating with real data from GIR.

2	Failure Forecast	24 hours	~3 hours (170 mins)	The system is able to detect failure 3 hours in advance which is below the target value but should be enough time to be able to start the twin pump and avoid stopping the system.
3	Availability	N/A	N/A	This KPI cannot be applied in the case of pumps as there is no sufficient information to calculate it.
4	Mean Time Between Failures	N/A	N/A	This KPI cannot be applied in the case of pumps as there is no sufficient information to calculate it.
5	Maintenance Costs	N/A	N/A	This KPI cannot be applied in the case of pumps as there is no sufficient information to calculate it.
6	Integration	1	0.9	Implemented all the pipeline the using the Barbara OS except IDS part that not working due to issues with proxy and communications. Still pending integration with PLATOON edge-cloud framework.

In order to validate the data analytic tool for predictive maintenance of hydraulic pump different size of training datasets have been considered. For illustrative purposes the 100 vs 100 configuration has been represented. 100 vs 100 means that first 100 samples (files) have been used for training and last 100 samples (files) have been used to validate the outcomes. The expected result would release a failure scenario for no more that 10-15 final samples.

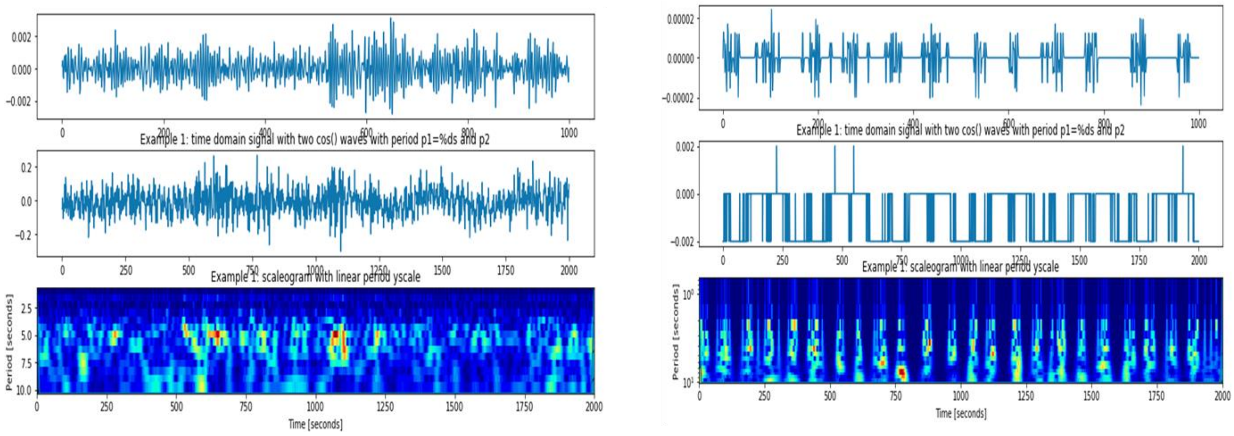


Figure 41: LLUC-3C-02-Hydraulic Pumps-Behavior for the first and final samples

Equally, different algorithms have been validated, namely a SVM OneClass Classifier, k-MEANS OneClass Classifier and DEEP AutoEncoder OneClass Classifier. Amongst all of them the one that produced the best results was the DEEP AutoEncoder OneClass Classifier. Equally, for each of the algorithms different hyperparameters have been attempted. The table below shows the validation results for different hyperparameter combination of the DEEP AutoEncoder OneClass Classifier. The dataset column represents the number of samples used for the validation and the train/test columns represent the train/test split. Finally, the result column represents the samples identified as failures by the algorithm.

Table 27: LLUC-3C-02-Hydraulic Pumps-Results for different hyperparameter combination of the DEEP AutoEncoder OneClass Classifier

Modelo	DataSet	Train	Test	Result
relu – mse	100	90	10	0, 0, 0, 10, 9, 8, 7, 5, 4, 3, 2
relu – mse	200	190	10	0, 0, 0, 10, 9, 8, 7, 5, 4, 3, 2
relu – mse	300	290	10	9, 0, 0, 0, 9, 8, 7, 5, 4, 3, 2
relu – mse	200	190	20	0, 0, 0, 9, 8, 5, 4, 3, 2
relu – mse	300	290	20	9, 0, 0, 0, 9, 8, 7, 5, 4, 3, 2
relu – mse	100	80	20	0,0,0,20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2
relu – mse	200	180	20	9, 0, 0, 0, 20, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2
relu – mse	300	280	20	9, 18, 0, 0, 0, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2
relu – mse	300	290	20	0, 0, 0, 13, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2
relu – msle	600	530	50	0, 0, 0, 42, 41, 40, 39, 38 ….
relu – msle	600	600	50	0, 0, 0, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3
relu – msle	600	600	100	0, 0, 0, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3

Analysing the results, it can be noted that the Autoencoder with “ReLU” activation function and “MSLE” reconstruction metric provides the best results. As it can be seen this algorithm identifies as failure all the samples since 17 sample before failure. The origin of the samples (0) represents the failure time and each of the samples is separated by 10 mins. Therefore, the algorithm is able to diagnose with almost 3 hours (170 mins) before the failure occurs. This value is below the threshold value of 24 hours but above the target value of 2 hours. However, it should be enough time to start the twin pump with enough time to avoid HVAC system to stop.

As a conclusion it can be noted that the current algorithm is looking into the symptom (vibration) rather than the cause (bearing crack due to fatigue). The bearing crack is a sudden phenomena, so, it is difficult to predict much in advance just looking into vibrations. Thus, in order to be able to predict failure in advance, we should look into the cause by using some type damage accumulation formula (e.g. Palmgren-Miner). However, this is a totally new approach that is out of the scope of the project but could be explored in a new project.

Regarding the integration KPI, all the pipeline has been validated using the Barbara OS and everything is working except the IDS part due to issues with proxy and communications. The corresponding evidence is included in the test report as part of Open Call deliverable part of WP7. For the end of the project, we are still pending integration with PLATOON edge-cloud framework as an alternative to the Barbara OS system.

Finally, regarding the pending aspects towards the end of the project, the first priority is to validate with real data from GIROA. However, this has 2 main limitations: on the one hand, we will have limited data as the sensors were recently installed. On the other hand, we won´t have failure data and we will just be able to test that the algorithm predicts well normality.

8.3.1.2 Chillers

Table 28: LLUC-3C-02-Chillers-KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Health Monitoring	0 – 100%	0 – 100%	This is an aggregated Health Status view of the machine, based on the Health Status of the different elements of the machine. It is based on a weighted average formula
1.1	Energy Variator	R^2 >= 0,85	R^2 = 0,92	Digital twin models show high accuracy with the real data. The system can detect whether a fault has occurred.
1.2	Evaporator Outlet Temp	R^2 >= 0, 85	R^2 = 0,92	Digital twin models show high accuracy with the real data. The system can detect whether a fault has occurred.
1.3	Flow Meter	MAE <= 2	MAE = 1,53	The high variability of this scenario requires different Scorer for the validation. Using a MAE verification below 2% we can distinguish a Bias of 1,8% (112m^3h vs 2$m^{3h)}$
1.4	Power Consumption Increase	R^2 >= 0, 85	R^2 = 0,96	Digital twin models show high accuracy with the real data. The system can detect whether a fault has occurred.
1.5	Temp Increase	R^2 >= 0, 85	R^2 = 0,915	Digital twin models show high accuracy with the real data. The system can detect whether a fault has occurred.
1.6	Phase Imbalance	Imbalance %	Imbalance %	Rule based indicator detects health status problem if the imbalance of the voltage of phases is over 3%
1.7	Power Supply	R^2 >= 0, 85	R^2 = 0,915	Digital twin models show high accuracy with the real data. The system can detect whether a fault has occurred.
1.8	Starter	MAE <= 2	MAE = 1,42	The high variability of this scenario requires different Scorer for the validation. Using a MAE verification below 2% we can distinguish a Bias of 1,8% (112m^3h vs 2$m^{3h)}$
2	Availability	Calculated KPI	Calculated KPI	Thank to CMMS Integration, we are taking required information to generate and import availability KPI to the main dashboard
4	Mean Time Between Failures	Calculated KPI	Calculated KPI	Thank to CMMS Integration, we are taking required information to

				generate and import MTBF KPI to the main dashboard
5	Maintenance Costs	Calculated KPI	Calculated KPI	Thank to CMMS Integration, we are taking required information to generate and import MTBF KPI to the main dashboard
6	Integration	1	n/a	For the Pilot 3C, the integration with IDS has been developed for the low-level use case LLUC-3C-02-Hydraulic Pumps. The IDS approach does not make sense with Promind because it will always be running as on-premise architecture.

The main KPIs to be validated and the ones with greater interest for the Giroa business are the ones related whit the Health Status of the chiller.
Each failure modes have been thought as a process. Having in mind the failure mode of the machine, an output signal has been selected to represent the output of the process. The same principle has been applied for the input signals, so the ones which are representative of the failure mode process has been selected as input.

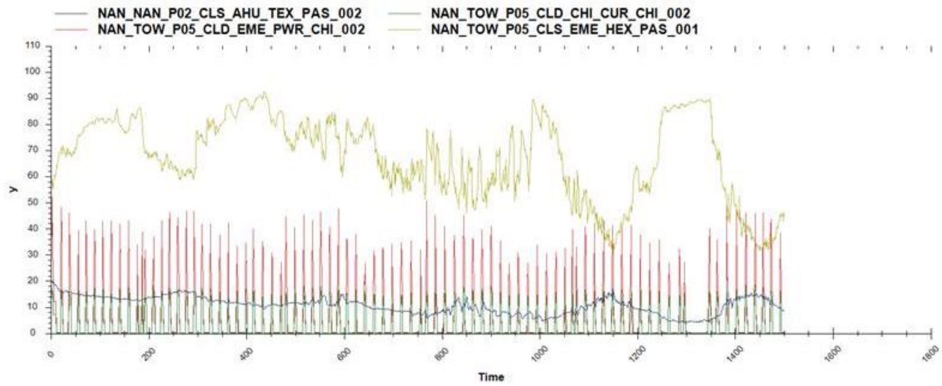


Figure 42: LLUC-3C-02-Chillers-input signals for Energy Variator model

Two ML approaches have been used to achieve best accuracy. Multilayer perceptron (MLP) and Random Forest models. After some benchmark testing the MLP model achieved the best accuracy. Hyperparametrization review has been performed with a result having the best performance with one hidden layer of 10 neurons, decay of 0,001 and learning rate of 0,01. The available real data has been split in 70% for training data and 30% for testing data.

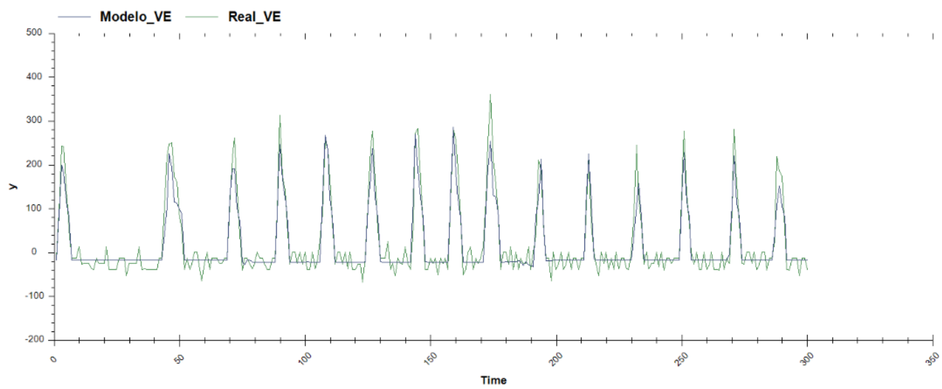


Figure 43: LLUC-3C-02-Chillers-Measured ThermicPower vs Predicted ThermicPower (Energy Variator Output)

All the KPIs are calculated and consolidated into a custom dashboard as shown in the figure below.

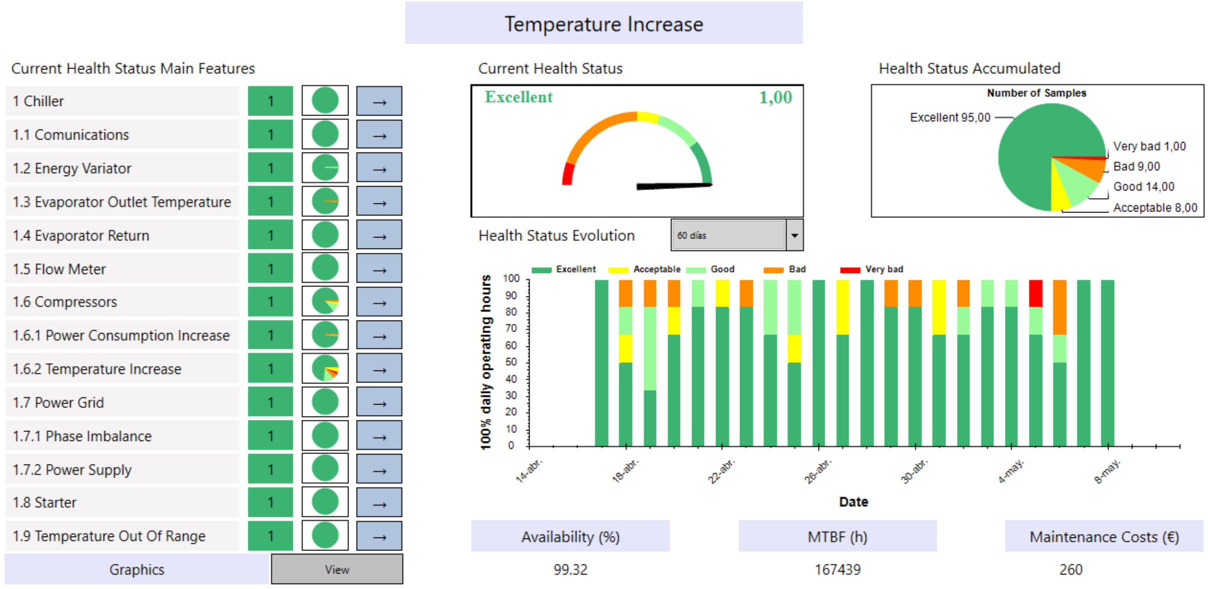


Figure 44: LLUC-3C-02-Chillers-Interactive dashboard for the hierarchical view of Health Status. Temperature Increase detail

8.4 Conclusion

As a result of the first validation it can be concluded that the validation results for the Predictive maintenance use case have been completed satisfactorily. On the one hand, the results for the Hydraulic Pumps predictive maintenance are acceptable in terms of health monitoring and failure detection. However, the results have been obtained with Open Source vibration data and need to be validated with real data from Giroa. In addition, some of the KPIs cannot be computed due to the lack of necessary data from Giroa. On the other hand, the results obtained from the Health Status analysis for the Chiller are of great interest. The system is able to determine whether the machine is working properly or if there is a malfunction problem not only at a machine level, but also identifying the machine-part or failure mode which is causing it.

Regarding the pending work, the validation of Advanced EMS use case tools are still pending and need to be completed for V2 of the deliverable due on M36. In addition, it is still pending integration with PLATOON edge-cloud framework.

9. Pilot 4A Evaluation & Validation Report

9.1 Introduction

This pilot takes place at the Multi-Good Microgrid Laboratory (MG2lab) in Politecnico di Milano, Italy. There is a single use case focused on Energy Management of Micro-grids (LLUC-4A-01) which aims to study data-driven energy management able to deal with increased complexity of the energy systems and to assess the advantages of innovative strategies: EMS with real-time processing and optimization for small-scale/renewable electricity generation, generation and load forecast, smart storage/generation.

9.2 LLUC-4A-01 Energy Management of Micro-grids

This use case focuses on data analytics tools aimed at the optimal exploitation of distributed renewable energy resources by means of an Energy Management Systems (EMS) with real-time processing and optimization for small-scale/renewable electricity generation, including specific implementation of day-ahead load consumption/generation forecast and nowcast capability. Indeed, the EMS is the algorithm that manages the forecasting modules for loads consumption and renewable energy production in view of the real-time management of all the energy assets in the micro-grid. The final aim is the optimization of unit commitment and scheduling of the energy resources on the base of these predicted profiles.

9.2.1 Evaluation and Validation

Table 29: : LLUC-4A-01- KPIs evaluation

KPI #	Description	Target Value	Actual Value	Comments
1	Energy availability	90%	88%	percentage of energy provided by renewable sources with respect to the measured consumption.
2	Cost	10%	12%	reduction of efforts and costs in terms of percentage of energy from the electrical grid with respect to total energy consumption
3	Forecast Accuracy (%error)	20%	17%	accuracy of forecasting in terms of percentage error with respect to the daily measured energy,
4	Realtime	80%	81%	ability of the system to monitor, forecast and optimize data in real time

In order to validate the data analytic tools for Energy Management of Microgrids developed in WP4, the different tools have been trained and tested with data from the MG2lab of Politecnico di Milano. The collection of the real time measurements of the MG2lab and the results of the implementation of this pilot specific tools is ongoing.
To ensure the performance of the microgrid energy management, optimization and control, and to measure its efficiency, suitable key performance indicators have been defined to assess the meeting of the requirements and the targets defined for pilot 4a. In particular, 4 KPIs have been specifically defined, which are currently under evaluation; these validation tests are also important to provide feedback for improvement of the optimization of the energy management system.

