6G Vertical trials for Sustainability

D5.1
6G Trials Plan and Validation/Evaluation Framework

Date of delivery:	31/10/2025	Version:	1.0
Project reference:	101192633	Call:	HORIZON-JU-SNS-2024-STREAM-D-01-01
Project Start date:	01/01/2025	Duration:	36 months

Co-funded by the European Union

6G-VERSUS project has received funding from the Smart Networks and Services Joint Undertaking (SNS JU) under the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101192633.

Document properties:

Number:	D5.1
Title:	6G Trials Plan and Validation/Evaluation Framework
Editor(s):	Haoxin Sun (KEYS); Javier Rivas (KEYS and UMA)
Authors:	Luís Cordeiro (ONE); Jorge Proença (ONE); Duarte Fevereiro (ONE); João Fernandes (ONE); Antti Pauane (UOULU); Hamid Malik (UOULU); Sanna Tuomela (UOULU); Haoxin Sun (KEYS); Issam Maaz (KEYS); Germán Corrales Madueño (KEYS); Maryam Irekeola (KUL); Shanaka Perera (UBRAD); George Makropoulos (NCSRD); Harilaos Koumaras (NCSRD); Spyridon Georgoulas (NCSRD); Elina Theodoropoulou (OTE); Fofy Setaki (OTE); George Lyberopoulos (OTE); Aimilios Leftheriotis (ICCS); Dimitrios Fragkos (ICCS); George Lentaris (ICCS); Konstantinos V. Katsaros (ICCS); Nikos Mitro (ICCS); Spyros Athanasiadis (ICCS); Thanasis Douklias (ICCS); Xenofon Vasilakos (ICCS); Vagelis Anagnostopoulos (INF); Evgenia Vergi (INF); Iraklis Spyrou (INF); George Koumaras (INF); Delia Rico (UMA); Francisco Luque (UMA); Javier Rivas (KEYS and UMA); Pablo Herrera (UMA); Pedro Merino (UMA); Álvaro Curto Merino (TID); Arturo José Torrealba Ferrer (TID); Pablo Iglesias Sanuy (TID); Sofiane Messaoudi (EUR); Ralitsa Rumenova (EE); Verzhinia Ivanova (EE); Georgi Manev (EE); Irina Ciornei (UCY); Kyriakos Karlettis (UCY); Panagiotis Chrysanthou (UCY); Afonso Castanheta (ITAV); Daniel Corujo (ITAV); David Santos (ITAV); Jorge Carapinha (ITAV); Rui Silva (ITAV); Tiago Barros (ITAV); João Gonçalves (IPN); Rafael Gonçalves (IPN); Sérgio Figueiredo (IPN); Davide Montagno Bozzone (HPE); Nicola di Pietro (HPE); Veronique Capdevielle (TSGF); Doan Khuc (TSGF); Ángel Cañete (CET); Sabina Torralba (HIDRA); Antonio Jesús Herrera (HIDRA)
Delivery Date:	31/10/2025
Dissemination:	PU
Status:	Final
Version:	1.0
File Name:	6G-VERSUS_D5.1_v1.0

Revision History

Revision	Date	Issued by	Description
0.1	14/05/2025	KEYS	Document with ToC
0.2	09/09/2025	ONE	Transposed KPI and Test case definitions from Blueprints
0.3	09/10/2025	KEYS	Document with all contributions from partners
0.4	10/10/2025	ONE	WP5 leader document review
0.5	23/10/2025	KEYS	Document updated addressing comments received by quality review partners
0.6	27/10/2025	ICCS	TM and PC review
1.0	31/10/2025	OULU	Final version

Abstract

This deliverable presents the initial trial plan and validation/evaluation framework for the 6G-VERSUS project. It describes the methodologies and processes guiding the preparation, execution, and reporting of large-scale trials across six European clusters, and it outlines the first version of the frameworks for Key Performance Indicators (KPIs) and Key Value Indicators (KVIs). As this document reflects the early stages of the project, both the trial plans and the defined indicators should be regarded as preliminary. They are expected to evolve as knowledge is gained, as trials progress, and as insights from stakeholders and technical developments are incorporated.

Keywords

6G-VERSUS Trial; Key Performance and Value Indicators (KPIs/KVIs); Validation Framework; Evaluation Methodology; Test Cases

Disclaimer

Executive Summary

This deliverable (D5.1) presents the first iteration of the 6G-VERSUS trial planning and validation/evaluation framework. As part of the project's initial phase, this document should be regarded as a baseline that will undergo refinement as the project advances. This is the first deliverable reporting on trial execution and evaluation. Deliverables D5.2 and D5.3 will update these activities at the end of the first and second phases of the trials, reporting trial results and also providing feedback to the technical WPs (3,4) and feeding WP6 activities with trial results for the impact assessment activities.

The trial planning section outlines the general methodology and preparation steps for the Bulgarian, Finnish, French, Greek, Portuguese, and Spanish clusters. Guided by these introduced principles, each cluster has identified use cases targeting sustainability-driven verticals such as energy systems, agriculture, water management, and smart cities. This section also provides a first version of each cluster trial plan, including a timeline with the main activities, a description of the trials and their main goals, and a list of stakeholders relevant to the cluster's trials.

This deliverable also introduces the Key Value Indicator (KVI) validation framework, which is designed to assess the broader impacts generated beyond purely technical performance of the 6G-VERSUS project. By considering environmental, societal, and economic dimensions, this framework enables a more holistic evaluation of technology design, deployment, and operation in 6G-VERSUS, in alignment with global and European sustainability agendas.

In parallel, another major focus is on the Key Performance Indicator (KPI) validation framework, enabling consistent measurement of technical performance across heterogeneous testbeds. Building on relevant projects, the framework is demonstrated to define performance metrics, measurement methodologies, and supporting tools. For each cluster, the relevant KPIs of interest are summarised alongside associated measurement tools, aiming at providing comprehensive evaluations. Additionally, initial test case plans are also shown outlined for each cluster, which lays a solid foundation for the upcoming trial executions in 6G-VERSUS.

In summary, this deliverable establishes a valuable reference point for preparing and coordinating the project's trial activities, ensuring coherence with project objectives, testbed developments, stakeholder engagement, and evaluation needs. However, the trial plans, KPIs, and KVIs presented here are initial and will be adapted, refined, or extended as project knowledge increases and as feedback from early trial executions and stakeholder engagements becomes available. This iterative and adaptive approach is central to ensuring that 6G-VERSUS delivers both technical excellence and societal relevance.