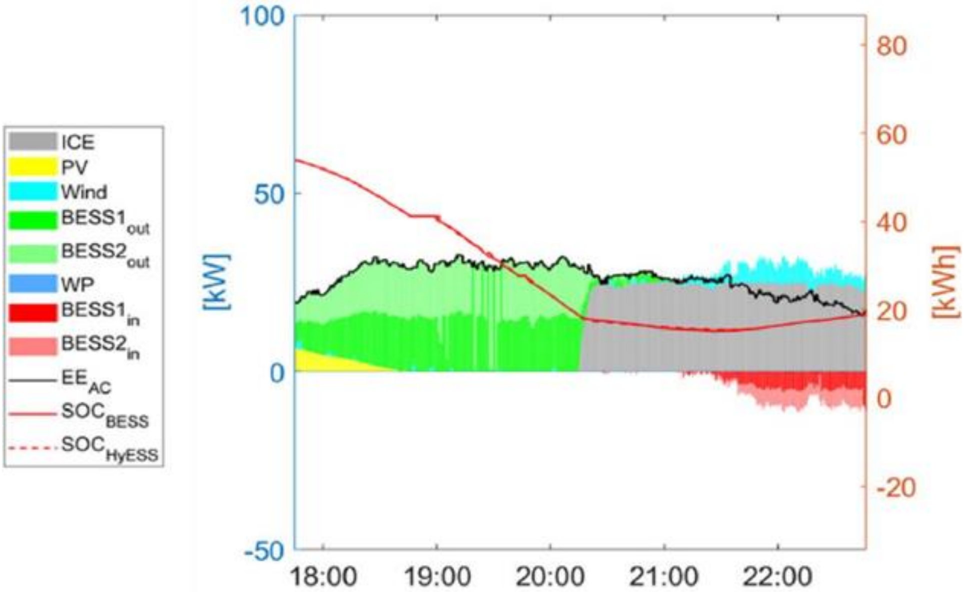


Figure 45: LLUC-4A-01-power production, storage and consumption of the microgrid.

Regarding the KPI related to energy availability, this indicator measures the percentage of energy provided by renewable sources with respect to the measured energy consumption, when the optimization for renewable electricity generation is performed considering smart storage and generation. In order to be able to evaluate this KPI, real power production ($P_{PV,t}$) and consumption ($P_{load,t}$) measurements are collected in real time at each considered time step from the micro-grid monitoring system, as described in the formula reported in the Annex I. The testing period spans across 24 hours, thus this KPI is computed with daily frequency by summing up the measured power values over the last 24 hours. The related results (in percentage) are reported into the consolidated dashboard. While higher percentage values correspond to a successful result in terms of energy availability, with an ideal target of 100% for this indicator, a threshold of 90% can be considered satisfactory.

Regarding the KPI related to costs, this indicator measures the reduction of maintenance effort and costs in terms of percentage of energy from the electrical grid with respect to the total energy consumption, when the optimization for renewable electricity generation is performed considering smart storage and generation.
In order to be able to evaluate this KPI, real power production ($P_{PV,t}$) and consumption ($P_{load,t}$) measurements are needed, as described in the formula reported in Annex I. These measurements are stored within PLATOON platform data storage, obtained in real time at each considered time step from the micro-grid monitoring system and the related results (in percentage) displayed into the consolidated dashboard. The minimum testing period is over 24 hours, but additional time horizons can be considered to provide an additional report on the performance of the system, thus this KPI is to be computed with daily frequency by summing up the measured power values listed above over the last 24 hours, but results will be also aggregated to longer time ranges with the increasing of collected data.
While lower percentage values correspond to a successful result, with an ideal target of 0% for this indicator, a threshold of 10% can be set as a satisfactory target in terms of energy cost.

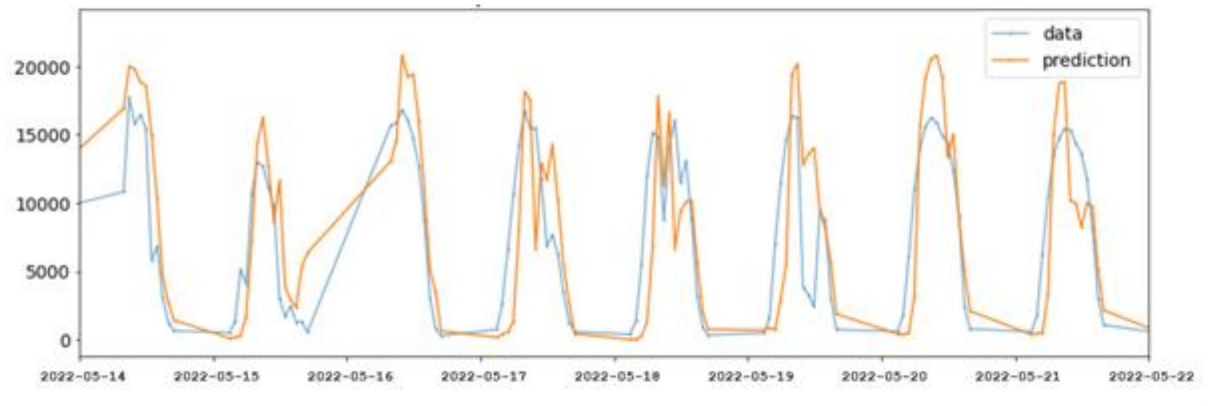


Figure 46: LLUC-4A-01-renewable power production and related forecasting.

Regarding the KPI related to forecast accuracy, this indicator measures the accuracy of forecasting in terms of percentage error with respect to the daily measured energy, both for production and consumption, in particular considering the well-known normalized Root Mean Square Error indicator (nRMSE) and the recently introduced Envelope Mean Absolute Error indicator (EMAE), as described in the formula reported in Annex I.

In particular, this KPI is computed considering the daily power forecast with respect to its daily measurement, as reported in Figure 46. In order to be able to evaluate this KPI, power forecast ($P_{f,t}$) and real measurements ($P_{m,t}$) are collected and stored within PLATOON platform data storage, obtained in real time at each considered time step from the micro-grid monitoring system. The testing period is over 24 hours, thus this KPI is computed with daily frequency, by summing up the power values over the last 24 hours, and the related results (in percentage) displayed into the consolidated dashboard.

While lower percentage values correspond to a successful result in terms of forecasting accuracy, with an ideal target of 0% for this indicator, a threshold of 20% can be considered satisfactory for this forecasting accuracy.

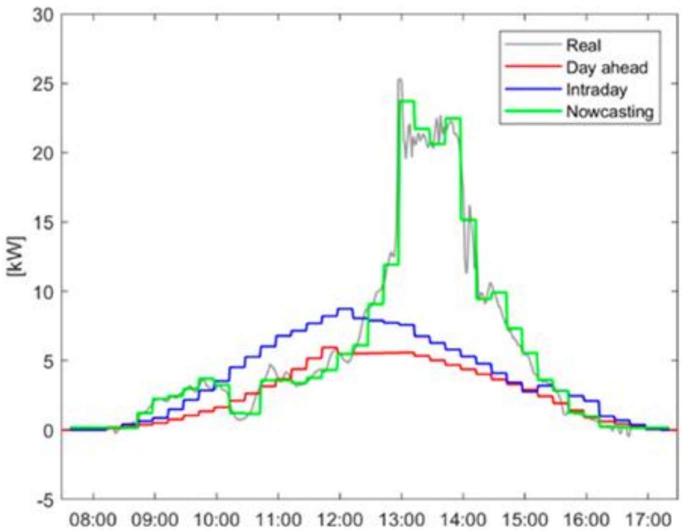


Figure 47: LLUC-4A-01-real time power forecasting adjustments by means of nowcasting technique.

Regarding the KPI related to realtime capability, this indicator measures the ability of the system to monitor, analyse and optimize forecasting results at real time rate, when the prediction for renewable electricity generation is performed considering current weather conditions. In particular, the KPI

related to realtime capability is measured considering the forecast skill of the nowcasting feature with respect to day-ahead forecasts, according to the formula reported in Annex I.
In order to be able to evaluate this KPI, the day-ahead power forecast ($P_{f,t}$) the most updated nowcast values ($P_{n,t}$) and the corresponding real measurements ($P_{m,t}$) are needed, as shown in Figure 47. These data and measurements are stored within PLATOON platform data storage, obtained in real time at each considered time step from the micro-grid monitoring system. The testing range spans across the last 24 hours, thus this KPI is computed with daily frequency and the related results (in percentage) displayed into the consolidated dashboard.
While higher percentage values will correspond to a successful result, with an ideal target of 100% for this indicator, a threshold of 80% can be set as a satisfactory target in terms of realtime capability of nowcasting.

9.3 Conclusion

As a conclusion of this preliminary validation, it can be drawn that the implemented energy management system (EMS) for the experimental microgrid of Politecnico di Milano is reaching the target KPIs regarding the renewable energy generation management and forecasting capabilities. However, some of the KPIs need to be improved and their consistency validated during a longer time range. Finally, the display of all the results needs to be completed in the dashboard to visualize the KPI aggregation on different time horizons.

10. PLATOON Common Components Evaluation & Validation Report

10.1 Introduction

This section covers the validation results of the cross-pilot PLATOON common components.

10.2 Marketplace - IDS Metadata Registry (Broker/Appstore), Clearing House, DAPS and Vocabulary Provider

The PLATOON Marketplace is one common endpoint to access the data and energy services provided by all pilots. PLATOON Marketplace comprehends the following IDS components:

- Metadata Registry
- Graphical User Interface (GUI)
- Clearing House
- Dynamic Attribute Provisioning Service (DAPS)
- Vocabulary Provider

10.2.1Evaluation and Validation

KPI #	Description	Target Value	Actual Value	Comments
1	Metadata Registry Integration	1	0.8	Metadata Registry has been successfully integrated with IDS DAPS and Connectors. The validation of App message handler is still pending which

				will be done before end of the project.
2	GUI Integration	1	0.9	GUI was successfully developed for the Marketplace. Some pilot partners have tested and interacted with the UI and raised some feedback which needs to be considered in future.
3	Clearing House Integration	1	0.6	A new instance of Clearing House based on the latest release by Fraunhofer AISEC has been successfully deployed in the marketplace. Exposing the instance publicly and integration with the connectors is still pending.
4	DAPS Integration	1	1	DAPS has been successfully integrated in the Clearing House and Metadata registry. All components interacting with the marketplace are authenticated through DAPS.
5	Vocabulary Provider Integration	1	0.8	Vocabulary Provider has been successfully integrated with IDS DAPS and Connectors. Integration with PLATOON datamodels is still pending as they have not been uploaded yet to a repository.

The PLATOON Metadata Registry has been successfully integrated with the TRUE Connector. As Figure 48 shows, the locally installed connector is running at `https://localhost:8084` and the Metadata registry is running at `https://localhost:8080`. The "Forward-To" in the header of the message contains the URL of the Metadata Registry. The "ids:MessageProcessedNotificationMessage" shows that the local connector is successfully registered in the Metadata Registry.

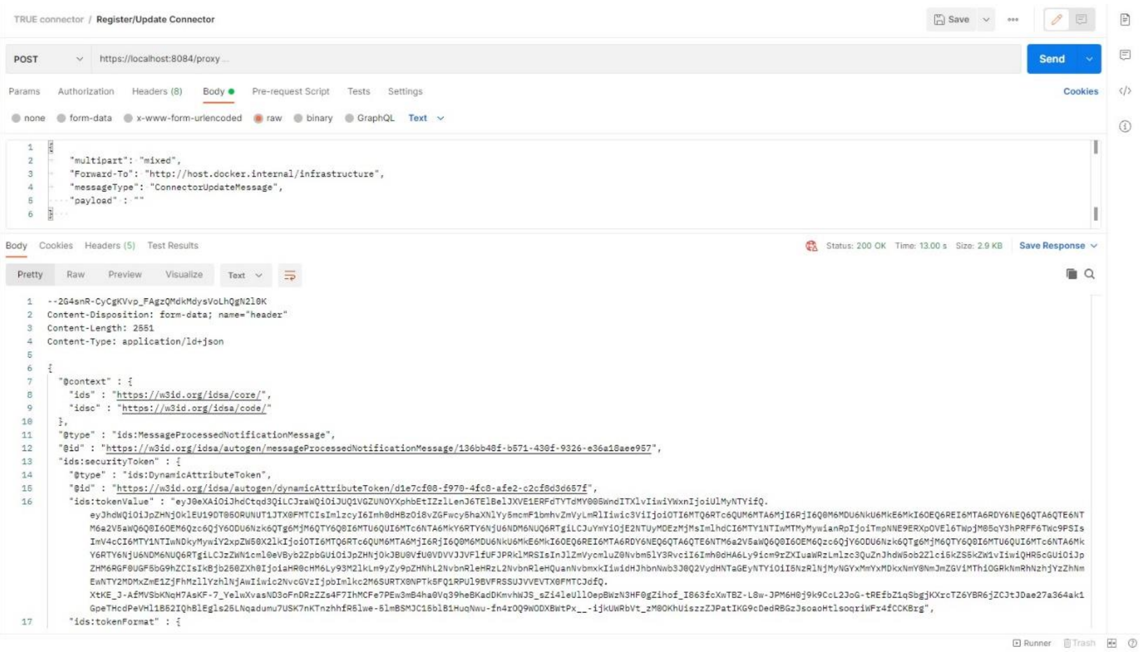


Figure 48:PLATOON Common Components - Integration of TRUE connector into the Metadata Registry

All the IDS components are authenticated through the DAPS. Figure 49 shows that when the upcoming Token from the connector is not Valid, Metadata Registry sends "ids:RejectionMessage" with the "Error verifying token".

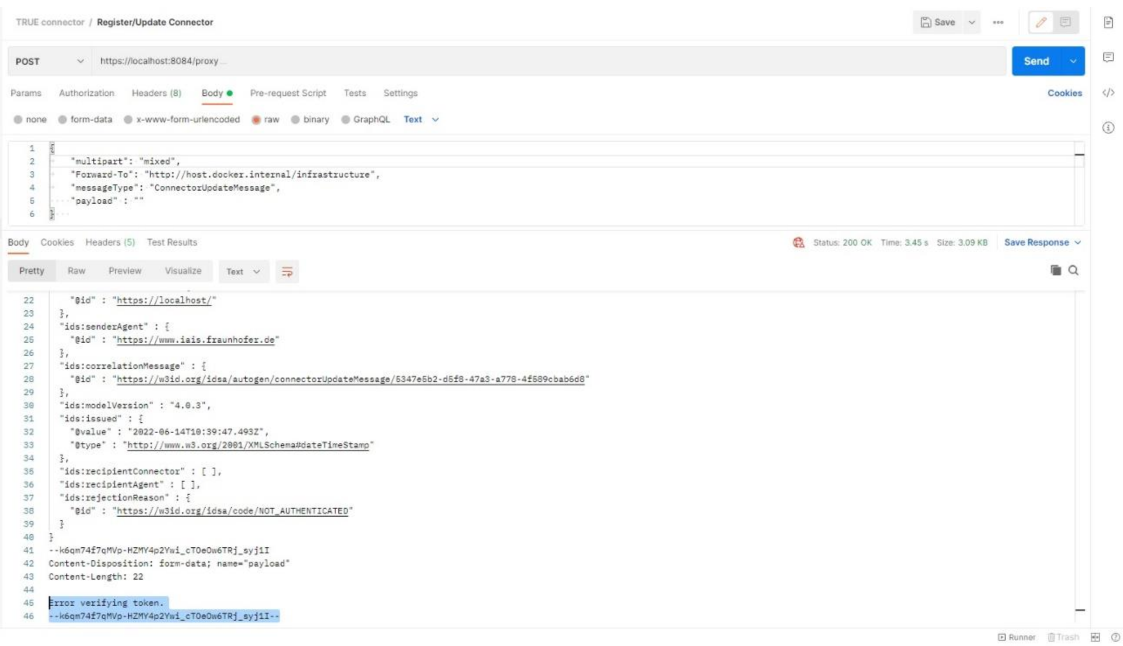


Figure 49: PLATOON Common Components - Integration of DAPS into the Metadata Registry

A GUI specific for PLATOON Marketplace has been developed and integrated into the Metadata Registry. The User Interface contains a Dashboard that shows the summary of all the registered Connectors, Resources, and Apps (services) in the Metadata Registry as shown in Figure 50. The **Datasets** and **Apps** windows shows the list of all the Resources present in the UI. If one clicks on any a dataset or an app, the window will show the details of it as shown in Figure 51.