Table of Contents

List of Tables

List of Figures

Acronyms and abbreviations

Term	Description
5QI	5G Quality of Service Identifier
AGV	Automated Guided Vehicles
AI	Artificial Intelligence
AI-App	AI-Assisted App
ALB	Altice Labs, S.A.
AMF	Access And Mobility Management Function
APA	Administração Do Porto De Aveiro, S.A.
API	Application Programming Interface
APN	Access Point Name
AR	Augmented Reality
ASAJA	Asociación Agraria Jóvenes Agricultores
B5G	Beyond 5G
BESS	Battery Energy Storage System
BTS	Base Transceiver Station
BWP	Bandwidth Part
BYOL	Bring-Your-Own-License
CAMARA	Common Api Framework by GSMA
CAPEX	Capital Expenditures
CAPIF	Common Api Framework
CES	Centro De Estudos Sociais
CET	Cetaqua
CI/CD	Continuous Integration And Continuous Development/Delivery
CN	Core Network
COTS	Commercial-Off-The-Shelf
CPU	Central Processing Unit
CSIC	Agencia Estatal Consejo Superior De Investigaciones Cientificas
CV-RMSE	Coefficient Of Variation RMSE
C-HIL	Control Hardware In The Loop
DCS	Distributed Control System
DER	Distributed Energy Resource
DL	Downlink
DMP	Data Management Plan
DNN	Data Network Name
DRX	Discontinuous Reception
DSS	Decision System Support
DT	Digital Twin
E2E	End-to-End
eCPRI	Enhanced Common Public Radio Interface
EE	Entra Energy
EEaaS	Energy Efficiency As A Service
eMBB	Enhanced Mobile Broadband
eNB	Evolved NodeB (4G base station)
EPC	Evolved Packet Core
EV	Electric Vehicle
FFR	Fast Frequency Support
FPS	Frame Per Second
FRT	Fault Ride Through
FTP	File Transfer Protocol
FW	Firmware

gNB	Next Generation Node B (5G base station)
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GPU	Graphical Processing Unit
GREEN	Greencityzen
GUI	Graphical User Interface
HD	High Definition
HIDRA	Hidralia
HPE	Hewlett-Packard Italiana S.R.L.
HTTP	Hypertext Transfer Protocol
HW	Hardware
ICCS	Erevnitiko Panepistimiako Institouto Systimaton Epikoinonion Kai Ypologiston
ICMP	Internet Control Message Protocol
INC	Incites Consulting Sa
INF	Infolysis P.C.
INUS	Intelligent Uxv Surveillance
IoT	Internet of Things
IPN	Instituto Pedro Nunes
ITAV	Instituto De Telecomunicacoes
JCAS	Joint Communication And Sensing
JSIO	JSIO - IoT Hassle Free
KEYS	Keysight Technologies
KPI	Key Performance Indicator
KV	Key Value
KVE	Key Value Enabler
KVI	Key Value Indicator
LAN	Local Access Network
MAE	Mean Absolute Error
MASE	Mean Absolute Scaled Error
MDI	Media Delivery Index
MEC	Multi-access Edge Computing
MIMO	Multiple Input Multiple Output
ML	Machine Learning
ML-Ops	Machine Learning Operations
mMTC	Massive Machine Type Communications
MOS	Mean Opinion Score
MTU	Maximum Transmission Unit
MV	Medium Voltage
NCSRD	National Center For Scientific Research "Demokritos"
NDM	Nemo Diagnostic Module
NEF	Network Exposure Function
NFVI	Network Functions Virtualization Infrastructure
NMS	Network Management System
NOC	Network Operations Center
NR	New Radio
NTN	Non-Terrestrial Networks
NWDAF	Network Data Analytics Function
N-App	Network App
OAI	OpenAirInterface
ONE	One Source Consultoria Informatica Lda
OPEX	Operating Expense
OSS	Operations Support System

OTE	Organismos Tilepikoinonion Tis Ellados Ote Ae
OVA	Open Virtual Appliance
PDCP	Packet Data Convergence Protocol
PDU	Protocol Data Unit
PMU	Phasor Measurement Unit
PRB	Physical Resource Block
PV	Photovoltaic
QoD	Quality On Demand
O-RAN	Open RAN
PLMN	Public Land Mobile Network
QoS	Quality of Service
RAN	Radio Access Network
RAPL	Running Average Power Limit
RAT	Radio Access Technology
RC	Regional Controller
RedCap	Reduced Capability
REI	Relative Efficiency Index
RES	Renewable Energy Source
RF	Radio Frequency
RGB	Red, Green, and Blue
RIC	RAN Intelligent Controller
RMSE	Root Mean Square Error
RoCoF	Rate of Change of Frequency
ROI	Return On Investment
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RSSI	Received Signal Strength Indicator
RTT	Round Trip Time
RU	Radio Unit
SA	Standalone
SAR	Search And Rescue
SCBG	Software Company
SDG	Sustainable Development Goal
SG	Sustainable Goal
SINR	Signal To Interference Plus Noise Ratio
SISO	Single Input Single Output
SME	Small and Medium-sized Enterprises
SMF	Session Management Function
SMS	Short Message Service
SoC	State-of-Charge
SoH	State-of-Health
SRIA	Strategic Research And Innovation Agenda
SRS	Sounding Reference Signal
SW	Software
TAC	Test Automation Cloud
TC	Test Case
TID	Telefónica Innovación Digital S.L.
TNLCM	Trial Network Life Cycle Manager
TWAMP	Two-Way Active Measurement Protocol
UAV	Unmanned Aerial Vehicle
UBRAD	University Of Bradford
UC	Use Case

UCY	University Of Cyprus
UE	User Equipment
UI	User Interface
UL	Uplink
UMA	Universidad De Málaga
UOULU	Oulun Yliopisto
UPF	User Plane Function
URLLC	Ultra-Reliable Low Latency Communications
USRP	Universal Software Radio Peripheral
VM	Virtual Machine
VNF	Virtual Network Function
VPS	Virtual Private Server
V-App	Vertical App
WAC	Wide Area Control
WAM-C	Wide Area Monitor And Control
YOY	Year-On-Year

1 Introduction

This document provides a high-level description of both cluster trial planning and evaluation frameworks for Key Value Indicator (KVI) and Key Performance Indicator (KPI). This has been conceived around an iterative approach that seeks incremental refinement as the project progresses. As a result, some of the initial planning efforts reported in this deliverable are linked to ongoing activities that might be refined and updated during the project and in future deliverables. Specifically, some trial timelines may have dependencies on external factors and stakeholders, potentially requiring adaptation. Similarly, the identification and evaluation of both KVIs and KPIs can be refined as the project evolves, utilising feedback from initial trials to improve methodologies and data processing processes, ultimately enhancing the impact of results. As a result, KVIs, KPIs, test cases, or evaluation methods could be added or removed to a certain extent if they contribute to better fulfilling the project objectives.

1.1 Scope and Intended Audience

This deliverable establishes the initial trial planning and validation/evaluation framework for the 6G-VERSUS project, which will lead WP5 and WP6 activities. It outlines the methodologies, processes, indicators, and test plans that will guide the trial execution across the six clusters. Thus, it provides a baseline reference for preparing and coordinating the project’s trial activities, ensuring consistency across project objectives, testbed developments, as well as stakeholder engagement and evaluation requirements.

The intended audience for this deliverable includes project partners contributing to WP5 activities, technical teams in WP3 and WP4 preparing testbeds and enablers for trials, cluster leads responsible for execution, and stakeholders engaged in validation and assessment. Besides, external reviewers and interested third parties may also use this document as a reference to understand the project’s initial approach to trial planning and KVI/KPI-based evaluation.

1.2 Relation to Milestones and Other Deliverables

This deliverable consolidates inputs from MS4 (Draft of the vertical use cases, requirements and KVIs/KPIs), MS5 (Draft of the reviewed architecture), and MS6 (6G-VERSUS testbeds description and upgrades & extensions roadmap), providing the foundation for both the methodological framework and the planning of the trials. Therefore, it ensures that validation goals are fully aligned with the project’s Use Cases (UCs), architecture evolution, and testbed capabilities. Furthermore, two milestones MS11 (Draft version of the validation and evaluation framework (KPIs), M10) and MS14 (Draft version of the validation and evaluation framework (KVIs), M10) are verified with the initial validation framework described in D5.1 chapters 3 and 4.

D5.1 is also closely interlinked with the technical WPs responsible for preparing the trials: WP3 on testbeds (with outcomes reported in D3.1, M12) and WP4 on enablers and X-Apps (with outcomes reported in D4.1, M16, and D4.2, M32). This alignment ensures that the trial plans take into account the testbed readiness, the integration of enabling technologies, and the availability of cluster-specific applications.

Furthermore, D5.1 serves as the reference point for the subsequent deliverables: D5.2 “6G trials validation and evaluation – Initial” (M18), which will report on the first trial executions, and D5.3 “6G trials validation and evaluation – Final” (M36), which will consolidate the evaluation of the complete set of trials. It also complements D2.1 (M10) on use cases and requirements as well as D2.2 (M20) on refined architecture, thus ensuring coherence between trial planning, requirements validation, and technical refinement.

1.3 Document Structure

The structure of this deliverable is designed to provide a coherent overview of the trial planning activities and the validation frameworks that will guide the 6G-VERSUS project. The document follows a progressive flow from general context to detailed cluster-specific descriptions and evaluation methodologies.

- Section 2 – Trial Planning and Preparation introduces the methodology applied to plan and prepare the 6G-VERSUS trials. It describes the four main phases (planning, preparation, execution, and reporting) and then details the preliminary plans for each of the six national clusters (Bulgaria, Finland, France, Greece, Portugal, Spain). For each cluster, the section outlines technical and business objectives, stakeholder engagement, and risk mitigation.
- Section 3 – KVI Validation Framework presents the first version of the framework for identifying and evaluating KVIs. This section explains the methodology for linking project outcomes with societal, environmental, and economic values in line with EU sustainability goals. It provides guidance for cluster partners on how to design and validate KVIs during trials.
  Section 4 – KPI Validation Framework describes the methodology for defining, measuring, and validating KPIs. It outlines references to existing frameworks from previous projects, defines the terminology for tests and experiments, and introduces recommended practices for measurement and monitoring. The section also includes initial cluster-specific KPI definitions with extended context, to enable effective trial evaluation and fair benchmarking with other projects such as 6G-SANDBOX, 6G-PATH, and 5G!DRONES.
- Section 5 – Conclusions summarises the main contributions of this deliverable, highlights its preliminary nature, and sets expectations for future updates. It also provides an outlook on how the frameworks and plans will be refined during subsequent phases of the project.

This structure ensures that readers can understand the practical implementation aspects of the trials (Section 2), and both strategic frameworks for the KVIs and KPIs (Sections 3 and 4), while keeping in mind that this deliverable is an initial reference point subject to iterative refinement throughout the project lifecycle.

2 Trial Planning and Preparation

This chapter details the planning and preparation activities carried out for the 6G-VERSUS trials. The objective is to provide a high-level overview of the actions taken to define the scope, objectives, methodology, and logistics necessary for the successful execution of each cluster's use case. Trial preparation is a critical step in ensuring alignment between technical capabilities, stakeholder expectations, and the validation goals set by the project. Additional details on each cluster's use cases, requirements, and KPI/KVI definitions are reported in D2.1 (M10) [6GV-D21]. Trial planning is also closely aligned with technical WPs (WP3, WP4). Alignment with WP3 is essential to ensure that the clusters' trial plans align with the work being done in the testbeds' improvements and extensions, to ensure their preparedness for the trials' execution dates, as well as X-Apps development in WP4. Details on the testbeds' preparation will be reported in D3.1 (M12) [6GV-D31], and the report on the initial X-Apps will be in D4.1 (M16) [6GV-D41].

The following sub-sections provide a structured overview of the current status and approach for each cluster in terms of trial planning and preparation. First, section 2.1 provides a detailed description of the methodology followed to create the trial plans. Section 2.2 lists the potential risks that may affect the successful lifecycle of the trials and pilots. The preliminary status of each cluster planning is presented across sections 2.3 to 2.8.

2.1 Methodology

Trials in 6G-VERSUS are organised into two main phases: phase 1 (M10-M18) and phase 2 (M22-M32). Each cluster will run multiple trials within each phase, following a progressive path of validation and scaling. Clear objectives have been defined in each phase to guide a progressive validation of methodologies and impact assessment. The key objectives of each phase are presented below:

Phase 1 — Methodology validation and first trial results (M10–M18).

The main goal of this phase is to reduce the risk of the overall project execution by validating end-to-end trial methodologies and procedures and stress-testing their planning, while providing training for all personnel involved in trial execution. It also includes the first trial phase and its results. Data-collection instruments are validated (such as questionnaires, interviews and focus-group guides), as well as the KPI/KVI analysis workflows. Clusters run "test-drive" demonstrations of the scenarios planned for the second phase, using smaller, lower-stakes events with stakeholders who can be engaged more than once, ensuring rapid iteration and learning before scaling. From the beginning of the first trials up to the end of this phase, each cluster will provide feedback to the technical WPs (3 and 4) so they can make any necessary adjustments to the testbeds and V/N/AI-Apps for the second phase.

Phase 2 — Scaling and consolidation (M22–M32).

The second phase focuses on scaling up the trials by increasing stakeholder attendance numbers and their relevance, the number of devices deployed, and scenario complexity by integrating a broader set of 5G/Beyond 5G (B5G) technologies. Before this phase, each cluster trial plan will be updated to address challenges observed in the first phase trials and make any necessary adjustments, enabling higher-impact demonstrations with well-prepared stakeholders and robust, comparable data and results that will feed WP5 and WP6 deliverables. Insights from the first phase will be used to revise KPI measurements and the analysis pipelines for KVI evaluation, including questionnaires, interviews and focus groups.

In each phase, the methodology adopted for the trials is structured around four main steps: **planning**, **preparation**, **execution** and **reporting**. These serve as a framework for the organisation, implementation, and evaluation of each use case trial across the participating clusters. The following subsections outline the primary activities and principles that govern each step, as outlined in the project guidelines.

Figure 1 provides a reference timeline for the trials, enabling clusters to run and evaluate them in line with project milestones and the reporting schedule, specifically for WP5 and WP6. Based on this, each cluster will define its individual timeline tailored to its use case and internal deadlines.