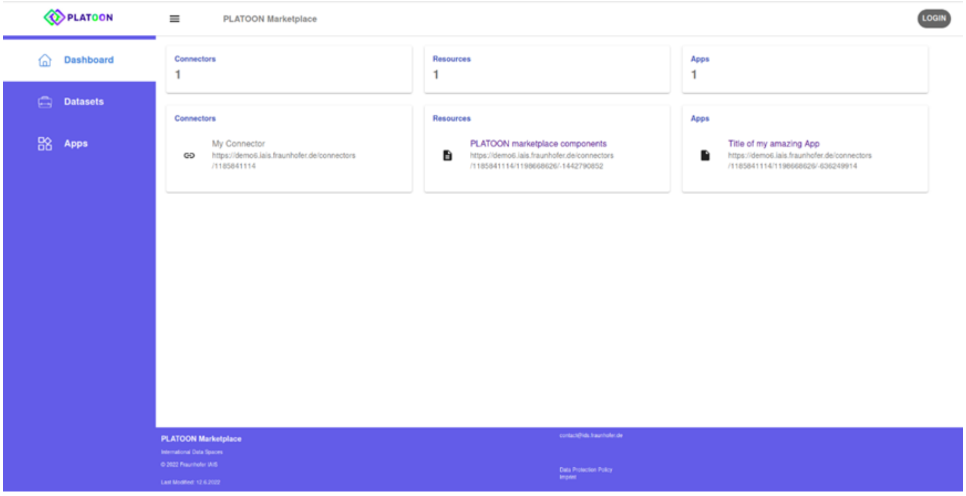


Figure 50: PLATOON Common Components - UI Dashboard

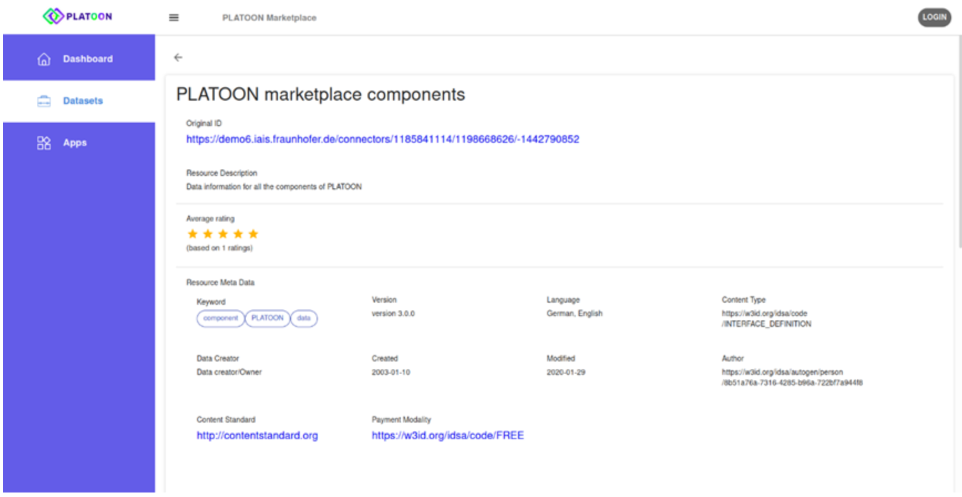


Figure 51: PLATOON Common Components - Dataset Window of the UI

Besides, the newest IDS Clearing House has been deployed for PLATOON. This component successfully integrates the DAPS. Without a proper Dynamic Attribute Token (DAT) coming from the DAPS, the Clearing House will not log anything and through a rejection message as shown in Figure 52.

Figure 52: PLATOON Common Components - Rejection log from the Clearing House when token is not valid

With a proper token with the integration of a Connector, the Clearing House responds with a successful message as an example shown in Figure 53.

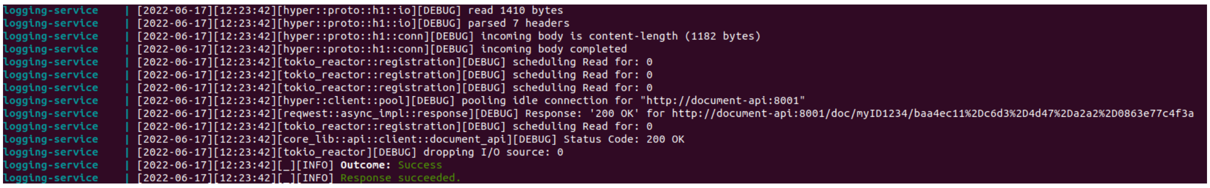


Figure 53: PLATOON Common Components – Successful response message from the Clearing House with respect to the Connector's incoming message

Finally, the PLATOON IDS Vocabulary provider has been successfully integrated with the IDS DAPS as it is able to manage the tokens generated. Equally it has been successfully integrated with the IDS connectors and can receive several IDS messages according to the IDS information model. The figure below shows the response of a query message that allows to interrogate a specific ontology directly from an IDS connector.

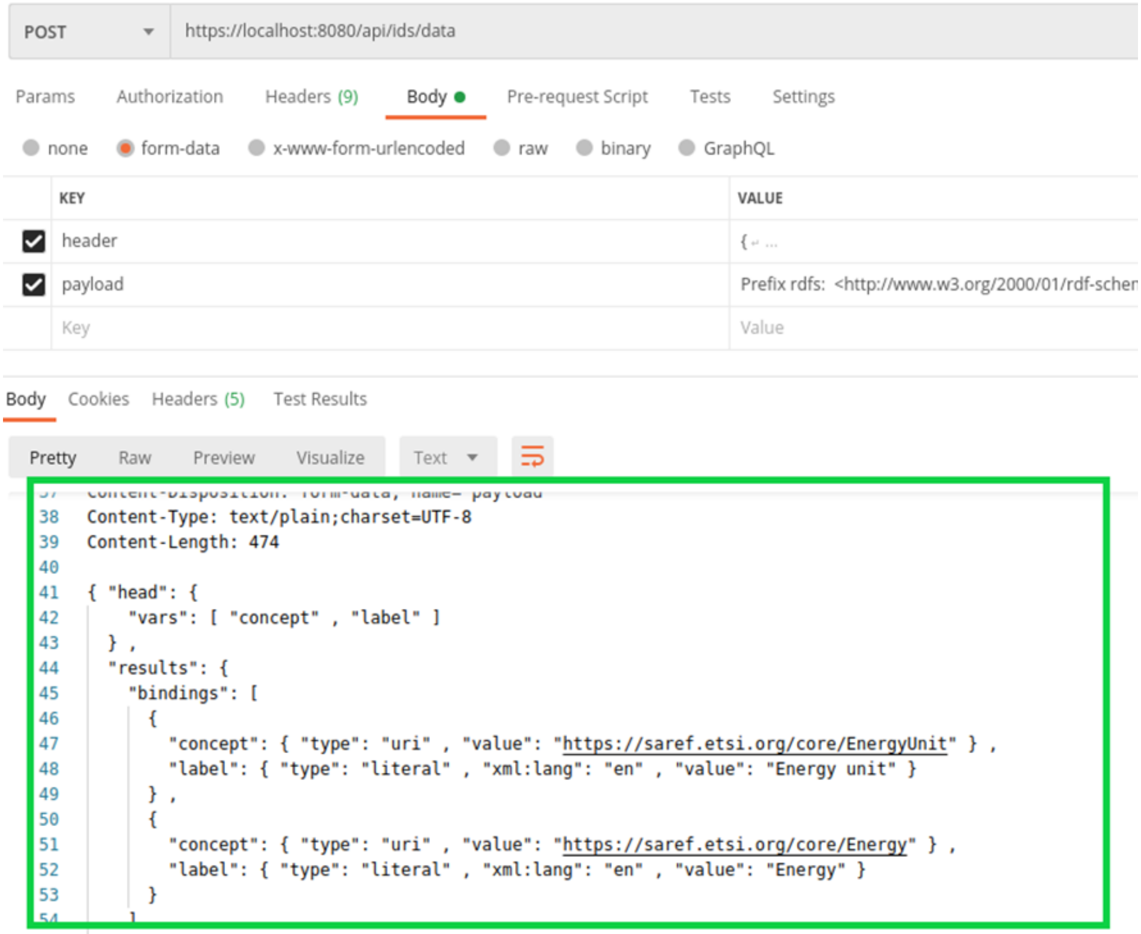


Figure 54: PLATOON Common Components - IDS Vocabulary Provider Validation Results

However, the integration of the PLATOON IDS vocabulary provider with the PLATOON data models is still pending as they have not been uploaded yet to a repository. As part of the exploitation plan we are defining the best way to publish the PLATOON data models. This will be decided by the end of the project and evidence of integration with the Vocabulary Provider will be shown in the V2 deliverable due M36.

10.3 Conclusion

As a result of the validation of the PLATOON Common Components it can be concluded that most functionalities of the Metadata Registry, Clearing House, GUI, DAPS and Vocabulary provider have been successfully validated. However, there are still some issues pending. One of the pending issues is to test the functionality of AppMessageHandler. The plan is to customize a dataspace connector so that the provider connector can register an app in the metadata registry and a consumer can automatically download and implement that app in their connector. Moreover, since a new version of Clearing House has been deployed, the integration with the TRUE connector is still pending and will be done once the Clearing House endpoint is exposed publicly. Regarding the PLATOON IDS Vocabuary provider the integration with the PLATOON data models is still pending as they have not been uploaded yet to a repository. All these pending issues will be solved by the V2 deliverable due M36.

11. Conclusion

As a result of the first validation performed in the different pilots and the PLATOON common components it can be concluded that most of the functionalities have been validated, but there are still some components that need to be validated. In general, the pilots face two main barriers to complete the validation:

1. Implementation of IDS connector and semantic pipeline.
2. Lack of sufficient data.

Regarding the first barrier, the corresponding technical partners are working on it as a high priority task and are planning to solve the pending issues before the summer.
Regarding the second barrier, all the necessary sensors are now installed and the data is being collected, thus, there should be enough data to complete the validation by the final version (V2) due by month M36.

In addition, it can be seen that the situation on the different pilots is not the same. There are some that are more advanced than others. Below it is shown a summary of the status of the different pilots using a colour code (green-on track; yellow – minor pending aspects; red -major pending aspects).

Table 30: Overall Validation Status Summary

Pilot	Status	Pending Aspects
1A	Minor pending aspects	• IDS connector scenario where ENGIE acts as data provider • Semantic adaptation of results from TECN and VUB • Validation of synthetic data and power converter needs to be completed.
2A	On track	All the methods will be validated during different seasons (summer, winter…) and in case of performance degradation, they will be accordingly updated
2B	Minor pending aspects	• IDS connector • Semantic pipeline • Some of the KPIs for LLUC-2B-01 need to be calculated

3A	Major pending aspects	Still ongoing work to get some relevant results and to be able to assess the KPIs
3B-PI	Major pending aspects	KPIs for the LLUC3B-PI-02 use case have not yet been produced
3B-ROM	Major pending aspects	Still ongoing work to get some relevant results and to be able to assess the KPIs
3C	Major pending aspects	KPIs for the LLUC3C-01 use case have not yet been produced
4A	Minor pending aspects	• Some of the KPIs need to be improved and their consistency validated during a longer time range. • The display of all the results needs to be completed in the dashboard to visualize the KPI aggregation on different time horizons.
Common Components	Minor pending aspects	• Metadata registry- Test the functionality of AppMessageHandler. • Clearing House - integration of with the TRUE connector. • Vocabulary provider - integration with the PLATOON data models.

Annex I: KPI Templates

Pilot 1a Predictive Maintenance of Wind Farms

<table>
<tr><td colspan="5">KPI N°1</td></tr>
<tr><td>KPI-Name</td><td colspan="2">Modelling quality</td><td>KPI-ID</td><td>1</td></tr>
<tr><td>KPI-Type</td><td colspan="4">Technical (specific to the pilot use case) or business (refer to D8.1/ PLATOON KPIs)
Technical</td></tr>
<tr><td>Description</td><td colspan="4">Accuracy of the predicted value compared to real value in healthy operating conditions using the Mean Absolute Percentage Error (MAPE).</td></tr>
<tr><td>Target Value</td><td>Target value:
3%</td><td>Threshold Value
5%</td><td colspan="2">The value used to assess the effectiveness/efficiency performance of the monitored process.
RMS error</td></tr>
<tr><td>Rounding</td><td colspan="4">Round to 1%</td></tr>
<tr><td>Unit</td><td colspan="4">Percentage error</td></tr>
<tr><td>Formula</td><td colspan="4">(Abs(predicted value of modelled parameter – true value)/true value) * 100</td></tr>
<tr><td>Calculating frequency</td><td colspan="4">Upon retraining of the model</td></tr>
<tr><td colspan="5">Calculation Methodology</td></tr>
<tr><td>Step</td><td colspan="4">Description</td></tr>
<tr><td>01-</td><td colspan="4">Predict the value of modelled parameter</td></tr>
<tr><td>02</td><td colspan="4">Compare to the real value according to the formula above.</td></tr>
<tr><td colspan="5">Data Source</td></tr>
<tr><td>Data description</td><td>Data source</td><td>Data collection frequency</td><td>Data collection time range</td><td>Data Owner</td></tr>
<tr><td>Signals used as input for the models</td><td>SCADA data</td><td>10 min</td><td>Data corresponding to training range for the model.</td><td>ENGIE</td></tr>
</table>

KPI N°2			
KPI-Name	Integration	**KPI-ID**	2
KPI-Type	Technical		
Description	Metric targeted at the validation of the fact that the tools of this pilot are able to work together.		
Target Value	1	**Threshold Value**	1
Rounding	Not applicable		
Unit	Binary 1 or 0		
Formula	If all tools to complete the pilot data analysis can interact and send data to each other than this KPI is 1. Otherwise, it is 0.		
Calculating frequency	At each pipeline release		
Calculation Methodology			
Step	**Description**		
01-	based on unit tests the input-output functioning of each pipeline is validated.		
02	Test data is exchanged between the pilot analytics blocks		

Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Test data	Predefined set of validation data.			Each pilot party involved with specific tools

KPI N°3			
KPI-Name	Fault detection	**KPI-ID**	3
KPI-Type	Technical		
Description	Anomaly detection speed + accuracy (false vs true positive). The accuracy is expressed using a confusion matrix. For the speed this is expressed in time to catastrophic failure.		
Target Value	Compared to the current failure detection the speed should improve with at least 25%, while keeping false positives below 10%	**Threshold Value**	All improvement compared to current situation is already useful.
Rounding	Not applicable for accuracy. Each element in the confusion matrix is binary. For the speed rounded to the next day.		

Unit	time			
Formula	Confusion matrix for each day block in time			
Calculating frequency	Once per day			
Calculation Methodology				
Step	**Description**			
01-				
02				
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
E.G. energy consumption	E.g. BMS	E.g. 15 min	E.g. Monthly	

KPI N°4			
KPI-Name	Processing capability	**KPI-ID**	4
KPI-Type	Technical		
Description	There are two aspects being tested in this KPI. The first is the speed at which one complete data analysis of the complete pipeline can be done. The second is the number of turbines that are feasible to be analysed using the approach.		
Target Value	Full processing chain for a farm should be able to run on a standard server.	**Threshold Value**	Full processing chain for a farm should be able to run on a standard server.
Rounding	Rounding up of CPU and RAM to next unit		
Unit	Nbr of s on CPU of type X with X Gb RAM for 1 turbine		
Formula	Cores and Gb		
Calculating frequency	Upon changes in the pipelines		
Calculation Methodology			
Step	**Description**		
01-			

02				
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Input data sources for the analytics methods	Input data sources for the analytics methods	Same as inputs for the analytics methods		

KPI N°5			
KPI-Name	Maintenance costs reduction	**KPI-ID**	5
KPI-Type	Business		
Description	The reduction in the maintenance cost of the wind turbine due to early fault detection. Less consequent damages are present and maintenance actions are clustered. Costs will be estimated by comparing cost of component replacement at detection to catastrophic failure. Revenues during additional time that the machine was able to run are subtracted from the maintenance costs.		
Target Value	10-20%	**Threshold Value**	10%
Rounding	Round to 0.01%		
Unit	%		
Formula	Euro maintenance cost with early detection/Euro maintenance cost run to failure		
Calculating frequency	yearly		
Calculation Methodology			
Step	**Description**		
01-			
02			

Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Maintenance records containing the maintenance actions	Maintenance records	continuously	yearly	ENGIE

performed on the wind turbines under investigation				

KPI N°6			
KPI-Name	Availability increase	**KPI-ID**	6
KPI-Type	**Technical (specific to the pilot use case) or business (refer to D8.1/ PLATOON KPIs)**		
Description	The increase of the turbine availability due to faster actions triggered by better predictive maintenance. We focus on machines with an error.		
Target Value	2-5%	**Threshold Value**	2%
Rounding	Round to 0.01%		
Unit	% of the time		
Formula	Abs(Availability as is situation – Availability after usage of Platoon toolbox)		
Calculating frequency	yearly		
Calculation Methodology			
Step	**Description**		
01-	Isolation of the availability reductions linked to the subcomponents within focus in Platoon.		
02	Comparison of the estimated availability with and without the fault detection knowledge of platoon analytics tools.		

Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Annotated stops	Maintenance records	continuously	yearly	ENGIE

Pilot 2a Electricity Balance and Predictive Maintenance

LLUC P 2a-03

KPI N°1a				
KPI-Name	Load Forecasting Mean Absolute Error		**KPI-ID**	LLUC 2a-03 KPI 1a
Description	This KPI is supposed to provide precision performance estimation for Load Forecasting models.			
Unit	[W]			
Formula	$$= \frac{1}{n}\sum_{i=1}^{n} \lvert e_i \rvert$$ where e_i is difference between estimated and real load and n is number of samples for which KPI is calculated.			
Calculating frequency	This KPI should be evaluated daily or monthly			
Calculation Methodology				
Step	**Description**			
01-	Estimated and real load from the PLATOON platform should be obtained and KPI should be calculated according to the formula above.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy consumption	MySQL	hourly	daily, monthly, yearly	IMP

KPI N°1b			
KPI-Name	Load Forecasting Mean Absolute Percentage Error	**KPI-ID**	LLUC 2a-03 KPI 1b
Description	This KPI is supposed to provide precision performance estimation for Load Forecasting models, similarly to the previous one, but normalized.		
Unit	[%]		
Formula	$$= \frac{1}{n}\sum_{i=1}^{n} \frac{\lvert e_i \rvert}{d_i}$$ where e_i is difference between estimated and real load d_i, and n is number of samples for which KPI is calculated.		

Calculating frequency	This KPI should be evaluated daily or monthly			
Calculation Methodology				
Step	**Description**			
01-	Estimated and real load from the PLATOON platform should be obtained and KPI should be calculated according to the formula above.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy consumption	MySQL	hourly	daily, monthly, yearly	IMP

KPI N°2a			
KPI-Name	Load Forecasting Root Mean Square Error	**KPI-ID**	LLUC 2a-03 KPI 2a
Description	This KPI is supposed to provide precision performance estimation for Load Forecasting models.		
Unit	[W]		
Formula	$$= \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}$$ where e_i is difference between estimated and real load, and n is number of samples for which KPI is calculated.		
Calculating frequency	This KPI should be evaluated daily or monthly		
Calculation Methodology			
Step	**Description**		
01-	Estimated and real load from the PLATOON platform should be obtained and KPI should be calculated according to the formula above.		

Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy consumption	MySQL	hourly	daily, monthly, yearly	IMP

KPI N°2b			
KPI-Name	Load Forecasting Root Mean Square Error Percentage	**KPI-ID**	LLUC 2a-03 KPI 2b
Description	This KPI is supposed to provide precision performance estimation for Load Forecasting models, similarly to the previous one, but normalized.		
Unit	[%]		
Formula	$$= \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}}{\frac{\sum_{i=1}^{n} d_i}{n}}$$ where e_i is difference between estimated and real load (d_i), and n is number of samples for which KPI is calculated.		
Calculating frequency	This KPI should be evaluated daily or monthly		
Calculation Methodology			

Step	**Description**			
01-	Estimated and real load from the PLATOON platform should be obtained and KPI should be calculated according to the formula above.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy consumption	MySQL	hourly	daily, monthly, yearly	IMP

LLUC P 2a-04

KPI N°1a			
KPI-Name	Production Forecasting Mean Absolute Error	**KPI-ID**	LLUC 2a-04 KPI 1a
Description	This KPI is supposed to provide precision performance estimation for Production Forecasting models.		
Unit	[W]		
Formula	$$= \frac{1}{n}\sum_{i=1}^{n}\lvert e_i \rvert$$ where e_i is difference between estimated and real production and n is number of samples for which KPI is calculated.		
Calculating frequency	This KPI should be evaluated daily or monthly		
Calculation Methodology			

Step	**Description**			
01-	Estimated and real production from the PLATOON platform should be obtained and KPI should be calculated according to the formula above.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy production	MySQL	hourly	daily, monthly, yearly	IMP

<table>
<tr><td colspan="5">KPI N°1b</td></tr>
<tr><td>KPI-Name</td><td colspan="2">Production Forecasting Mean Absolute Percentage Error</td><td>KPI-ID</td><td>LLUC 2a-04 KPI 1b</td></tr>
<tr><td>Description</td><td colspan="4">This KPI is supposed to provide precision performance estimation for Production Forecasting models, similarly to the previous one, but normalized.</td></tr>
<tr><td>Unit</td><td colspan="4">[%]</td></tr>
<tr><td>Formula</td><td colspan="4">$$= \frac{1}{n}\sum_{i=1}^{n}\frac{|e_i|}{p_i}$$ where e_i is difference between estimated and real production p_i, and n is number of samples for which KPI is calculated.</td></tr>
<tr><td>Calculating frequency</td><td colspan="4">This KPI should be evaluated daily or monthly</td></tr>
<tr><td colspan="5">Calculation Methodology</td></tr>
<tr><td>Step</td><td colspan="4">Description</td></tr>
<tr><td>01-</td><td colspan="4">Estimated and real production from the PLATOON platform should be obtained and KPI should be calculated according to the formula above.</td></tr>
<tr><td colspan="5">Data Source</td></tr>
<tr><td>Data description</td><td>Data source</td><td>Data collection frequency</td><td>Data collection time range</td><td>Data Owner</td></tr>
<tr><td>Energy consumption</td><td>MySQL</td><td>hourly</td><td>daily, monthly, yearly</td><td>IMP</td></tr>
</table>

<table>
<tr><td colspan="4">KPI N°2a</td></tr>
<tr><td>KPI-Name</td><td>Production Forecasting Root Mean Square Error</td><td>KPI-ID</td><td>LLUC 2a-04 KPI 2a</td></tr>
<tr><td>Description</td><td colspan="3">This KPI is supposed to provide precision performance estimation for Production Forecasting models.</td></tr>
<tr><td>Unit</td><td colspan="3">[W]</td></tr>
<tr><td>Formula</td><td colspan="3">$$= \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}$$ where e_i is difference between estimated and real production, and n is number of samples for which KPI is calculated.</td></tr>
<tr><td>Calculating frequency</td><td colspan="3">This KPI should be evaluated daily or monthly</td></tr>
<tr><td colspan="4">Calculation Methodology</td></tr>
<tr><td>Step</td><td colspan="3">Description</td></tr>
</table>

01-	Estimated and real production from the PLATOON platform should be obtained and KPI should be calculated according to the formula above.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy consumption	MySQL	hourly	daily, monthly, yearly	IMP

KPI N°2b			
KPI-Name	Production Forecasting Root Mean Square Error Percentage	**KPI-ID**	LLUC 2a-04 KPI 2b
Description	This KPI is supposed to provide precision performance estimation for Production Forecasting models, similarly to the previous one, but normalized.		
Unit	[%]		
Formula	$= \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2}}{\frac{\sum_{i=1}^{n} p_i}{n}}$ where e_i is difference between estimated and real production (p_i), and n is number of samples for which KPI is calculated.		
Calculating frequency	This KPI should be evaluated daily or monthly		

Calculation Methodology				
Step	**Description**			
01-	Estimated and real production from the PLATOON platform should be obtained and KPI should be calculated according to the formula above.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy consumption	MySQL	hourly	daily, monthly, yearly	IMP

LLUC P 2a-05

KPI N°1			
KPI-Name	Increase in PV insertion capacity	**KPI-ID**	KPI-8

Description	Estimate how many PVs can be integrated into LV grid (and where) before a grid limitation is reached (e.g., overvoltage limit). Increase is compared to actual installed PV capacity on LV grid.			
Unit	%			
Formula	$\frac{P_{max}(Vmax)}{P_{installedPV}} * 100\%$ V_{max} according to EN-50160			
Calculating frequency	Once per installation			
Calculation Methodology				
Step	**Description**			
01-	Obtain the maximal daily grid voltage from PMU			
02	For certain period and for estimated worst case scenario condition estimate max grid Voltage.			
03	Calculate the capacity			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Grid voltage	EMS / PMU	50 Hz	months	

LLUC P 2a-07

KPI N°1			
KPI-Name	Saving costs	**KPI-ID**	KPI-8
Description	Algorithms detects abnormal behaviour and predicts the degreation constant. Reduces maintainance costs. It also detects failures.		
Unit	€		
Formula	1. Binary 0 1 Trigger's detection of failure, immediate replacement $\left(N_{days\ estimate} - N_{days\ after\ detecting failure}\right) * E_{daily} * price_{of\ electricity}$ 2. Prediction of failure Reduction of Asset Investment costs by minimizing the number of elements to be replaced (PV modules). $\left(\sum_{i=0}^{N_{total}} i - \sum_{i=0}^{N_{string}} i\right) * cost\_of\_module$		

Calculating frequency	daily			
Calculation Methodology				
Step	**Description**			
01-	Obtain correction factor for PV from the service			
02	Obtain historical degradation parameter from the service			
03	Check the values for PV plant/string or inverter level			
04	Compared to the predefined threshold (eg. 75% for module efficiency), 0 or 1 for the inverters			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
E.G., energy consumption	E.g., EMS	daily	daily	

Pilot 2b Electricity Grid Stability, Connectivity and Life cycle

LLUC P 2b-01

KPI N°1				
KPI-Name	Temperature estimation accuracy (%)		**KPI-ID**	01
Description	Hourly temperature accuracy estimation based on estimated temperature (ET) and actual (measured) temperature (AT) for top oil.			
Target Value	5%	**Threshold Value**	10%	
Unit	None			
Formula	(Estimated Temperature-Actual Temperature)/ Actual Temperature (%)			
Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Model the top oil temperature using machine learning/deep learning.			
02	Compare the prediction obtained using our model with the real values obtained from the sensor.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°2			
KPI-Name	True positives (TP)	**KPI-ID**	02
Description	Number of anomalies detected with early warnings and confirmed with a corrective work order		
Unit	None		
Formula			

Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Obtain the warnings of needed corrective order given by the model.			
02	Calculate the number of corrective orders that are predicted and applied.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°3				
KPI-Name	False positives (FP)		**KPI-ID**	03
Description	Early warnings with no associated corrective work order			
Unit	None			
Formula				
Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Obtain the warnings of needed corrective order given by the model.			
02	Calculate the number of corrective orders that are predicted but not applied.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°4				
KPI-Name	False negatives (FN)		**KPI-ID**	04
Description	Corrective work order without a previous early warning.			
Unit	None			
Formula				
Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Obtain the warnings of needed corrective order given by the model.			
02	Calculate the number of corrective orders that are not predicted and applied.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°5			
KPI-Name	True Negatives (TN)	**KPI-ID**	05
Description	No early warning and no work order		
Unit	None		
Formula			
Calculating frequency	Hourly		
Calculation Methodology			
Step	**Description**		

01-	Obtain the warnings of needed corrective order given by the model.			
02	Calculate the number of corrective orders that are not predicted and not applied.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°6					
KPI-Name	Specificity (%)		**KPI-ID**	06	
Description	Proportion of true negatives relative to all negative cases.				
Unit					
Formula	(TN/(TN+FP))				
Calculating frequency	Hourly				
Calculation Methodology					
Step	**Description**				
01-	Obtain the proportion of transformers that does not need a corrective order that are correctly identified.				
Data Source					
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner	
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL	

KPI N°7			
KPI-Name	Sensitivity (%)	**KPI-ID**	07
Description	Proportion of actual needed corrective order correctly identified		
Unit	None		

Formula	(TP/(TP+FN))			
Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Obtain the proportion of transformers that need a corrective order that are correctly identified.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°8				
KPI-Name	Cohen's Kappa (%)	**KPI-ID**	08	
Description	Measurement of matches in the predictive tool discounting the probability of randomly matching			
Unit	None			
Formula	$K = \frac{p_0 - p_e}{1 - p_e}$, where $p_0 = \frac{TP+TN}{TP+TN+FP+FN}$ and $p_e = p_{Yes} + p_{No} = \frac{TP+FP}{TP+TN+FP+FN} \frac{TP+FN}{TP+TN+FP+FN} + \frac{FP+TN}{TP+TN+FP+FN} \frac{FN+TN}{TP+TN+FP+FN}$			
Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Calculate the TP,TN,FP,FN.			
02	Apply the formula to obtain the needed corrective orders not well predicted randomly.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°9				
KPI-Name	Savings (€)		**KPI-ID**	09
Description	Cumulative measurement of savings associated to True Positives considering: a) Avoided breakdown consequences + b) Downtime cost			
Unit	€			
Formula				
Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Calculate the breakdown caused by the failure that has been predicted and corrected and the downtime that it should have caused.			
02	Obtain the monetary compensation that this downtime and breakdown should have caused.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°10			
KPI-Name	Additional Costs (€)	**KPI-ID**	10
Description	Increased costs due to maintenance activities associated to False Positives. They should be subtracted from Savings to get the net value.		
Unit	€		
Formula			
Calculating frequency	Hourly		
Calculation Methodology			
Step	**Description**		

01-	Obtain the cost of maintenance caused due to false positives.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°11				
KPI-Name	Anticipation time (days)		**KPI-ID**	11
Description	For each True Positive it represents the delta Time between the moment of detection and the time of failure.			
Unit	Seconds|Minutes|Days			
Formula				
Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Predict the failure dates of the transformers and obtain the difference between the predicted date and the real failure dates.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°12			
KPI-Name	Risk decrease (€)	**KPI-ID**	12
Description	Risk decrease comparing risk-based maintenance supported by the tool to the ordinary preventive maintenance (equal maintenance expenditure is assumed in both cases)		
Unit	€		
Formula			

Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Calculate the ordinary risk of failure and predicted risk of failure. Multiply this by the cost of maintenance.			
02	Obtain the difference between the cost * risk between the tool and the actual maintenance strategy.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°13				
KPI-Name	Maintenance cost savings (€)		**KPI-ID**	13
Description	Maintenance cost savings comparing risk-based maintenance supported by the tool to the ordinary preventive maintenance (equal risk level is assumed in both cases)			
Unit	€			
Formula				
Calculating frequency	Hourly			
Calculation Methodology				
Step	**Description**			
01-	Calculate the costs of ordinary maintenance and predicted maintenance.			
02	Obtain the difference between predicted maintenance cost and ordinary maintenance cost.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

KPI N°14				
KPI-Name	Useful Life Extension (years)		**KPI-ID**	14
Description	Based on the estimation of the RUL (Remaining Useful Time) it indicates the achievable extension of life relative to that indicated by the manufacturer			
Unit	Years/months			
Formula	Previous RUL- loss of life since last RUL calculation			
Calculating frequency	Daily			
Calculation Methodology				
Step	**Description**			
01-	Apply the standards to obtain the HST from the TOT			
02	Apply the standards to calculate the useful life decreasement from the TOT.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Transformer Temperature and load	Transformer Temperature sensors database S02	15 min	unknown	SAMPOL

LLUC P 2b-02

KPI N°1			
KPI-Name	Global Losses Energy Percentage	**KPI-ID**	NTL-KPI-01
Description	Percentage of the energy that is provided from a MV substation or LV CT that is not settle to any consumer and is therefore lost. To be averaged in long periods (at least months).		
Target Value	15%	**Threshold Value**	20%
Unit	None		
Formula	NTL-KPI-01 = NTL-KPI-02 + NTL-KPI-03		
Calculating frequency	Hourly/Daily		

Calculation Methodology				
Step	**Description**			
01-	Calculate the total consumption of all customers.			
02	Calculate the percentage of the customers consumptions over the energy provided by the power transformer.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°2			
KPI-Name	NTL Energy Percentage	**KPI-ID**	NTL-KPI-02
Description	Percentage of the energy that is provided from a MV substation or LV CT that is lost due to NTL		
Target Value	5%	**Threshold Value**	10%
Unit	None		
Formula	NTL-KPI-02 = NTL-KPI-04 + NTL-KPI-05		
Calculating frequency	Hourly/Daily		

Calculation Methodology				
Step	**Description**			
01-	Calculate the NTL caused by consumers and non-consumers.			
02				
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc	S02	1 hour	2016-10-19 00:00:00,	SAMPOL

Bit distribution Grid			2020-10-16 04:00:00,	

KPI N°3			
KPI-Name	TL Energy Percentage	**KPI-ID**	NTL-KPI-03
Description	Percentage of the energy that is provided from a MV substation or LV CT that is lost due to TL		
Unit	None		
Formula	None		
Calculating frequency	Hourly/Daily		
Calculation Methodology			
Step	**Description**		
01-	Obtain the characteristics of the distribution grid.		
02	Calculate the expected technical loses.		

Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°4			
KPI-Name	Customer NTL Energy Percentage	**KPI-ID**	NTL-KPI-04
Description	Percentage of the energy that is provided from a MV substation or LV CT that is lost due to fraud executed by customers. This portion of NTL is more likely to be avoided after it is detected, as legal actions can be taken against the connection point contractors.		
Unit	None		
Formula	None		
Calculating frequency	Hourly/Daily		

Calculation Methodology				
Step	**Description**			
01-	Subtract the technical loses to the total loses.			
02	Obtain the part of the result that can be imputed to customers.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°5					
KPI-Name	Non-Customer NTL Energy Percentage		**KPI-ID**	NTL-KPI-05	
Description	Percentage of the energy that is provided from a MV substation or LV CT that is lost due to fraud executed by non-customers. This energy is stolen by non-permitted connections to the grid, which are difficult to be located physically.				
Unit	None				
Formula	None				
Calculating frequency	Hourly/Daily				
Calculation Methodology					
Step	**Description**				
01-	Subtract the technical loses to the total loses.				
02	Obtain the part of the result that can not be imputed to customers.				
Data Source					
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner	
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL	

KPI N°6				
KPI-Name	True positives (TP)		**KPI-ID**	NTL-KPI-06
Description	Number of customers identified as fraud authors in the NTL identification scenario which are verified to be committing fraud			
Unit	None			
Formula	None			
Calculating frequency	Hourly/Daily			
Calculation Methodology				
Step	**Description**			
01-	Obtain the customers that can be causing NTL using the developed models and identify if they are really causing NTL			
02	Calculate the number of customers that are predicted as causing NTL and are really causing NTL.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°7			
KPI-Name	False positives (FP)	**KPI-ID**	NTL-KPI-07
Description	Number of customers identified as fraud authors in the NTL identification scenario which are not committing fraud, as result of a verification action		
Unit	None		
Formula	None		
Calculating frequency	Hourly/Daily		
Calculation Methodology			
Step	**Description**		
01-	Obtain the customers that can be causing NTL using the developed models and identify if they are really causing NTL		

02	Calculate the number of customers that are predicted as causing NTL and are not causing NTL.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°8				
KPI-Name	False negatives (FN)		**KPI-ID**	NTL-KPI-08
Description	Number of customers which are not identified as fraud authors in the NTL identification scenario but are really committing fraud			
Unit	None			
Formula	None			
Calculating frequency	Hourly/Daily			
Calculation Methodology				
Step	**Description**			
01-	Obtain the customers that can be causing NTL using the developed models and identify if they are really causing NTL			
02	Calculate the number of customers that are predicted as not causing NTL and are really causing NTL.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°9			
KPI-Name	True negatives (TN)	**KPI-ID**	NTL-KPI-09
Description	Number of customers which are not identified as fraud authors in the NTL identification scenario, and are not really committing fraud.		

Unit	None			
Formula	None			
Calculating frequency	Hourly/Daily			
Calculation Methodology				
Step	**Description**			
01-	Obtain the customers that can be causing NTL using the developed models and identify if they are really causing NTL			
02	Calculate the number of customers that are predicted as not causing NTL and are really not causing NTL.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°10			
KPI-Name	Specificity (%)	**KPI-ID**	NTL-KPI-10
Description	Proportion of true negatives relative to all negative cases.		
Unit	None		
Formula	(TN/(TN+FP))		
Calculating frequency	Hourly/Daily		
Calculation Methodology			
Step	**Description**		
01-	Obtain the proportion of negative cases of NTL that are correctly identified.		
02			
Data Source			

Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°11				
KPI-Name	Sensitivity (%)		**KPI-ID**	NTL-KPI-11
Description	Proportion of actual positives cases of NTL correctly identified.			
Unit	None			
Formula	(TP/(TP+FN))			
Calculating frequency	Hourly/Daily			
Calculation Methodology				
Step	**Description**			
01-	Obtain the proportion of positives that are correctly identified.			
02				
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°12			
KPI-Name	Cohen's Kappa (%)	**KPI-ID**	NTL-KPI-12
Description	Measurement of matches in the NTL identification scenario discounting the probability of randomly matching.		
Unit	None		

Formula	$K = \frac{p_0 - p_e}{1 - p_e}$, where $p_0 = \frac{TP+TN}{TP+TN+FP+FN}$ and $p_e = p_{Yes} + p_{No} = \frac{TP+FP}{TP+TN+FP+FN}\frac{TP+FN}{TP+TN+FP+FN} + \frac{FP+TN}{TP+TN+FP+FN}\frac{FN+TN}{TP+TN+FP+FN}$			
Calculating frequency	Hourly/Daily			
Calculation Methodology				
Step	**Description**			
01-	Calculate the TP,TN,FP,FN.			
02	Apply the formula to obtain the NTL identifications not well predicted randomly.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

KPI N°13			
KPI-Name	Economic Savings	**KPI-ID**	NTL-KPI-13
Description	Economic savings due to detected non-technical losses.		
Unit	None		
Formula	None		
Calculating frequency	Hourly/Daily		
Calculation Methodology			
Step	**Description**		
01-	Obtain the costs of energy production and impute the percentage of NTL to this costs.		
02			
Data Source			

Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner

Data obtained from the Parc Bit distribution Grid	S02	1 hour	2016-10-19 00:00:00, 2020-10-16 04:00:00,	SAMPOL

Pilot # 3a Office building: operation performance thanks to physical models and IA algorithms

LLUC P-3a-01

KPI N°1			
KPI-Name	Deviation to target comfort during occupancy time	**KPI-ID**	KPI-1
KPI-Type	Technical/Business		
Description	The thermal comfort in the building is evaluated thanks to air temperature. During occupancy time, the objective is to be within the range of comfort defined by the building manager. The deviations to this range will be monitored during occupancy periods.		
Target Value	0.5°C to comfort range	**Threshold Value**	2°C to comfort range
Rounding	Rounding to 0.01		
Unit	°C		
Formula	During occupancy periods : $$\sum_{t=0}^{nb_{timestep}} \sum_{p=0}^{nb_{point}} \frac{error\left(T(t,p), Target_{range}\right) * w(p)}{nb_{timestep} * nb_{point} * \sum_{p=0}^{nb_{point}} w(p)}$$ T(t,p): temperature of the point p at the timestep t (during occupancy period) w(p) : weight of the point p (if any, default 0) Target_range : Interval of room temperature defined by the building manager that is considered as “acceptable”. Typically : [20°C-25°C] nb_timestep : number of regular timestep (hourly or less) in the period analyzed. nb_point : number of temperature sensor points		
Calculating frequency	According to need : daily, weekly, monthly …		
Calculation Methodology			
Step	**Description**		

01	Choice of a period, or calculation for default periods (days, weeks, months, years)
02	For the given period considered (week, month, year), identification of the occupancy periods for the different zones defined in the building.
03	Request of the temperature for the different occupancy periods of the different zones and application of the formula

Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Occupancy in the different zones	ENGIE IT – data occupancy	15 min	Ongoing in real time	ENGIE
Temperature in the different zones	BMS	15 min ?	Ongoing in real time	ENGIE
Config pilot	config	-	-	ENGIE

KPI N°2			
KPI-Name	Unnecessary HVAC heating emission	**KPI-ID**	KPI-2
KPI-Type	Technical/Business		
Description	Evaluate the amount of energy emission (heating or cooling) that could be considered as unnecessary regarding the actual building occupancy, especially when : ■ Preheating or precooling time over-anticipation ■ Heating/cooling but no one present for the rest of the day. The percentage of valve opening, attributed to a specific weight will be considered as the measure of the unnecessary heating or cooling emission.		
Target Value	<10%	**Threshold Value**	30%
Rounding	Rounding to 0.1%		
Unit	%		
Formula	$$\frac{\sum_{v=0}^{nb_{nb\_valve}} \sum_{t\,\epsilon\,[unecessary\ heating]} Op_h(v,t) * P_{max,h}(v)}{\sum_{v=0}^{nb_{nb_{valve}}} \sum_{t\,\epsilon\,[whole\ period]} Op_h\,(v,t) * P_{max,h}(v)}$$ With : [Unnecessary heating] : ▪ Last period at the end of the day when the zone is unoccupied but heating still happening.		

	▪ First period of the day when the zone is unoccupied, heating is happening, but preheating period is finished (Tzone-Tsetpoint<Tref_lim) Op_h (v,t) : opening of the valve v for heating during the time step t $P_{max,h}$ (v) : Maximum power of the heat emissions behind the valve v
Calculating frequency	According to need : daily, weekly, monthly …
Calculation Methodology	
Step	**Description**
01-	Choice of a period, or calculation for default periods (days, weeks, months, years)
02	Identification of time periods for each valve where : - Last period at the end of the day when the zone is unoccupied, but heating still happening (over anticipation) - First period of the day when the zone is unoccupied, heating is happening, but preheating period is finished (Tzone-Tsetpoint<Tref_lim)
03	For the different periods identified, and for the different zones and valves considered, the above formula can be calculated

Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Temperature in the different zones	BMS	15 min ?	Ongoing in real time	ENGIE
Valve opening in the different zones	BMS	15 min ?	Ongoing in real time	ENGIE
Temperature setpoints in the different zones	BMS	15 min ?	Ongoing in real time	ENGIE
Config pilot	config	-	-	ENGIE

KPI N°3			
KPI-Name	Unnecessary HVAC cooling emission	**KPI-ID**	KPI-2bis
KPI-Type	Technical/Business		
Description	Evaluate the amount of energy emission (heating or cooling) that could be considered as unnecessary regarding the actual building occupancy, especially when : ■ Preheating or precooling time over-anticipation ■ Heating/cooling but no one present for the rest of the day. The percentage of valve opening, attributed to a specific weight will be considered as the measure of the unnecessary heating or cooling emission.		

Target Value	<10%	**Threshold Value**	30%
Rounding	Rounding to 0.1%		
Unit	%		
Formula	$$\frac{\sum_{v=0}^{nb_{nb\_valve}} \sum_{t\, \epsilon\, [necessary\ heating]} Op_c(v,t) * P_{max,c}(v)}{\sum_{v=0}^{nb_{nb_{valve}}} \sum_{t\, \epsilon\, [whole\ period]} Op_c\,(v,t) * P_{max,c}(v)}$$ With : [Unnecessary heating periods] : ▪ Last period at the end of the day when the zone is unoccupied, but heating still happening. ▪ First period of the day when the zone is unoccupied, heating is happening, but preheating period is finished (Tzone-Tsetpoint<Tref_lim) Op_c (v,t) : opening of the valve v for cooling during the time step t $P_{max,c}$ (v) : Maximum power of the cooling emissions behind the valve v		
Calculating frequency	According to need : daily, weekly, monthly …		

Calculation Methodology	
Step	**Description**
01-	Choice of a period, or calculation for default periods (days, weeks, months, years)
02	Identification of time periods for each valve where : - Last period at the end of the day when the zone is unoccupied, but cooling still happening (over anticipation) - First period of the day when the zone is unoccupied, cooling is happening, but precooling period is finished (Tzone-Tsetpoint<Tref_lim)
03	For the different periods identified, and for the different zones and valves considered, the above formula can be calculated

Data Source

Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Temperature in the different zones	BMS	15 min ?	Ongoing in real time	ENGIE
Valve opening in the different zones	BMS	15 min ?	Ongoing in real time	ENGIE
Temperature setpoints in the different zones	BMS	15 min ?	Ongoing in real time	ENGIE

Config pilot	config	-	-	ENGIE

KPI N°4			
KPI-Name	Gain on heating consumption	**KPI-ID**	KPI-3
KPI-Type	Technical/Business		
Description	Climate corrected gain on heating energy consumption in comparison with the consumption of the previous year		
Target Value	>10%	**Threshold Value**	0%
Rounding	0.1%		
Unit	%		
Formula	For a given period : $$\frac{Cs_{ht,p} * HDD\big(p, Text(p)\big) - Cs_{ht,p\_py} * HDD\big(p\_py, Text(p\_py)\big)}{Cs_{ht,p\_py} * HDD\big(p\_py, Text(p\_py)\big)}$$ With : $Cs_{ht,p}$: energy consumption for heating during the period p Cs_{ht,p_py} : energy consumption for heating during the period p but the previous year HDD(p,Text(p)) : Heating degree day for the period p and the external temperature over the period HDD(p,Text(p)) : Heating degree day for the period p of the previous year and the external temperature during this period ***à cf. formula of calculation HDD at the end of the document***.		
Calculating frequency	On request		

Calculation Methodology	
Step	**Description**
01-	Choice of a period, or calculation for default periods (days, weeks, months, years) *à Data of the previous year over the same period has to be available*
02	Application of the formula

Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Energy consumption for heating	BMS	15 min ?	Ongoing in real time	ENGIE

External temperature setpoint	BMS/	1h or less	Defined periods	ENGIE

<table>
<tr><td colspan="5">KPI N°5</td></tr>
<tr><td>KPI-Name</td><td colspan="2">Gain on cooling consumption</td><td>KPI-ID</td><td>KPI-4</td></tr>
<tr><td>KPI-Type</td><td colspan="4">Technical/Business</td></tr>
<tr><td>Description</td><td colspan="4">Climate corrected gain on cooling energy consumption in comparison with the consumption of the previous year</td></tr>
<tr><td>Target Value</td><td>>10%</td><td>Threshold Value</td><td colspan="2">0%</td></tr>
<tr><td>Rounding</td><td colspan="4">0.1%</td></tr>
<tr><td>Unit</td><td colspan="4">%</td></tr>
<tr><td>Formula</td><td colspan="4">For a given period :
$$\frac{Cs_{c,p} * CDD(p, Text(p)) - Cs_{c,p\_py} * CDD(p\_py, Text(p\_py))}{Cs_{c,p\_py} * CDD(p\_py, Text(p\_py))}$$
With :
$Cs_{ht,p}$: energy consumption for cooing during the period p
Cs_{ht,p_py} : energy consumption for cooling during the period p but the previous year
CDD(p,Text(p)) : cooling degree day for the period p and the external temperature over the period
CDD(p,Text(p)) : cooling degree day for the period p of the previous year and the external temperature during this period
⯑ cf. formula of calculation CDD at the end of the document.</td></tr>
<tr><td>Calculating frequency</td><td colspan="4">On request</td></tr>
<tr><td colspan="5">Calculation Methodology</td></tr>
<tr><td>Step</td><td colspan="4">Description</td></tr>
<tr><td>01-</td><td colspan="4">Choice of a period, or calculation for default periods (days, weeks, months, years) ⯑ Data of the previous year over the same period has to be available</td></tr>
<tr><td>02</td><td colspan="4">Application of the formula</td></tr>
<tr><td colspan="5">Data Source</td></tr>
<tr><td>Data description</td><td>Data source</td><td>Data collection frequency</td><td>Data collection time range</td><td>Data Owner</td></tr>
<tr><td>Energy consumption for cooling</td><td>BMS</td><td>15 min ?</td><td>Ongoing in real time</td><td>ENGIE</td></tr>
</table>

External temperature setpoint	BMS/	1h or less	Defined periods	ENGIE

LLUC P-3a-02: Provide Demand Response services through building inertia and HVAC controls

KPI N°1			
KPI-Name	Mean error on heating load prediction	**KPI-ID**	KPI-1
KPI-Type	Technical/Business		
Description	Mean error (%) on the HVAC heating load prediction calculated every 30min as the errors between the predicted and the realized energy consumption and the predicted one (when HVAC is operating).		
Target Value	Error <10%	**Threshold Value**	Mean error above 20%
Rounding	0.1%		
Unit	%		
Formula	$$\sum_{t=0}^{nb_{timetep}} \frac{Cs_{ht,model}(t) - Cs_{ht,real}(t)}{Cs_{ht,real}(t) * nb_{timestep}}$$ With : $Cs_{ht,model}(t)$: heating consumption predicted by the model for the timestep t $Cs_{ht, real}(t)$: real heating consumption measured for the timestep t		
Calculating frequency	Once, daily, weekly, monthly …		
Calculation Methodology			
Step	**Description**		
01-	Choice of a period, or calculation for default periods (days, weeks, months, years)		
02	Application of formula		
Data Source			

Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Energy consumption for heating	BMS	30 min	Ongoing in real time	ENGIE

Predicted energy consumption	Platoon tool	30 min	-	ENGIE

KPI N°2			
KPI-Name	Mean error on cooling load prediction	**KPI-ID**	KPI-1bis
KPI-Type	Technical/Business		
Description	Mean error (%) on the HVAC cooling load prediction calculated every 30min as the errors between the predicted and the realized energy consumption and the predicted one (when HVAC is operating).		
Target Value	Error <10%	**Threshold Value**	Mean error above 20%
Rounding	0.1%		
Unit	%		
Formula	$$\sum_{t=0}^{nb_{timetep}} \frac{Cs_{c,model}(t) - Cs_{c,real}(t)}{Cs_{c,real}(t) * nb_{timestep}}$$ With : $Cs_{c,model}(t)$: cooling consumption predicted by the model for the timestep t $Cs_{c, real}(t)$: real cooling consumption measured for the timestep t		
Calculating frequency	Once, daily, weekly, monthly …		

Calculation Methodology	
Step	**Description**
01-	Choice of a period, or calculation for default periods (days, weeks, months, years)
02	Application of formula

Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Energy consumption for cooling	BMS	30 min	Ongoing in real time	ENGIE
Predicted energy consumption	Platoon tool	30 min	-	ENGIE

KPI N°3

KPI-Name	95-percentile error on heating load prediction	**KPI-ID**	KPI-2
KPI-Type	Technical/Business		
Description	95-percentile Error on the HVAC heating load prediction calculated every 30min as the errors between the predicted and the realized energy consumption and the predicted one (when HVAC is operating).		
Target Value	Error <20%	**Threshold Value**	Mean error above 40%
Rounding	0.1%		
Unit	%		
Formula	Error on each timestep $$Err(t) = \frac{Cs_{ht,model}(t) - Cs_{ht,real}(t)}{Cs_{ht,real}(t)}$$ Then, identification of the 95-percentile of the Err(t) over the period With : $Cs_{ht,model}(t)$: heating consumption predicted by the model for the timestep t $Cs_{ht,\ real}(t)$: real heating consumption measured for the timestep t		
Calculating frequency	Once, daily, weekly, monthly …		
Calculation Methodology			
Step	**Description**		
01-	Choice of a period, or calculation for default periods (days, weeks, months, years)		
02	Application of the formula		

Data Source

Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Energy consumption for heating	BMS	30 min	Ongoing in real time	ENGIE
Predicted energy consumption	Platoon tool	30 min	-	ENGIE

KPI N°4			
KPI-Name	95-percentile error on cooling load prediction	**KPI-ID**	KPI-2bis
KPI-Type	Technical/Business		

Description	Error (%) on the HVAC cooling load prediction calculated every 30min as the errors between the predicted and the realized energy consumption, divided by the predicted one (when HVAC is operating). The error can be characterized over the period: mean, standard deviations, daily distribution, seasonal distribution.		
Target Value	Error <20%	**Threshold Value**	Mean error above 40%
Rounding	0.1%		
Unit	%		
Formula	Error on each timestep $$Err(t) = \frac{Cs_{c,model}(t) - Cs_{c,real}(t)}{Cs_{c,real}(t)}$$ Then, identification of the 95-percentile of the Err(t) over the period With : $Cs_{c,model}(t)$: cooling consumption predicted by the model for the timestep t $Cs_{c, real}(t)$: real cooling consumption measured or the timestep t		
Calculating frequency	Once, daily, weekly, monthly …		

Calculation Methodology

Step	**Description**
01-	Choice of a period, or calculation for default periods (days, weeks, months, years)
02	Application of formula

Data Source

Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Energy consumption for cooling	BMS	30 min	Ongoing in real time	ENGIE
Predicted energy consumption	Platoon tool	30 min	-	ENGIE

KPI N°5

KPI-Name	Error on the flexibility prediction	**KPI-ID**	KPI-4
KPI-Type	Technical/Business		
Description	Error (%) on the prediction of “flexibility available” on the building, in term of time of interruption of heating or cooling in the building, during flexibility event implemented in the building.		