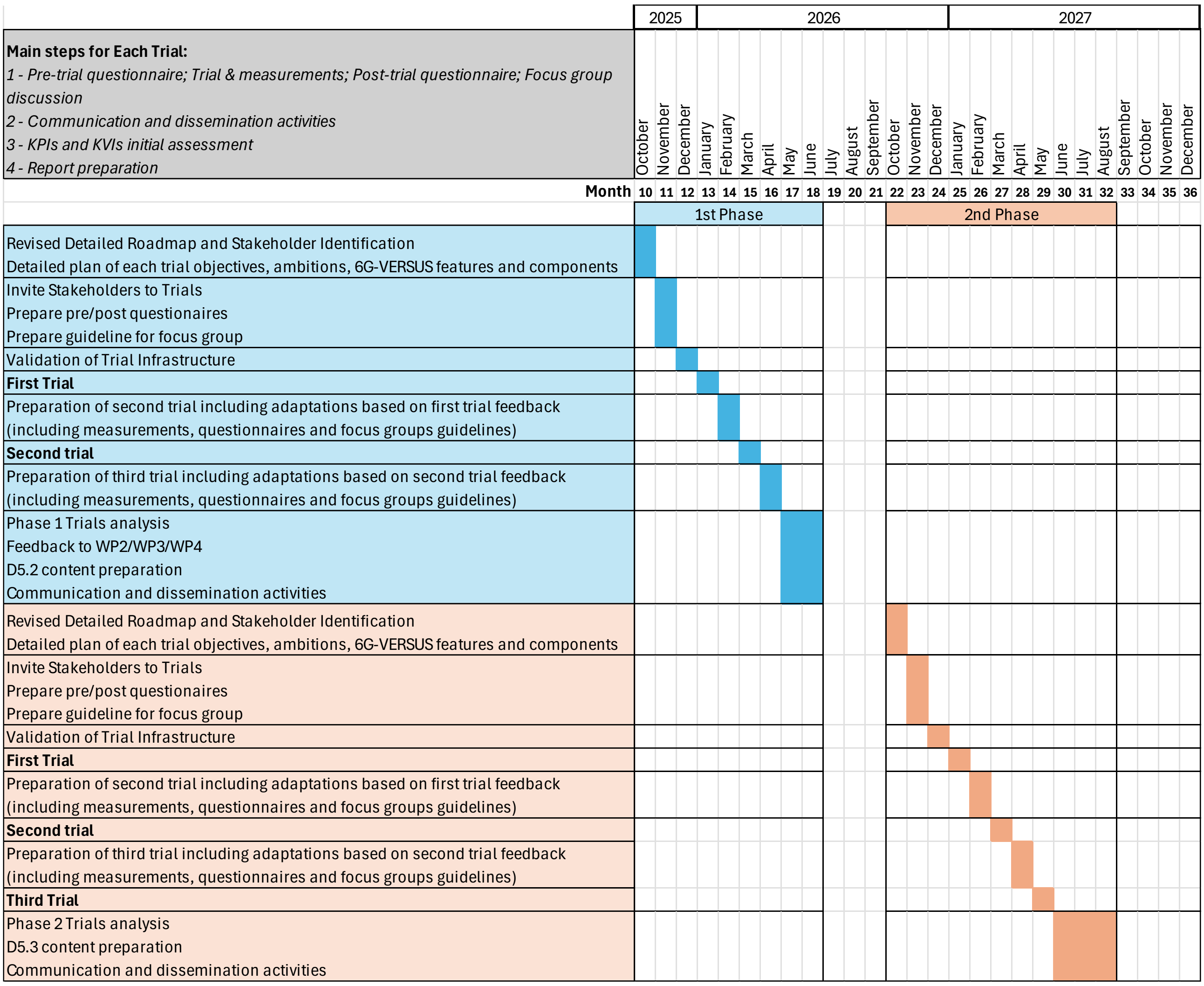

Figure 1: Template for trials timeline.

2.1.1 Planning

The planning phase establishes the foundation for each trial. The first step is to identify the type, scope, and target dates of the trial, i.e., whether it is a technical showcase or a stakeholder-driven demonstration. Each trial must have clearly defined objectives that combine technical and business goals, specifying which scenarios and use cases will be implemented.

The planning process involves the creation of a detailed timeline, which covers all stages of the trial lifecycle. These stages range from initial preparations to post-execution evaluation. Milestones, integration points, and deadlines must be clearly documented to support coordinated action. At this stage, a list of relevant stakeholders for each trial must be prepared, considering the trial storyboard and its technical, business, and societal impact. Another key element of this phase is the coordination of all partners involved in the trial as a pre-trial preparation. This includes assigning roles (e.g., network configuration, application provisioning, measurement collection) and ensuring a shared understanding of the trial's objectives and flow. The plan should also detail the execution steps, including small-scale or laboratory testing, the setup of the network and applications, execution of the trial scenario, possible fault or variation injections, and the procedures for data measurement and collection.

For the first phase, some trials may have some incomplete components which are planned to be finished at a later stage. Nevertheless, it is expected that each cluster demonstrates meaningful progress through a clear storyboard that reflects the objectives of the trial.

Finally, each plan must incorporate a risk assessment and suitable mitigation strategies. These may cover technical risks (e.g., integration delays), environmental risks (e.g., weather impact on outdoor demonstrations), and operational challenges, ensuring that the trial remains viable under varying conditions.

2.1.2 Preparation

This phase is crucial in ensuring that all technical, procedural and logistical elements are in place before the execution of the trial. The process involves the integration of technical components and infrastructure in accordance with the specifications outlined during the planning phase. Before full-scale implementation, rigorous testing and validations are carried out in controlled environments (e.g., in-lab testing, sandbox setups) to ensure system stability and to identify and resolve potential issues. These tests are essential to optimise the end-to-end flow and ensure that all services perform as expected under live conditions during trials.

Upholding the highest standards of data protection, privacy, and regulatory compliance is essential, especially when handling sensitive data or third-party systems. Trials must ensure that the infrastructure adheres to relevant standards and security practices, as outlined in the guidelines from the DMP activities (WP1). Finally, logistics and contingency planning are addressed, including fallback scenarios, backup components, and procedures to deal with unexpected disruptions or failures during the live trial.

2.1.3 Execution

This phase involves executing trials as planned, typically in real-world conditions or realistic simulated environments. Execution should follow the timeline and sequence outlined in the planning phase, with real-time monitoring and responsiveness to any deviations.

During the execution phase, data is systematically collected from all relevant sources, including KPIs from network and system performance, and KVIs from stakeholder interactions using the trial-specific data collected during the trials, including questionnaires, interviews, and focus groups. The details of which data will be collected and analysed. KPI collection must align with definitions established in T2.3 to ensure technical comparability across trials.

The KVI assessment is conducted using a combination of tools, including survey instruments, interviews, questionnaires, and focus groups. The objective is to obtain a minimum of 30 responses for each trial. Each cluster is responsible for designing its questionnaires in collaboration with the KVI assessment team (UBRAD, CES, INC) to ensure methodological alignment with the broader project evaluation framework. Execution also includes stakeholder engagement and showcases, where the trials are presented to external participants. These sessions are essential to validate the practical impact of 6G-VERSUS innovations. Detailed records must be kept, including attendee numbers, profiles and feedback. All trials in both phases are expected to involve external stakeholders. This is crucial for gathering meaningful feedback and assessing the relevance and applicability of the solutions being tested. The first version of the KVI validation framework is described in the section 3, which will support the activities in WP6 and the impact reporting in deliverables D6.1 and D6.2.

2.1.4 Reporting

Following the execution of the trial, each cluster is required to prepare a comprehensive report. This report should summarise the trial's goals, setup details, execution process, results, and outcomes. This includes a detailed analysis of KPIs and KVIs, lessons learned, and any issues encountered, which, in the case of the first phase trials, will serve as input to refine the second phase ones. The report should follow the reporting template, thereby ensuring consistency and comparability across all trials. These reports serve not only as internal documentation but also as a basis for broader dissemination activities, including publications and public communications.