Target Value	Target : 10%	**Threshold Value**	30%
Rounding	0.1%		
Unit	%		
Formula	$$\frac{\mathrm{Time}_{\mathrm{int,model}} - \mathrm{Time}_{\mathrm{int,real}}}{\mathrm{Time}_{\mathrm{int,real}}}$$ Time_(int,model) : time of interruption planned in the model Time_(int, real) : actual time of interruption that was actually implemented in the building.		
Calculating frequency	After interruption event …		

Calculation Methodology	
Step	**Description**
01-	Application of the formula

Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Predicted time of interruption	Platoon tool	30 min	-	ENGIE
Interruption	BMS	-	-	ENGIE

KPI N°6			
KPI-Name	Mean error on HVAC load prediction for days with load shifting programs	**KPI-ID**	KPI-5
KPI-Type	Technical/Business		
Description	Mean error (%) on the HVAC load prediction calculated every 30min as the errors between the predicted and the realized energy consumption and the predicted one (when HVAC is operating), in case of the implementation of a load shifting program (not the usual building operation)		
Target Value	Error <10%	**Threshold Value**	20%
Rounding	0.1%		
Unit	%		
Formula	$$\sum_{t=0}^{nb_{timetep}} \frac{Cs_{ht,c,model}(t) - Cs_{ht,c,real}(t)}{Cs_{ht,c,real}(t) * nb_{timestep}}$$		

	With : $Cs_{ht,c,model}(t)$: heating or cooling consumption predicted by the model for the timestep t $Cs_{ht, real}(t)$: real heating or cooling consumption measured for the timestep t
Calculating frequency	For a day after implementation of load shifting program
Calculation Methodology	
Step	**Description**
01	Application of the formula

Data Source

Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Energy consumption for heating	BMS	30 min	Ongoing in real time	ENGIE
Energy consumption for cooling	BMS	30 min	Ongoing in real time	ENGIE
Predicted energy consumption	Platoon tool	30 min	-	ENGIE

Pilot 3b - PI Advanced Energy Management System and Spatial (Multi-Scale) Predictive Models in the Smart City

LLUC-01

<table>
<tr><td colspan="5">KPI N°1</td></tr>
<tr><td>KPI-Name</td><td colspan="2">Forecast Error</td><td>KPI-ID</td><td>PI_KPI01</td></tr>
<tr><td>KPI-Type</td><td colspan="4">Technical</td></tr>
<tr><td>Description</td><td colspan="4">The KPI calculates the % of deviation between the energy consumption forecast and the actual consumption in the building.
The KPI checks how closely the predictive model adheres to reality - <u>Effectiveness</u></td></tr>
<tr><td>Target Value</td><td>+-5%</td><td>Threshold Value</td><td colspan="2">+-20%</td></tr>
<tr><td>Rounding</td><td colspan="4">round off to 0% for values between 0.00 and 0.49 and to 1% for values above</td></tr>
<tr><td>Unit</td><td colspan="4">Kilowatt per hour (KWh)</td></tr>
<tr><td>Formula</td><td colspan="4">$$\mathbf{FE_i} = \frac{(F_{F,i} - F_{A,i})}{F_{A,i}} * 100$$
$$\mathbf{FE_M} = \frac{1}{N}\sum_i FE_i$$
FE_i= Forecast Error % of building “i”
$F_{F,i}$ = Forecasted value of building “i”
$F_{A,i}$ = Actual value of building “i”
N = number of buildings utilized for the KPI calculation</td></tr>
<tr><td>Calculating frequency</td><td colspan="4">Weekly (Alert if Threshold Value is exceeded)</td></tr>
<tr><td colspan="5">Calculation Methodology</td></tr>
<tr><td>Step</td><td colspan="4">Description</td></tr>
<tr><td>01</td><td colspan="4">Select the time range and the specific building and the perimeter of calculation:
1. Total energy consumption
or
2. Total energy consumption for Specific line (cooling/heating) in the building</td></tr>
<tr><td>02</td><td colspan="4">Calculate the forecast taking into account:
– Real data consumption
– Temperature and Humidity (internal and external)
– Number of Customers and Employees
– Building open hours and shift
– Building Climate Zone, m^3</td></tr>
</table>

03	Get the Real consumption data (of the target month) taking into account: - The full month active energy consumption (Total Active Energy) of the selected building or of a specific line (Detailed Energy Consumption)			
04	Apply the formula			
05	The formula will be applied for each one of the selected buildings, then arithmetic mean will be calculated from these selected values.			
Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Office Registry	Building Data	At starting up and then occasionally, when changes occur	No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Total Active Energy consumption	Energy Consumption	Monthly		Poste Italiane
Detailed Energy Consumption DL_102	Energy Consumption	Monthly	TBD	Poste Italiane
Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2018	Poste Italiane
Temperature, Humidity	Weather		From 01/01/2018	External Services
Customers Number	Occupancy	Monthly	From 01/01/2018	Poste Italiane
Employees Number	Occupancy	Monthly	From 01/01/2018	Poste Italiane

KPI N°2			
KPI-Name	Building Benchmarking BtI_LY	**KPI-ID**	PI_KPI02
KPI-Type	Business		
Description	The KPI calculate, in % value, the difference in Energy consumption of a building with itself during the time. The comparison will be made with the previous year consumption		
Target Value	+-10%	**Threshold Value**	+-20%
Rounding	round off to 0% for values between 0.00 and 0.49 and to 1% for values above		
Unit	Kilowatt per hour (KWh)		
Formula			

	$$\mathbf{B}_{\mathbf{BTI_{LY},i}} = \frac{(\mathrm{EC}_{\mathrm{y,i}} - \mathrm{EC}_{\mathrm{y-1,i}})}{\mathrm{EC}_{\mathrm{y-1,i}}} * 100$$ $$\mathbf{B}_{\mathbf{BTI_{LY},M}} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}} \mathrm{B}_{\mathrm{BTI_{LY},i}}$$ BBTILY,i = Building “i” last year comparison (with itself) ECy,i = Energy Consumption in the time range for building “i” ECy-1, i= Energy Consumption in the same time range of the previous year for building “i” N = number of buildings utilized for the KPI calculation			
Calculating frequency	Weekly			
Calculation Methodology				
Step	**Description**			
01	The calculation takes into account: a. The time range (reference week) b. The time range for benchmark (the same week of the previous year) c. The building d. The perimeter of the analysis: Total energy consumption, or, where available the energy consumption of heating or cooling, lighting			
02	The formula will be applied for each one of the selected buildings, then arithmetic mean will be calculated from these selected values.			
03	Normalize both the consumptions by the comfort level (where available) Comfort level is a range of internal temperature and humidity that must be complied Comfort level = f (internal temperature, internal humidity) (Internal humidity, internal temperature) = f (external humidity, external temperature)			
Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Office Registry	Building Data	At starting up and then occasionally, when changes occur	No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2018	Poste Italiane
Detailed Energy Consumption DL_102	Energy Consumption	Monthly	TBD	Poste Italiane
Total Active Energy consumption	Energy Consumption	Monthly	From 01/01/2018	Poste Italiane
Temperature, Humidity	Weather			External Services

KPI N°3			
KPI-Name	Building Benchmarking BtI_LWs	**KPI-ID**	PI_KPI03
KPI-Type	Business		
Description	The KPI calculate, in % value, the difference in Energy consumption of a building with itself during the time. The comparison will be made with the two previous weeks comsumptions		
Target Value	+-10%	**Threshold Value**	+-20%
Rounding	round off to 0% for values between 0.00 and 0.49 and to 1% for values above		
Unit	Kilowatt per hour (KWh)		
Formula	$$\mathbf{B}_{\mathbf{BTI_{LW},i}} = \frac{(\mathrm{EC_{W,i}} - [\frac{\mathrm{EC_{W-1,i}} + \mathrm{EC_{W-2,i}}}{2}]}{(\mathrm{EC_{W-1,i}} + \mathrm{EC_{W-2,i}})/2} * 100$$ $$\mathbf{B}_{\mathbf{BTI_{LW},M}} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}} \mathrm{B_{BTI_{LW,i}}}$$ BBTILY,i = Building “i” comparison with two last weeks (with itself) ECw,i = Energy Consumption in the time range for building “i” ECW-1,i , ECW-2,i = Energy Consumption of the two previous weeks for building “i” N = number of buildings utilized for the KPI calculation		
Calculating frequency	Weekly		
Calculation Methodology			
Step	**Description**		
01	The calculation takes into account: 2.1.1. The time range (reference week) 2.1.2. The time range for benchmark (the two previous weeks) 2.1.3. The building 2.1.4. The perimeter of the analysis: Total energy consumption, or, where available the energy consumption of heating or cooling, lighting		
02	The formula will be applied for each one of the selected buildings, then arithmetic mean will be calculated from these selected values.		
03	Normalize both the consumptions by the comfort level (where available) Comfort level is a range of internal temperature and humidity that must be complied Comfort level = f(internal temperature, internal humidity) (Internal humidity, internal temperature) = f (external humidity, external temperature)		

Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**

Office Registry	Building Data	At starting up and then occasionally, when changes occur	No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2018	Poste Italiane
Detailed Energy Consumption DL_102	Energy Consumption	Monthly	TBD	Poste Italiane
Total Active Energy consumption	Energy Consumption	Monthly	From 01/01/2018	Poste Italiane
Temperature, Humidity	Weather			External Services

KPI N°4			
KPI-Name	Building Benchmarking BtB	**KPI-ID**	PI_KPI04
KPI-Type	Business		
Description	The KPI calculate, in % value, the difference in Energy consumption between a cluster of buildings.		
Target Value	+-10%	**Threshold Value**	+-20%
Rounding	round off to 0% for values between 0.00 and 0.49 and to 1% for values above		
Unit	Kilowatt per hour (KWh)		
Formula	$$\mathbf{B_{BTB,i}} = \frac{\frac{B_i}{m_i^3} - (\sum_j \frac{B_j}{m_j^3})/(n-1)}{(\sum_j \frac{B_j}{m_j^3})/(n-1)} * 100$$ $$\mathbf{B_{BTB,M}} = \frac{1}{N}\sum_i B_{BTB,i}$$ $B_{BTB,i}$ = Building Energy Consumption comparison with the mean of the same cluster B_i = Energy Consumption of Building “i” B_j = = Energy Consumption of Building “j” (Some cluster of “i”, i.e., some typology and destination use) n = number of buildings in the cluster m^3 = volume of building N = number of buildings utilized for the KPI calculation		

Calculating frequency	Weekly (Alert if Threshold Value is exceeded)
Calculation Methodology	
Step	**Description**
01	The calculation takes into account: 1. The time range (current year last week) 2. The building types 3. The building's destination uses 4. the perimeter of the analysis: Total energy consumption or, where available the energy consumption of heating or cooling, lighting
02	The formula will be applied for each one of the selected buildings, then arithmetic mean will be calculated from these selected values.
03	Normalize the consumptions of both the buildings by the comfort level (where available). Comfort level* is a range of internal temperature and humidity that must be complied. *Comfort level = f(internal temperature, internal humidity) (Internal humidity, internal temperature) = f(external humidity, external temperature)

Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Office Registry	Building Data	At starting up and then occasionally, when changes occur	No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2018	Poste Italiane
Detailed Energy Consumption DL_102	Energy Consumption	Monthly	TBD	Poste Italiane
Total Active Energy consumption	Energy Consumption	Monthly	From 01/01/2018	Poste Italiane
Temperature, Humidity	Weather		From 01/01/2018	External Services

KPI N°5			
KPI-Name	CO2 emission reduction	**KPI-ID**	PI_KPI05
KPI-Type	Business		
Description	The KPI calculate the impact of energy consumption reduction on CO2 emissions in a time range		
Target Value	≥ 10%	**Threshold Value**	0≤ $\Delta(\mathbf{CO_2})_{\mathrm{y,M}}$ <10%

Rounding	round off to 0 for values between 0.00 and 0.49 and to 1 for values above
Unit	Kg
Formula	$\Delta(\mathbf{KWh})_{\mathrm{y,i}} = \frac{\text{Budget (Kwh)}_{\mathrm{y,i}} - \text{Consumption (Kwh)}_{\mathrm{y,i}}}{\text{Consumption (Kwh)}_{\mathrm{y,i}}} * 100$ $\Delta(\mathbf{KWh})_{\mathbf{y,i}} = \Delta(\mathbf{CO_2})_{\mathbf{y,i}}$ Because $C_{O2}\ (\text{Kg}) = 0{,}36099 * \text{Energy (KWh)}$ Finally $\Delta(\mathbf{CO_2})_{\mathrm{y,M}} = \frac{1}{\mathrm{M}} \sum_{\mathrm{i}} \Delta(\mathrm{CO_2})_{\mathrm{y,i}}$ Budget (kWh)y,i = yearly budget of building “i” Consumption (kWh)y,I = yearly consumption of building “i” Δ(KWh) y,i = consumption saving percentage of building “i” Δ (CO_2) y,I = CO_2 saving percentage of building “i” M = number of buildings utilized for the KPI calculation
Calculating frequency	Yearly
Calculation Methodology	
Step	**Description**
01	Calculate the yearly total consumption of the building
02	The formula will be applied for each one of the selected buildings, then arithmetic mean will be calculated from these selected values.

Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Office Registry	Building Data	At starting up and then occasionally, when changes occur	Static - No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Total Active Energy consumption	Energy Consumption	Monthly	From 01/01/2018	Poste Italiane

LLUC-02

KPI N°6				
KPI-Name	Recall		**KPI-ID**	PI_KPI06
KPI-Type	Technical			
Description	The KPI measures the number of cases which correctly classified as problematic (True Positives) by the algorithm divided by the sum of the cases that were classified as normal but actually were problematic (False Positives) plus the number of True Positives.			
Target Value	90%	**Threshold Value**	>=80%	
Rounding	N/A			
Unit	Adimensional			
Formula	$$Recall = \frac{TruePositives}{TruePositives + FalsePositives}$$			
Calculating frequency	Monthly			
Calculation Methodology				
Step	**Description**			
01	The time range comprises the historical data up to the month chosen for the analysis			
02	2.5.3. Identify all those cases where correctly identified (TruePositives) as abnormalities in the Heating and Cooling system and those which are classified as normal but are cases with anomalous behaviors (False Negatives).			
03	Apply the formula			
04	The formula will be applied for each one of the selected buildings.			
Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Office Registry	Building Data	At starting up and then occasionally, when changes occur	Static - No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane

Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2018	Poste Italiane
Alarms of abnormal behaviours of the systems	System Fault	Daily	From June 2021	Poste Italiane
Temperature, Humidity	Weather		From 01/01/2018	External Services
Systems Registry	Building Systems	At starting up and then occasionally, when changes occur	No Temporal Range	Poste Italiane

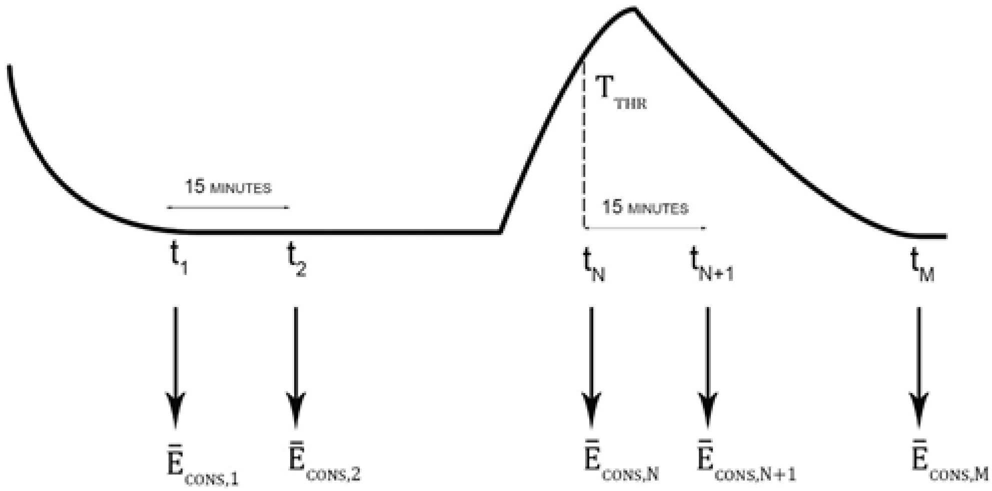
KPI N°7			
KPI-Name	Precision	**KPI-ID**	PI_KPI07
KPI-Type	Technical		
Description	The KPI measures Pre-MATE's performance. Precision, is defined as the ratio of all cases that are correctly identified as problematic (True Positives) to all cases that are identified as problematic, even if they are not, actually (All Positives-True and False).		
Target Value	90%	**Threshold Value**	>=80%
Rounding	N/A		
Unit	Adimensional		
Formula	$$Precision = \frac{TruePositives}{TruePosives + FalsePositives}$$		
Calculating frequency	Monthly		
Calculation Methodology			
Step	**Description**		
01	The time range comprises the historical data up to the month chosen for the analysis		
02	Identify all those cases where correctly identified (True Positives) as abnormalities in the Heating and Cooling system and those which are classified as problematic but are actually normal behaviors (False Negatives)		
03	Apply the formula		
04	The formula will be applied for each one of the selected buildings.		
Data Source			

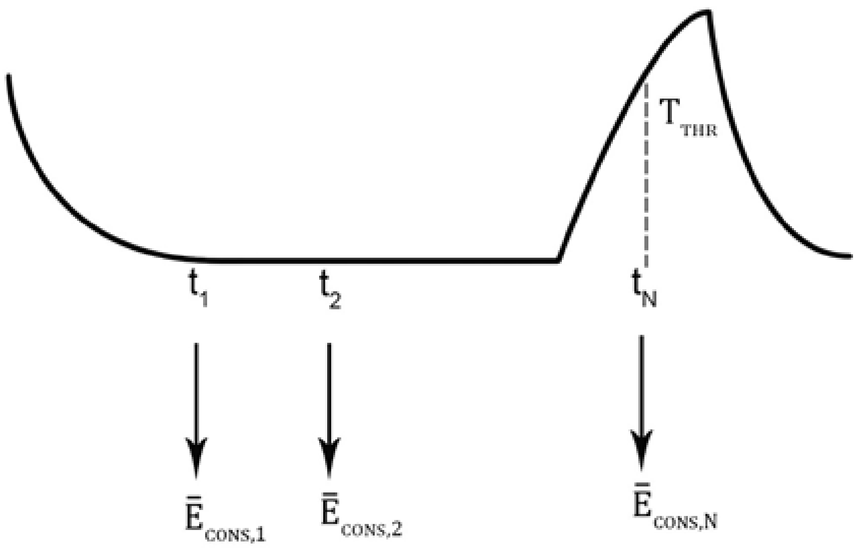
Data description	Data source	Data collection frequency	Data collection time range	Data Owner
Office Registry	Building Data	At starting up and then occasionally, when changes occur	No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2018	Poste Italiane
Alarms of abnormal behaviours of the systems	Systems Fault	Daily	From June 2021	Poste Italiane
Temperature, Humidity	Weather		From 01/01/2018	External Services
Systems Registry	Building Systems	At starting up and then occasionally when changes occur	No Temporal Range	Poste Italiane

KPI N°8			
KPI-Name	F1-Score	**KPI-ID**	PI_KPI08
KPI-Type	Technical		
Description	The KPI is used in cases where the best combination of precision and recall is desired. F_1 score could be used to combine the two criteria. The F_1 score is the harmonic mean of precision and recall, using the formula below to account for both metrics		
Target Value	90%	**Threshold Value**	>=80%
Rounding	N/A		
Unit	Adimensional		
Formula	$$F_1 = 2\frac{Precision \cdot Recall}{Precision + Recall}$$		
Calculating frequency	Bi-Monthly		
Calculation Methodology			
Step	**Description**		

01	The time range comprises the historical data up to the month chosen for the analysis			
02	Calculate Recall and Precision KPIs before			
03	Apply the formula			
04	The formula will be applied for each one of the selected buildings			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Office Registry	Building Data	At starting up and then occasionally when changes occur	No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2018	Poste Italiane
System Registry	Building Systems	At starting up and then occasionally, when changes occur	No Temporal Range	Poste Italiane
Employees Number	Occupancy	Monthly	From 01/01/2018	Poste Italiane

KPI N°09			
KPI-Name	Performances Analysis	**KPI-ID**	PI_KPI09
KPI-Type	Technical/Business		
Description	This KPI measures the energy consumed by the conditioning systems for returning to optimal internal temperature, normalized for the temperature recover range.		
Target Value	5% (month to month increase)	**Threshold Value**	10% (month to month increase)
Rounding	No		
Unit	KWh		
Formula			

	$$\mathbf{E_{cond,k}} = \frac{1}{M-N+1}\sum_{j=N}^{M}\frac{(\bar{E}_{cons,k,j} - E_{bias,k})}{\lvert T_{thr} - T_{int,k}\rvert \times volume \times p_j}$$ *where* $$\mathbf{E_{bias,k}} = \frac{1}{N}\sum_{n=1}^{N}\bar{E}_{cons,n}$$ $$\mathbf{p_j} = \frac{\#\ of\ sensor\ out\ of\ range, at\ the\ time\ 'j'}{total\ number\ of\ sensors\ in\ the\ building}$$ *finally* $$\mathbf{E_{cond i\_m}} = \frac{1}{K}\sum_{k}\mathbf{E_{cond,ki\_m}}$$ = normalized energy the conditioning systems consume per unit of volume and temperature to bring the internal temperature back to the normal range, for the temperature violation 'k'. (is costituited by M-N+1 fifteen minutes interval see figure below)  = energy consumed for conditioning in the optimal range of temperature (from t_1 to t_2) and till the threshold reached(from t_2 to t_N), when is on range or above (see figure below) (19° or above in the heating tabulated period, 27° or below in the cooling tabulated period), for the temperature violation 'k')

	 when is on range or above (19° or above in the heating tabulated period, 27° or below in the cooling tabulated period), for the temperature violation 'k') = Energy consumption at time 'j' = temperatute threshold : they are tabulated : 19° (in the heating tabulated period) ; 27° (in the cooling tabulated period) = max out of range internal temperature to recover for the temperature violation 'k' volume= building volume = normalized energy the conditioning systems consume per unit of volume and temperature to bring the internal temperature back to the normal range, for the month 'm' and the building 'i' K = number of violations for the month 'm' and the building 'i' = ratio between the number of sensor out of range at the time 'j and the total number of sensors in the building
Calculating frequency	Monthly
Calculation Methodology	
Step	**Description**
01	Data have to be taken on building type = Smart building, in the total time window of availability of data. Data must be considered only for days and hours in which the buildings are open
03	Calculate the plants consumption taking into account, for each temperature violation, the energy consumed for the conditioning, normalized by bias energy, temperature interval, and number of sensors outside range in that moment
04	Apply the formulas. We will have a value for each smart building and for each month, so could be compared the performance of different months of the same building for degradation analysis or could be compared performance of different buildings (in which case will be useful compare the same month)
Data Source	

Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**

Office Registry	Building Data	At starting up and then occasionally, when changes occur	Static - No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Detailed Energy Consumption	Energy Consumption (and internal temperatures)	Daily	From when they are available	Poste Italiane

LLUC-03

KPI N°10			
KPI-Name	Lighting Estimation	**KPI-ID**	PI_KPI10
KPI-Type	Technical		
Description	The KPI calculates the % of deviation between the actual and the estimated lighting consumption.		
Target Value	+/- 5%	**Threshold Value**	+/- 10%
Rounding	round off to 0% for values between 0.00 and 0.49 and to 1% for values above		
Unit	Kilowatt per hour (KWh)		
Formula	$$\mathbf{LE_i} = \frac{L_{e,i} - L_{a,i}}{L_{a,i}} * 100$$ $$\mathbf{LE_M} = \frac{1}{N}\sum_i LE_i$$ LE_i = Lighting Estimation Error % of building "I" L_e = Lighting consumption estimated of building "I" L_a = Lighting consumption actual of building "I" N = number of buildings utilized for the KPI calculation		
Calculating frequency	Weekly		
Calculation Methodology			
Step	**Description**		

01	Select the time range and the building (month)			
02	Calculate the estimated consumption considering the following information: 1. Total energy consumption in the rime range 2. Parameter on % of incidence of consumption form Heating and Cooling systems 3. Building open hours and shift			
03	Lighting consumption estimation will be compared to the real consumption (where available) and then will be exploited by buildings for which is no available.			
03	Calculate the real consumption value			
04	Apply the formula			
05	The formula will be applied for each one of the selected buildings, then arithmetic mean will be calculated from these selected values.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Office Registry	Building Data	At starting up and then occasionally when changes occur	No Temporal Range	Poste Italiane
Building Calendar	Building Data	Monthly	From 01/01/2018	Poste Italiane
Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2018	Poste Italiane
Detailed Energy Consumption DL_102	Energy Consumption	Monthly	TBD	Poste Italiane
Total Active Energy consumption	Energy Consumption	Monthly		Poste Italiane
System Registry	Building Systems	At starting up and then occasionally, when changes occur	No Temporal Range	Poste Italiane
Employees Number	Occupancy	Monthly	From 01/01/2018	Poste Italiane

Pilot 3b – ROM Advanced Energy Management System and Spatial (Multi-Scale) Predictive Models in the Smart City

KPI N°01			
KPI-Name	**Total Energy Savings (TES)** **[kWh / Y]**	**KPI-ID**	ROM_Kpi_R01
KPI-Type	Technical - Energy Savings		
Description	The analysis and the improved management of the meters data (historical and current) will produce a series of measures and interventions ("EVENTS") that should reduce the yearly total energy consumptions, such as dismission of un-useful meters, maintenance and interventions on buildings following some anomalies detection, contractual re-definition resulting from Platoon analysis, other measures impacting on behaviours. Component indicators are the Total Energy Savings in terms of Gas (TES-G) and in terms of Electricity (TES-E), that gives a better picture of the impact of Platoon services. TeS can be applied also to different reference or analysis period different from Year. This KPI calculates for example the difference between the energy consumption before and after reference EVENTS. It is always necessary to explicit the subset of buildings refereed to an instance calculation. This subset can range from n.1 meter/building to all meters/buildings		
Target Value	--% Target to be defined	**Threshold Value**	--%
Rounding	---% round off to 0% for values between 0.00 and 0.49 and to 1% for values above		
Unit	Kilowatt (KWh) / year		
Formula	$\mathbf{TES_Y = TES_G + TES_E}$ $TES_{AY} = \frac{TES_A}{\text{Ref.period}} = \frac{TES_E + TES_G}{\text{Ref. period}}$ TESY = Total Energy Saving (for one full year) TES_F = Forecasted value (calculated for 1 full year after the event, including future periods) TES_{AY} = Actual value normalized on 1 full year TES_A = Actual value (sum of the measured savings from the EVENT time to last data available, when total period is different from 1 year) EVENT time = the date of the intervention / action / event		
Calculating frequency	On demand … Monthly		
Calculation Methodology			
Step	**Description**		
01	Select the time range = Ref. Period (year) ; default is 1 (year) I.e 1 month = 1/12; 18 months = 18/12 Select/identify/Set the EVENT time, in order to verify which period covered by data is available after this EVENT time. Ref.Period is set to this period		

02	Select the specific building(s) and the perimeter of calculation: Total energy consumption for District / Area buildings or Total energy consumption for Specific building(s)			
03	Select the Energy typology: Electric (power meters) or Gas (gas meters or kWh derived from contatermie dataset) Or Both			
04	Calculate energy saving for selected typology/building(s) comparing consumption related to one full year before the EVENT (ECb) and consumptions after the EVENT (ECa) (normalized if necessary to one full year) : ECa – ECb = TES repeat for different energy typology if requested			
Data Source				
Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Building Calendar	Building Data	Monthly	TBD	
Total Electric Energy consumption	Energy Consumption	Monthly	TBD	ROM
Total Gas Energy consumption	Energy Consumption	Monthly	TBD	ROM
Detailed Energy Consumption	Energy Consumption (electric or gas)	daily	TBD	ROM
Temperature, Humidity	Weather	Monthly		External Services

KPI N°02			
KPI-Name	Saving Personnel Costs	**KPI-ID**	**ROM_04_Kpi_R02**
KPI-Type	Technical		
Description	The installation of a monitoring system shall reduce the costs for the personnel. This KPI is calculated from the difference of the saved personnel costs (per year) and the depreciation amount of the data monitoring system.		
Target Value	15%	**Threshold Value**	30%

Rounding	0% round off to 0% for values between 0.00 and 0.49 and to 1% for values above
Unit	% on Euro [per year]
Formula	$$SPC = \frac{(CS - CD)}{C_A} \times 100$$ CS= Personnel cost saving, based on the calculation of the avoided yearly worked days CD = depreciation amount of the data monitoring system in the same year C_A = Actual value of the personnel cost for the energy management (before Platoon implementation) Note: the calculation has to be limited to the personnel directly involved or impacted from the toolbox usage.
Calculating frequency	Monthly
Calculation Methodology	
Step	**Description**
01	
02	–
03	-
04	

Data Source

Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**

KPI N°0x

KPI-Name	Forecast Error	**KPI-ID**	
KPI-Type	Technical		
Description	The KPI calculates the % of deviation between the energy consumption forecast and the actual consumption in the building. This KPI checks how closely the predictive model adheres to reality - <u>Effectiveness</u>		
Target Value	5%	**Threshold Value**	10%
Rounding	0%		

	round off to 0% for values between 0.00 and 0.49 and to 1% for values above
Unit	Kilowatt (KW)
Formula	$$FE = \frac{(F_F - F_A)}{F_A} \times 100$$ FE= Forecast Error % F_F = Forecasted value F_A = Actual value
Calculating frequency	Monthly
Calculation Methodology	
Step	**Description**
01	Select the time range and the specific building and the perimeter of calculation: 3. Total energy consumption or 4. Total energy consumption for energy typology (electric meter or gas mater) in the building (or set of buildings)
02	Calculate the forecast considering: – Real data consumption – Temperature and Humidity (meteo conditions) – Number of Customers and Employees (when occupancy factors is available) – Building open hours and shift (hen occupancy factors is available) – Building Climate Zone, m^3
03	Calculate the Real consumption data (of the target month) taking into account: - Sum the daily energy consumption over the full month for the selected building or set of building
04	Apply the formula

Data Source

Data description	**Data source**	**Data collection frequency**	**Data collection time range**	**Data Owner**
Office Registry	Building Data	At starting up and then occasionally, when changes occur	No Temporal Range	ROM
Building Calendar	Building Data	Monthly	From 01/01/2016	ROM
Total Active Energy consumption	Energy Consumption	Monthly		ROM
Detailed Energy Consumption DL_102	Energy Consumption	Monthly	TBD	ROM
Detailed Energy Consumption	Energy Consumption	Daily	From 01/01/2016	ROM

Temperature, Humidity	Weather		From 01/01/2016	External Services
Customers Number	Occupancy	Monthly	From 01/01/2016	ROM if available
Employees Number	Occupancy	Monthly	From 01/01/2016	ROM if available

Pilot 3C Energy Efficiency and Predictive Maintenance in the Smart Tertiary Building Hub Grade

LLUC-3C-01

KPI N°1			
KPI-Name	Integration	**KPI-ID**	2
KPI-Type	Technical		
Description	Metric targeted at the validation of the fact that the tools of this pilot are able to work together. This includes: -Semantic pipeline: PLATOON data models mapping -Data Connectors with legacy databases, sensors, edge computing devices -IDS connector between Giroa and Tecnalia -IDS connector with Broker and Marketplace -Data Analytics Tools		
Target Value	1	**Threshold Value**	1
Rounding	Not applicable		
Unit	Binary 1 or 0		
Formula	If all tools to complete the pilot data analysis are able to interact and send data to each other then this KPI is 1. Otherwise it is 0.		
Calculating frequency	At each pipeline release		
Calculation Methodology			
Step	**Description**		
01-	based on unit tests the input-output functioning of each pipeline is validated.		
02	Test data is exchanged between the pilot analytics blocks		
Data Source			

Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Test data	Energy consumption/generation data, energy price data, meteo data and operational parameters.	Mins	2021-2022	Giroa

KPI N°2			
KPI-Name	Energy Bill reduction	**KPI-ID**	2
KPI-Type	Business		
Description	The KPI will evaluate the energy bill reduction achieved		
Target Value	20% reduction	**Threshold Value**	All improvement compared to current situation is already useful.
Rounding	first decimal		
Unit	% and euros		
Formula	(Current energy bill (euros)- New energy bill(euros))/ Current energy bill (euros)		
Calculating frequency	Once per day		
Calculation Methodology			
Step	**Description**		
01-	Calculate the energy generation and consumption forecast		
02	Calculate corrected energy price taking into account energy production excess and selling/buying poll pricre		
03	Optimise HVAC on/off		
04	Calculate HVAC energy consumption		
05	Calculate energy bill taking into account outputs from steps 2 and 4.		
Data Source			

Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Test data	Energy consumption/generation data, energy price data,	Mins	2021-2022	Giroa

	meteo data and operational parameters.			