Topic	Description
Setup and validation tools	
Infrastructure setup	Description of the infrastructure setup, including testbed details, 6G-VERSUS infrastructure, etc.
Technical setup	Description of the technical setup, including Vertical Applications, user equipment, and other relevant components.
Measurement and validation tools	Description of the measurement and validation tools used during the pilot to collect/store/provide metrics, KPIs, KVIs, etc.
DMP and Ethics	Description of how DMP and Ethics procedures are followed during the trial (see DMP for reference).
Trial execution and results	
Scenario(s)	Description of the trial's execution, including the multiple scenarios, if applicable and respective results, with pictures and videos.
KPI and KVI validation summary	Detailed KPI and KVI validation summary, with results, conclusions, etc.
Open Data availability	Detailed Open Data availability description. How it will be gathered and how it will be made available (using DMP as a reference)
Showcasing	
Description (per event)	Description of showcasing events performed, including narratives, key presences, number of attendees, targeted audiences, photos and videos, etc.
Stakeholder feedback and impact	
Stakeholder Feedback and Impact	Report on the impact of the pilot, including the number of attendees, a description of the stakeholders, feedback received (such as interviews, questionnaires, and quotes), and any relevant dissemination-related information.
Lessons Learned	
For each trial	Report on the impact of the pilot, including the number of attendees, a description of the stakeholders, feedback received (such as interviews, questionnaires, and quotes), and any relevant dissemination-related information.

Table 1: Main reporting topics

2.2 Risks

The following risks have been identified as potentially affecting the successful lifecycle of the trials and pilots, encompassing aspects such as planning, stakeholder engagement, technical readiness, execution, and data collection. These risks highlight possible challenges that may compromise the quality, reliability, or impact of the results, and will be monitored and mitigated throughout the project.

ID	Probability	Impact	Risk Description
01	Low	High	Clusters prioritise technological development over organising events and trials.
02	Medium	High	Misalignment between trial scenarios and the project objectives.
03	Medium	High	Challenges include identifying relevant stakeholders and ensuring their availability to participate.
04	Medium	Medium	Risk of low stakeholder attendance or lack of diversity among participants.
05	Low	Medium	Limited understanding or engagement from stakeholders attending trials.
06	Low	Medium	Difficulty in organising and coordinating trials across partners and locations.
07	Medium	Low	Adverse weather or environmental conditions affect the execution of outdoor trials.
08	Low	Medium	Equipment malfunction or failure during trial activities.
09	Low	High	Safety incidents involving personnel during trials.

10	Low	Medium	Delays due to administrative or regulatory approvals required for specific trials.
11	Low	High	Delays in receiving infrastructure, applications or components from technical WPs (e.g., WP3 and WP4).
12	Medium	Medium	Risk of interoperability issues when integrating components from multiple partners and testbeds.
13	Medium	High	Insufficient collection or analysis of KPIs/KVIs, limiting the value of trial outcomes.
14	Medium	High	Inadequately designed questionnaires or interview guides can lead to ambiguous or biased responses, reducing data quality.
15	Medium	Medium	Low response rates or limited participation in surveys and focus groups undermine representativeness.
16	Low	Medium	Misinterpretation of questions by participants can result in inconsistent or unreliable feedback.
17	Low	Low	Inadequate coordination or moderation of focus groups can lead to unbalanced discussions and limited stakeholder engagement.

Table 2: Trial risks

2.3 Bulgarian Cluster

The Bulgarian cluster will feature four trials, two per phase, aimed at demonstrating the benefits of B5G/6G technologies to enable Distributed Energy Resources (DERs) in remote areas and actively participate in ancillary service markets. This participation is expected to enhance revenue streams while supporting the grid with fast and coordinated responses during disturbances.

The trials focus on two different grid modes: normal operation, focused on real-time DER monitoring via 5G-enabled Phasor Measurement Units (PMUs) installed at Medium Voltage (MV) feeders, and abnormal operation, where the Wide Area Control (WAC) module of the V-App Central issues event-driven commands to DERs to provide rapid frequency and voltage regulation. The integration of V-App Central, WAM-C (Wide Area Monitor and Control) in the cloud and V-App Regional Controllers (RCs) at the Multi-access Edge Computing (MEC) infrastructure ensures both wide-area observability and localised, constraint-aware DER control. Across the four trials, the cluster progresses from baseline 5G performance validation to advanced Ultra-Reliable Low Latency Communications (URLLC)-enabled monitoring and control, and ultimately to a cognitive, self-optimising grid service leveraging the cluster X-Apps. This approach showcases how B5G/6G connectivity, network slicing, and AI-driven optimisation can transform DERs from passive energy providers into active assets for grid stability and resilience. The planned timeline for the Bulgarian cluster is represented in Figure 2.

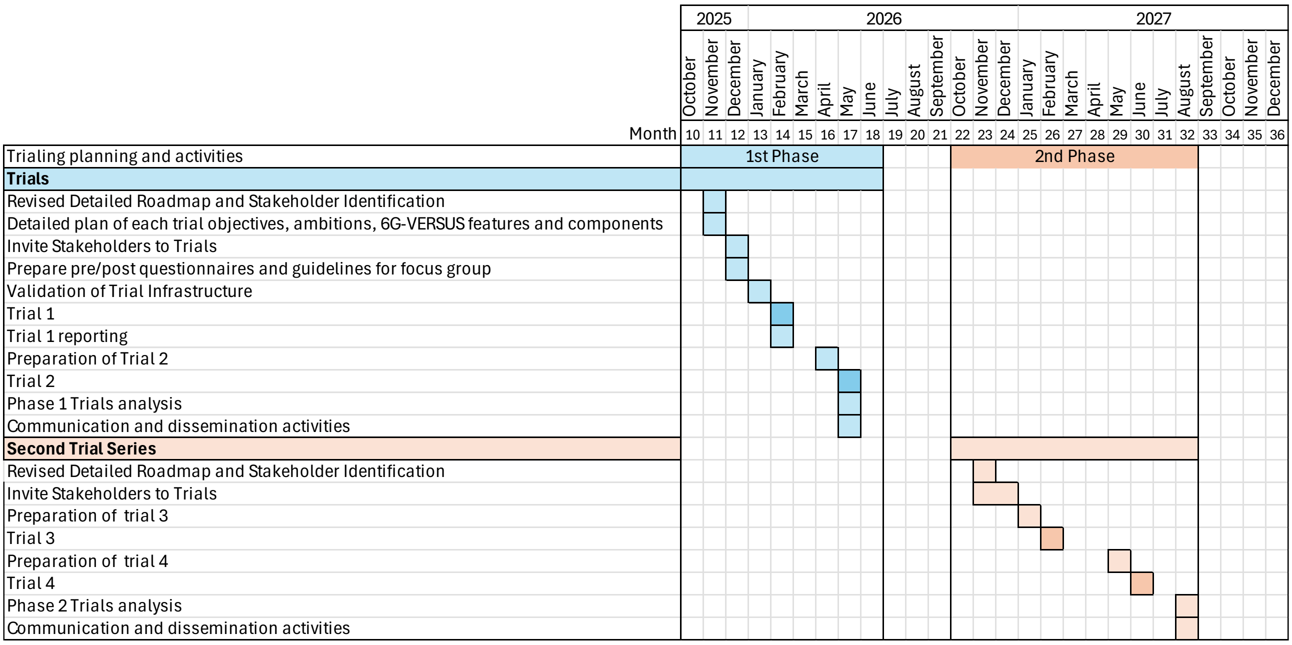

Figure 2: Bulgarian cluster timeline.