KPI N°3				
KPI-Name	RES utilization ratio		**KPI-ID**	3
KPI-Type	Technical			
Description	The KPI will evaluate the RES usage versus overall energy consumption.			
Target Value	30% increase	**Threshold Value**	Full processing chain for a farm should be able to run on a standard server.	
Rounding	1st decimal			
Unit	Percentage			
Formula	RES production usage/ overall energy consumption			
Calculating frequency	Once per day			
Calculation Methodology				
Step	**Description**			
01-	Calculate the energy generation and consumption forecast			
02	Calculate corrected energy price taking into account energy production excess and selling/buying poll pricre			
03	Optimise HVAC on/off			
04	Calculate HVAC energy consumption			
05	Calculate RES usage taking into account outputs from steps 1 and 4.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Test data	Energy consumption/generation data, energy price data, meteo data and operational parameters.	Mins	2021-2022	Giroa

LLUC-3C-02

<table>
<tr><td colspan="5">KPI N°1</td></tr>
<tr><td>KPI-Name</td><td colspan="2">Health Monitoring</td><td>KPI-ID</td><td>KPI-01</td></tr>
<tr><td>Description</td><td colspan="4">Monitoring the health status of each asset, using the PML (Process Mastery Level) indicator, in a range from 0 (Failure Status) to 1 (Optimal Status).</td></tr>
<tr><td>Target Value</td><td>100 %</td><td colspan="2">Threshold Value</td><td>0 – 100%</td></tr>
<tr><td>Unit</td><td colspan="4">Percentage indicator, set points, etc.</td></tr>
<tr><td>Formula</td><td colspan="4">Each defined failure mode will have specific Digital twin based on machine learning algorithms. From those models, the real time information grouped by time slots (for example, 8 hours) will be evaluated against the Digital twin.
Statistics for the digital twin testing:
R^2
MAE

The correlation will be evaluated in a range from 0 to 1 as a FTL

Always depending on the availability of signals, an attempt will be made to extract information about the following PMLs:
Energy Variator
Starter
Phase imbalance
Power Supply
Communications
Flow Meter
Temp Out of range
Evaporator Return
Temp Increase
Power consumption increase
Evaporator Outlet Temp</td></tr>
<tr><td>Calculating frequency</td><td colspan="4">Depending on the asset. From 4 to 24 hours</td></tr>
<tr><td colspan="5">Calculation Methodology</td></tr>
<tr><td>Step</td><td colspan="4">Description</td></tr>
<tr><td>01-</td><td colspan="4">Define the PML Formula for each asset</td></tr>
<tr><td>02-</td><td colspan="4">Monitoring the health status according to the values of the variables and its associated PML Value.</td></tr>
<tr><td colspan="5">Data Source</td></tr>
<tr><td>Data description</td><td>Data source</td><td>Data collection frequency</td><td>Data collection time range</td><td>Data Owner</td></tr>
<tr><td>Signals used as input for the models</td><td>Consolidated Data Base with PLC and SCADA data</td><td>Defined default time (every 15 min aprox.)</td><td></td><td>GIROA</td></tr>
</table>

KPI N°2				
KPI-Name	Availability		**KPI-ID**	KPI-02
Description	Availability of the asset over a period of time. Availability takes into account Availability Loss, which includes any events that stop planned production for an appreciable length of time (usually several minutes; long enough for an operator to log a reason). Used for OEE calculation.			
Unit	%			
Formula				
Calculating frequency	Daily			
Calculation Methodology				
Step	**Description**			
01-	Register events of unplanned stops.			
02	Calculate the availability for a determined period of time by using the above formula.			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Signals used as input for the models	Consolidated Data Base with PLC and SCADA data	Defined default time (every 15 min aprox.)		GIROA

KPI N°3			
KPI-Name	Mean Time Between Failures	**KPI-ID**	KPI-03
Description	Mean time between failures (MTBF) describes the expected time between two failures for a repairable system		
Unit	Hours		
Formula	MTBF = (Total Working Time – Total Breakdown Time) / Number of Breakdowns MTBF = Total Operational time / Number of Breakdowns		
Calculating frequency	Daily		
Calculation Methodology			
Step	**Description**		

01-	Acquire running operational time			
02-	Determine Number of breakdowns. Apply filters as needed to exclude micro stops, mini stops, or other criteria's			
03-	Apply formula			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Signals used as input for the models	Consolidated Data Base with PLC and SCADA data	Defined default time (every 15 min aprox.)		GIROA

KPI N°4				
KPI-Name	Maintenance Costs		**KPI-ID**	KPI-04
KPI-Type	Business			
Description	The maintenance cost of an asset is the sum of the costs of the work orders that have been carried out on that asset. It is important to indicate that maintenance costs may be higher in some assets that use predictive maintenance. Therefore, the goal should be achieving the lowest possible cost in the set of assets.			
Target Value	Not applicable	**Threshold Value**	Not applicable	
Rounding	Not applicable			
Unit	Euros			
Formula	Sum of the maintenance costs of the equipment selected for the use case.			
Calculating frequency	Daily			
Calculation Methodology				
Step	**Description**			
01-	Acquire necessary data from integration with Prisma (CMMS) Total cost associated to Work Orders related to the equipment			
02	Create total cost KPI associated to the equipment			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner

Test data	Operational parameters.	Mins	2021-2022	Giroa
Maintenance log data	Prisma	Daily	Daily	Giroa

KPI N°5				
KPI-Name	Integration		**KPI-ID**	KPI-05
KPI-Type	Technical			
Description	Metric targeted at the validation of the fact that the tools of this pilot are able to work together. This includes: -Semantic pipeline: PLATOON data models mapping -Data Connectors with legacy databases, sensors, edge computing devices -IDS connector between Sisteplant and Tecnalia -IDS connector with Broker and Marketplace -Data Analytics Tools			
Target Value	1	**Threshold Value**	1	
Rounding	Not applicable			
Unit	Binary 1 or 0			
Formula	If all tools to complete the pilot data analysis are able to interact and send data to each other then this KPI is 1. Otherwise it is 0.			
Calculating frequency	At each pipeline release			
Calculation Methodology				
Step	**Description**			
01-	based on unit tests the input-output functioning of each pipeline is validated.			
02	Test data is exchanged between the pilot analytics blocks			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Test data	Energy consumption/generation data, energy price data, meteo data and operational parameters.	Mins	2021-2022	Giroa

Pilot 4a Energy Management of Microgrids

LLUC P-4a-01

KPI N°1			
KPI-Name	Energy availability	**KPI-ID**	KPI-1
KPI-Type	Technical (specific to the pilot use case)		
Description	Optimal energy consumption (increase in energy availability) – Optimization for renewable electricity generation Smart storage/generation		
Target Value	Example: amount of daily load covered by renewable generation – Target value:100%	**Threshold Value**	The value used to assess the effectiveness/efficiency performance of the monitored process: 90%
Rounding	the criteria for rounding the calculated values (Example : For % calculation, round off to 0% for values between 0.00 and 0.49 and to 1% for values above)		
Unit	%		
Formula	$$\mathrm{KPI}_{01}(\%) = \frac{\sum_{t=1}^{N} P_{PV,t}}{\max\left(\sum_{t=1}^{N} P_{PV,t}, \sum_{t=1}^{N} P_{load,t}\right)} \cdot 100$$		
Calculating frequency	Daily		
Calculation Methodology			
Step	**Description**		
01-	Daily measurements of load consumption, renewable energy generation and battery		
02			

Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy consumption	mySQL db	10 min	daily	PDM
Energy production	mySQL db	10 min	daily	PDM

KPI N°2			
KPI-Name	Cost	**KPI-ID**	KPI-2
KPI-Type	Technical (specific to the pilot use case)		

Description	Reduction of maintenance effort and costs (optimization for renewable electricity generation)			
Target Value	Example: maintenance cost	**Threshold Value**	10%	
Rounding	the criteria for rounding the calculated values (Example : For % calculation, round off to 0% for values between 0.00 and 0.49 and to 1% for values above)			
Unit	%			
Formula	$\mathrm{KPI}_{02}(\%) = \frac{\sum_{t=1}^{N}\left(P_{load,t} - P_{PV,t}\right)}{\sum_{t=1}^{N} P_{load,t}} \cdot 100$			
Calculating frequency	daily, montly			
Calculation Methodology				
Step	**Description**			
01-	Daily measurements of load consumption, renewable energy generation and battery			
02				
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Failure rate	mySQL db	10 min	daily	PDM
Maintenance activity	mySQL db	10 min	daily	PDM

KPI N°3			
KPI-Name	Forecast accuracy	**KPI-ID**	KPI-3
KPI-Type	Technical (specific to the pilot use case)		
Description	Reduced forecasting errors (generation and load forecast)		
Target Value	Example: forecating error – Target value:0%	**Threshold Value**	The value used to assess the effectiveness/efficiency performance of the monitored process: 20%
Rounding	None		
Unit	%		

Formula	Standard forecating error indicators, such as nRMSE, WMAE, EMAE, OMAE $$\mathrm{KPI}_{03}(\%) = \mathrm{nRMSE} = \frac{1}{\max\,(P_{m,t})}\sqrt{\frac{\sum_{t=1}^{N}\left(P_{f,t} - P_{m,t}\right)^2}{N}} \cdot 100$$ $$\mathrm{EMAE} = \frac{\sum_{t=1}^{N}\left|P_{f,t} - P_{m,t}\right|}{\sum_{t=1}^{N}\max\left(P_{f,t}, P_{m,t}\right)} \cdot 100$$ [S. Leva, M. Mussetta, A. Nespoli and E. Ogliari, "PV power forecasting improvement by means of a selective ensemble approach," 2019 IEEE Milan PowerTech, 2019, pp. 1-5, doi: 10.1109/PTC.2019.8810921.]			
Calculating frequency	Daily, monthly, yearly			
Calculation Methodology				
Step	**Description**			
01-	Daily measurements of renewable energy generation			
02	Comparison with related forecasting and error measurement			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Energy consumption	mySQL db	10 min	daily	PDM
Energy production	mySQL db	10 min	daily	PDM
Production forecast	mySQL db	10 min	daily	PDM
Solar nowcast	mySQL db / edge node?	10 min	daily	PDM
Load forecast	mySQL db	10 min	daily	PDM

KPI N°4			
KPI-Name	Realtime	**KPI-ID**	KPI-4
KPI-Type	Technical (specific to the pilot use case)		
Description	Ability to monitoring/analyze/optimize data and the system at real time rate (EMS with real-time processing)		

Target Value	100%	**Threshold Value**	80%
Rounding	the criteria for rounding the calculated values (Example : For % calculation, round off to 0% for values between 0.00 and 0.49 and to 1% for values above)		
Unit	%		
Formula	$$\mathrm{KPI}_{04}(\%) = \frac{\sum_{t=1}^{N}\left(P_{m,t} - P_{f,t}\right)^2 - \sum_{t=1}^{N}\left(P_{m,t} - P_{n,t}\right)^2}{\sum_{t=1}^{N}\left(P_{m,t} - P_{f,t}\right)^2} \cdot 100$$		
Calculating frequency	Daily		
Calculation Methodology			
Step	**Description**		
01-	Daily measurements of renewable energy generation		
02	Comparison with related forecasting and error measurement		
Data Source			

Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Nowcast	mySQL db	10 min	daily	PDM
EMS schedule	mySQL db	10 min	daily	PDM

PLATOON Common Components

KPI N°1				
KPI-Name	IDS Metadata Registry (Boker/Appstore)Integration		**KPI-ID**	1
KPI-Type	Technical			
Description	Metric targeted at the validation of the fact that the IDS Metadata Registry (Broker/Appstore) is able to work together with IDS connectors and Data Analytics Tools Dockers and Marketplace.			
Target Value	1		**Threshold Value**	1
Rounding	Not applicable			
Unit	Binary 1 or 0			
Formula	If all tools to complete the pilot data analysis are able to interact and send data to each other then this KPI is 1. Otherwise it is 0.			
Calculating frequency	At each pipeline release			
Calculation Methodology				
Step	**Description**			
01-	based on tests the input-output functioning of each pipeline is validated.			
02	Test data is exchanged between the pilot analytics blocks			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Each pilot will have its own dataset connected to IDS connectors.				

KPI N°2			
KPI-Name	DAPS Integration	**KPI-ID**	3
KPI-Type	Technical		
Description	Metric targeted at the validation of the fact that the DAPS provided by Fraunhofer AISEC (not developed in PLATOON) is able to work together with PLATOON IDS connectors, IDS Metadata Registry and IDS Vocabulary Provider.		
Target Value	1	**Threshold Value**	1

Rounding	Not applicable			
Unit	Binary 1 or 0			
Formula	If all tools to complete the pilot data analysis are able to interact and send data to each other then this KPI is 1. Otherwise it is 0.			
Calculating frequency	At each pipeline release			
Calculation Methodology				
Step	**Description**			
01-	based on tests the input-output functioning of each pipeline is validated.			
02	Test data is exchanged between the pilot analytics blocks			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Each pilot will have its own dataset connected to IDS connectors.				

KPI N°3			
KPI-Name	Clearing House Integration	**KPI-ID**	4
KPI-Type	Technical		
Description	Metric targeted at the validation of the fact that the Clearing House provided by Fraunhofer (not developed in PLATOON) is able to work together with PLATOON IDS connectors, IDS Metadata registry, DAPS and Marketplace.		
Target Value	1	**Threshold Value**	1
Rounding	Not applicable		
Unit	Binary 1 or 0		
Formula	If all tools to complete the pilot data analysis are able to interact and send data to each other then this KPI is 1. Otherwise it is 0.		
Calculating frequency	At each pipeline release		
Calculation Methodology			
Step	**Description**		
01-	based on tests the input-output functioning of each pipeline is validated.		

02	Test data is exchanged between the pilot analytics blocks			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Each pilot will have its own dataset connected to IDS connectors.				

KPI N°4			
KPI-Name	PLATOON Marketplace GUI Integration	**KPI-ID**	5
KPI-Type	Technical		
Description	Metric targeted at the validation of the PLATOON Marketplace is able to work together with PLATOON IDS Metadata Registry, DAPs and Clearning House.		
Target Value	1	**Threshold Value**	1
Rounding	Not applicable		
Unit	Binary 1 or 0		
Formula	If all tools to complete the pilot data analysis are able to interact and send data to each other then this KPI is 1. Otherwise it is 0.		
Calculating frequency	At each pipeline release		
Calculation Methodology			
Step	**Description**		
01-	based on tests the input-output functioning of each pipeline is validated.		
02	Test data is exchanged between the pilot analytics blocks		

Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Each pilot will have its own dataset connected to IDS connectors.				

KPI N°5

KPI-Name	IDS Vocabulary Provider Integration		**KPI-ID**	2
KPI-Type	Technical			
Description	Metric targeted at the validation of the fact that the IDS Vocabulary Provider is able to work together with PLATOON datamodels repository, IDS connectors and DAPs.			
Target Value	1	**Threshold Value**	1	
Rounding	Not applicable			
Unit	Binary 1 or 0			
Formula	If all tools to complete the pilot data analysis are able to interact and send data to each other then this KPI is 1. Otherwise it is 0.			
Calculating frequency	At each pipeline release			
Calculation Methodology				
Step	**Description**			
01-	based on tests the input-output functioning of each pipeline is validated.			
02	Test data is exchanged between the pilot analytics blocks			
Data Source				
Data description	**Data source**	Data collection frequency	Data collection time range	Data Owner
Each pilot will have its own dataset connected to IDS connectors.				

References

[i] https://www.servisinfo.com/cena-struje

[ii] Messinis, G. M., Rigas, A. E., & Hatziargyriou, N. D. (2019). A hybrid method for non-technical loss detection in smart distribution grids. *IEEE Transactions on Smart Grid*, *10*(6), 6080-6091

[iii] Cochran, W.G. (1977) Sampling Techniques. 3rd Edition, John Wiley & Sons, New York.