The first trial within each phase (Trial 1 in 1st Phase and Trial 3 in 2nd Phase) will demonstrate the normal mode of grid operation (most common case), where the main goal is to demonstrate the ability to monitor remotely located DERs via PMU and 5G (help Renewable Energy Source (RES) and BESS (Battery Energy Storage System) owners prove compliance with grid code requirements. In this scenario, EE leads the end-to-end setup on A1's 5G testbed, with SCBG deploying the V-App Central (WAM-C) -central cloud, and V-App Regional (RC), at MEC, and the N-App for network exposure and control hosting; UCY contributes the WAC/RC control logic of the V-App central and V-App regional, and with the reference implementation, while HPE supports Open-CAPIF alignment. PMU streams from MV feeders are ingested in near real-time by WAM-C for wide-area observability, while the RC provides localised, constraint-aware views at the grid edge. The trial validates the interfaces between V-App and N-App, as well as the synchronisation of synchrophasor data, baseline End-to-End (E2E) latency/jitter, time synchronisation integrity, data completeness, and reliability. From a business perspective, the trials are expected to show how continuous, verifiable monitoring can help RES/BESS owners document grid code compliance and operational readiness for the future ancillary service market. The primary stakeholders' targets to attend these trials are industry and academia stakeholder in Bulgaria, such as TSO/DSO in Bulgaria, energy technology providers, students and academics from academic institutions with an interest in innovative solutions in the energy domain across Bulgaria, which participate in the Winter Energy School Event organized by Energy Department of Sofia University. The event is usually covered by specialized tech media in Bulgaria such as Utility Magazines. Expected attendance includes ≥40 from academia, ≥5 local media, and ≥10 participants from Small and Medium-sized Enterprises (SMEs)/industry. The first trial is expected to be organised as part of the Winter Energy School event (a national energy-focused event in Bulgaria), which is organised annually by Sofia University, Energy Department. The event is also covered by specialised media (e.g., Utility Magazine, University News Outlet, etc).

The second trial within each phase (Trial 2 in 1st Phase and Trial 4 in 2nd Phase) represent the second operation mode of the grid (abnormal, event-based and rare case), validating the first step toward event-driven control under abnormal conditions. The focus is on the X-Apps interfaces: the V-App-central (WAM-C) ↔ N-App control path, the Core Network (CN) Application Programming Interfaces (APIs) exposed via Common API Framework (CAPIF), and the initial development and training of AI models strictly for smart-grid use (e.g., disturbance detection and asset selection). Upon a simulated frequency excursion, the WAC issues event-triggered set-points that traverse the N-App to selected DERs for Fast Frequency Support (FFR). At the same time, the RC enforces local constraints and safe operational envelopes. The trial verifies message routing, authorisation, and discovery via CAPIF, command delivery latency, actuation reliability, and safe fallback behaviour. From a business standpoint, it demonstrates how DERs can shift from passive producers to controllable assets providing ancillary services, engaging RES/DER owners, grid operators, and the regulator. Expected attendance targets ≥40 from the scientific community, ≥5 local media, ≥10 SMEs/RES industry, ≥2 government bodies, and ≥2 grid operators. The highlights of the Bulgarian trials are summarised in Table 3.

Technical Goals	Validate the connection between V-App and N-App and the smart grid sensors, as well as the V-App/N-App interface logic.
Business Goals	Demonstrate the ability to monitor remotely located DERs via PMU and 5G (help RES and BESS owners prove compliance with Fault Ride Through (FRT) requirements.
Partners Involved & Roles	EE: leads the trial, UCY provides the V-App (WAC and RC) A1: provides the testbed and N-app (hosting) SCBG: provides the V-Apps (WAM&C - central, RC-regional), and the N-App) UCY: provides support to the trial setup
Target Stakeholders	Academic Community, SMEs (electricity traders, technology service providers, independent energy producers), grid operators, and the general public.
Expected Attendance (per trial)	> 40 participants, Local Media coverage in at least 3 local/national news articles.

Table 3: Bulgarian cluster trials highlights

2.4 Finnish Cluster

The Finnish cluster showcases self-sustainable B5G/6G base stations in extreme Arctic conditions, demonstrating how AI-driven energy management, predictive weather modelling, and dynamic Radio Access Network (RAN) control enable continuous connectivity despite harsh environments and no power grid access. Trials will focus on a remote site equipped with Radio Unit (RU), renewable energy sources (solar, wind, hydrogen), battery storage, and wireless fronthaul, highlighting energy optimisation and resilience. By leveraging edge Artificial Intelligence/Machine Learning (AI/ML) algorithms and advanced Open-RAN (O-RAN) architecture, the trials will validate energy net-positive operations, autonomous network adaptation, and resource-efficient connectivity. They will demonstrate societal impact through digital inclusion in remote Arctic areas, resilient communication under grid failure scenarios, and the practical integration of emerging technologies such as humanoid robots and high-frequency transceivers.

The Finnish cluster will run a 1-year-long test for the Remote BTS platform during the course of this project. The four trials organised in parallel to the platform test will demonstrate and showcase to relevant stakeholders the benefits of AI-enhanced, energy-aware 5G/6G networks for critical infrastructure, highlighting sustainable operations, stakeholder value, and readiness for extreme environments. The planned timeline for the Finnish cluster is represented in Figure 3. As presented, the Finnish trial will have four trials. The focus of these trials is:

Trial1 (M15): Focus on Remote BTS platform simulations, architecture options and selections. The questionnaire will be organised as part of this trial.

Trial 2 (M18): Focus on the first results of the remote BTS platform tests. This trial will include a live demo, and feedback from stakeholders will consist of questionnaires.

Trial 3 (M26): Focus on results from the winter period of the platform tests. This is the critical period of the test, which determines whether the platform architecture design and component selections were done correctly. This trial will include a live demo, and feedback from stakeholders will consist of questionnaires.

Trial 4 (M31): At this phase, the test of the remote BTS has been completed after executing the tests for 12 consecutive months. The focus of the fourth and final Trial is on completed test results, lessons learned, and possible changes proposed to the architecture. This trial will include a live demo, and feedback from stakeholders will consist of questionnaires.

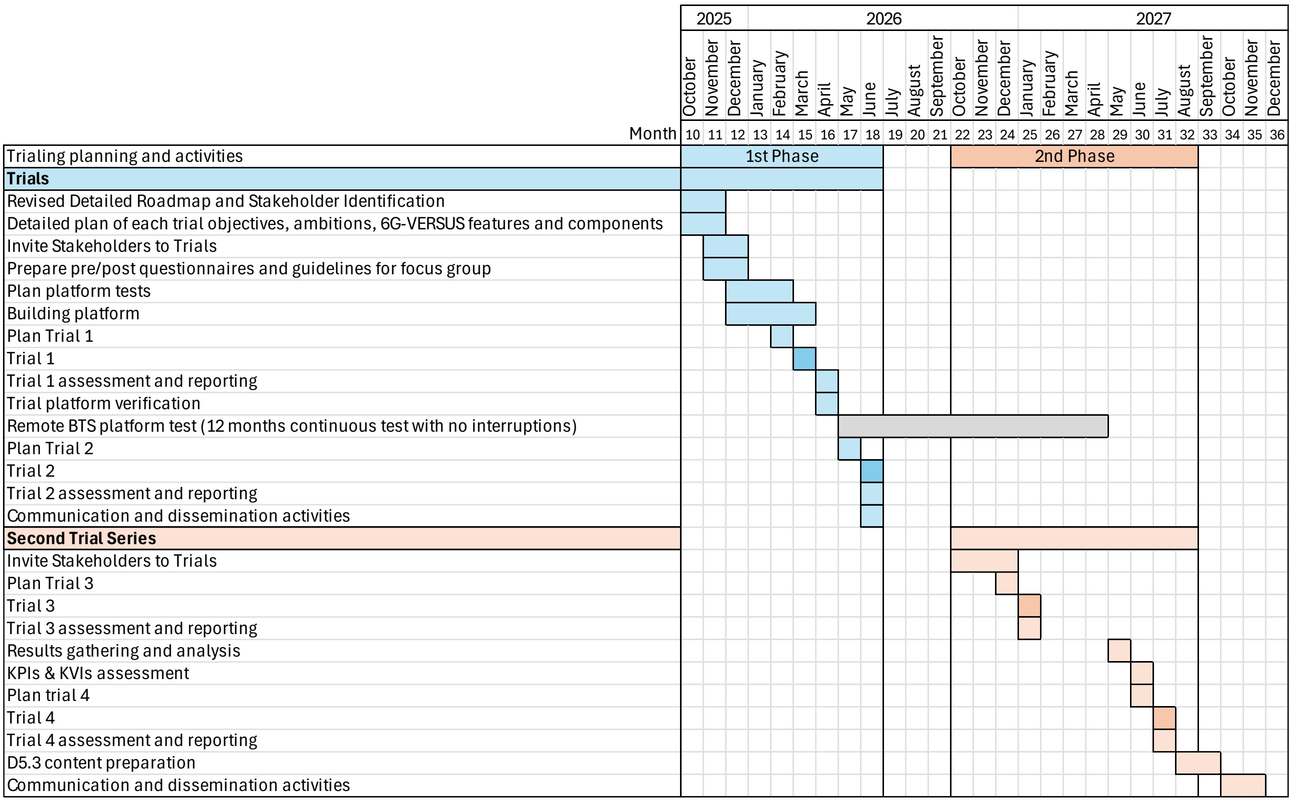

Figure 3: Finnish cluster timeline.

The Finnish cluster is planning a one-year test for the remote BTS platform under a single scenario: deploying and operating a 1-year run pilot of a self-sustaining 5G RAN in Arctic conditions, including months of limited sunlight, with the development and integration led by UOULU, the test progresses from establishing energy balancing of power consumption by introducing AI-based prediction and energy-aware control using the cluster X-Apps (V-App for policies, N-App for telemetry/exposure and actuation, AI-App for forecasting and decisions). During the winter months, the test expands its scope of operation through the polar night to experiment with deep sleep windows, Multiple Input Multiple Output (MIMO) muting, dynamic bandwidth usage and prioritised connectivity for control/telemetry, culminating in closed-loop autonomy with minimal human intervention and improved energy efficiency. The main business goals of Finnish trials focus on increasing revenue opportunities for remote-area actors (e.g., reindeer owners, tourism) while decreasing Operating Expense (OPEX) through reduced energy draw, fewer site visits, and evidence-based rightsizing. Stakeholder engagement targets telecom operators, network technology providers, Arctic-zone municipalities and organisations, local business representatives, the UOULU research community, national/international research networks, and media, delivered via on-site trials and guided visits. The highlights of the Finnish trials are summarised in Table 4.

Technical Goals	This use case addresses the deployment of a 5G radio access network in the challenging Arctic conditions where Year-on-Year (YoY) continuous connectivity must be maintained despite the harsh environment, weather, and a period of almost no direct sunlight for approximately 3 months. The network architecture is designed to enable an intelligent energy-aware decision-making system that combines AI-based predictions, real-time monitoring and optimised control over power saving measures to ensure self-sustainable operations.
Business Goals	Increase in revenue of businesses in remote areas (reindeer owners, tourism-related businesses in the Arctic region). Decrease in OPEX
Partners Involved & Roles	UOULU: Development, integration and verification/validation, trials

Target Stakeholders	Telecom operators, network technology providers, municipalities in the Arctic zone, the research community in UOULU, national and international research networks, and national and local newspapers
Expected Attendance (per trial)	Attendance greater than 30 persons, including telecom operators, network technology providers, municipality representatives and researchers.

Table 4: Finnish cluster trials highlights

2.5 French Cluster

The French cluster advances data-driven management of water networks through B5G/6G trials that pair real-world deployments with EURECOM's Open5GLab. The programme targets two use cases (sewer monitoring and sustainable irrigation) and is led by GreenCityzen with EURECOM as testbed provider and Thales Six GTS as developer/integrator. Trials run on an open, cloud-native 5G Standalone (SA) stack (OpenAirInterface (OAI) RAN/CN) with RAN Intelligent Controller (RIC)/xApps, Network Data Analytics Function (NWDAF) / Network Exposure Function (NEF) exposure and CAMARA northbound APIs. At the same time, Reduced Capability (RedCap) devices, Discontinuous Reception (DRX)/eDRX and Bandwidth-Part (BWP) adaptation are exploited to curb device and network energy use. Field activities are centred in Marseille, with lab integration and orchestration at Sophia Antipolis. The focus is on turning raw telemetry into actionable operations, reducing alert fatigue with context-aware filtering, virtualising sensors to shrink hardware footprint, and using edge-AI to prioritise interventions.

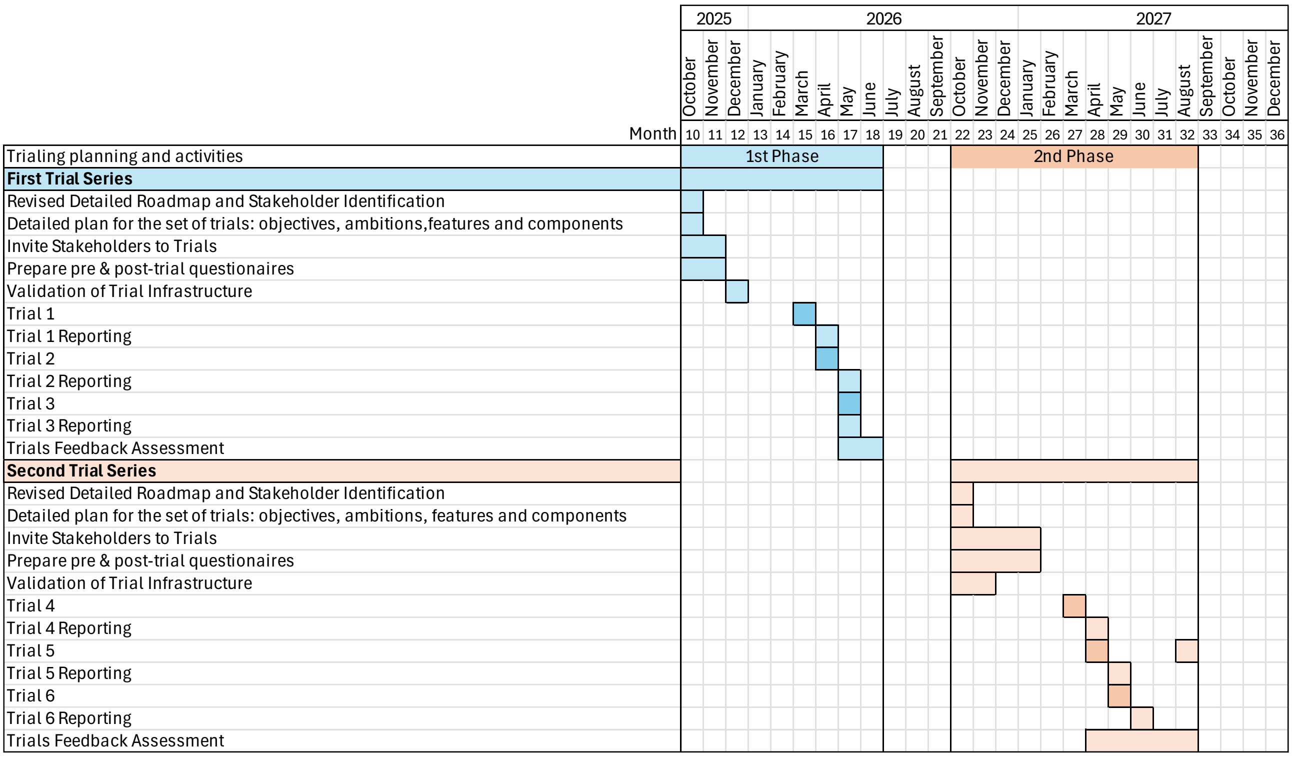

Figure 4: French cluster timeline.

From a technical perspective, the cluster showcases large-scale sensor virtualisation to halve physical probes in rainwater networks, abnormal-behaviour detection in sewers to lower inundation/pollution risk, leak and over-watering detection with automatic schedule optimisation, and AI-assisted RAN tuning (DRX/BWP) to extend battery life while meeting latency targets. The impact of B5G/6G technologies is assessed via quantified KPIs/KVIs. These trials are expected to show improved battery-lifetime gains through AI-optimised DRX/BWP, reduced E2E latency and false-alert reduction. Moreover, hardware footprint can be reduced via virtual sensors, and device-network energy-efficiency gains with RedCap. Outcomes emphasise higher Return On

Investment (ROI, fewer site visits, lower energy/chemicals), reduced electronic waste, and measurable environmental benefits communicated as real-time indicators to operators and public stakeholders. Measurement uses Keysight traffic/latency tooling and inline power metering for rigorous, repeatable evaluation. The planned timeline for the French cluster is represented in Figure 4.

The French cluster will run six trials (three per phase) covering three scenarios repeated across phases to demonstrate AI-native, standards-based water-network services at the city scale. The first scenario (Toulon rainwater network), EURECOM's 5G testbed with RedCap and CAPIF exposure integrates GREEN's platform to produce real-time environmental benefit indicators and estimate waste volumes with fewer physical sensors via AI-driven sensor virtualisation, complemented by a V-App alert dashboard with drill-down mapping for CITEO reporting. Scenario 2 (smart watering in Sète) employs the same V-App/AI-App/N-App stack and Machine Learning Operations (ML-Ops) pipeline to optimise interventions and reduce instrumentation/OPEX while preserving Quality of Service (QoS). Scenario 3 (energy-aware RAN + leakage detection, Marseille focus) introduces NWDAF exposure and THALES's AI-App for RAN, enabling dynamic tuning of DRX wake-up cycles based on sensor traffic, leakage detection, and the continuous flow of indicators for Agence de l'eau reports. Across all scenarios, EURECOM acts as a 5G testbed, CAPIF/NWDAF provider and ML-Ops owner; GREEN supplies the city playgrounds (Toulon, Sète, Marseille), historical datasets, and V-/AI-App features; THALES contributes real-time AI control for RAN reconfiguration. The cluster's business focus is to increase ROI for customers (less hardware, lower energy/use of spectrum, fewer field visits) while meeting mandatory reporting obligations. The target stakeholders include public/private water operators, public institutions, tool vendors, engineering firms, and media (specifically Toulon Métropole, CITEO, Agence de l'eau, Zug Consulting, Smart City Magazine). Expected attendance is >30 relevant stakeholders from public and private water network operators, water management technology providers, regional and municipal authorities, regulatory and funding bodies, utilities, industry representatives, and local media. The highlights of the planned trials in the French Cluster are summarised in Table 5.

Technical Goals	Demonstrate the capability to produce a real-time environmental benefit indicator and the ability to provide relevant information on waste volume in the rainwater network, reducing the number of deployed sensors. Demonstrate the capability to detect abnormal behaviour in the sewer network. Demonstrate the capability of saving energy by dynamically adjusting the RAN wake-up cycles in DRX, tailored to the sensors' traffic. Demonstrate the capability to produce in real-time environmental benefit indicators. Demonstrate the capability to detect abnormal behaviour such as leakages.
Business Goals	Provide the customer with the expected indicators to be provided to their own funder in mandatory reports (CITEO). Increase the return on investment for our customers. Provide advanced IA services better to manage the sewer network operation for our customers. To demonstrate how AI-driven applications and standardised 5G/6G interfaces (e.g., CAPIF, NWDAF) can enable smarter, more energy-efficient and interoperable vertical services.
Partners Involved & Roles	**EURECOM**: 5G Testbed, RedCap, CAPIF Provider, ML-Ops Owner **GREEN**: GREEN provides the playground for the trials. In scenario 1, the playground is a real-life deployment of rainwater network monitoring in the city of Toulon. GREEN provides its existing platform, historical dataset for models training and indicators feeding, as well as the V-App and AI-App feature developments carried out during the project. **THALES:** Optimised version of AI-App RAN with dynamic reconfiguration based on AI-App trained on ML-Ops platform; Tests with real-time inference based on multiple traffic profiles.
Target Stakeholders	Public and private water network operators, Water management tools and technology providers, Toulon Metropole, CITEO, City of Marseille, Agence de l'eau, SUEZ Sète (Water Utility Company), SUD Alsace (municipality), Zug consulting, Smart city magazine
Expected Attendance	> 30 attendees

(per trial)	

Table 5: French cluster trials highlights

2.6 Greek Cluster

The Greek cluster will validate the role of B5G/6G-enabled collaborative robotics and AI-driven automation in wildfire detection and Search and Rescue (SAR) operations. The cluster focuses on enhancing situational awareness, response efficiency, and responder safety in challenging environments, where wildfires and natural disasters present high risks to both populations and first responders. The trials exploit a combination of UAVs (Unmanned Aerial Vehicle, drones), robotic dog, AI-based fire detection, and AR-enabled interfaces, all supported by low-latency, high-reliability 5G/6G communication. This integration enables early fire detection through real-time High Definition (HD)/4K video analysis, wide-area reconnaissance, safe route identification, and the deployment of autonomous assets in hazardous areas. UAVs act as both aerial sensors and relay nodes, while robotic dogs provide ground-level inspection, delivery of supplies, and victim communication.

The trial program is structured in two phases. In Phase 1, trials focus on establishing baseline capabilities for pre-event surveillance, wildfire monitoring, and patrolling of high-risk areas. UAVs and robotic dogs are deployed to provide real-time video streaming, anomaly detection, and intrusion monitoring, ensuring early alerts and situational awareness. In Phase 2, the focus shifts to event-driven and post-event operations, including autonomous SAR missions with UAVs and robotic dogs, real-time detection and tracking of missing persons, safe route mapping for manned responders, and transport of lightweight supplies. Post-event trials address restoration, wide-area damage assessment, and fire-front progression tracking in alignment with INSARAG standards (ASR1–3).

The trials are supported by the Athens testbed, which provides end-to-end 5G connectivity with optimised bandwidth and latency management for emergency scenarios. Enablers include edge computing for real-time analytics, network slicing for prioritisation of SAR traffic, and an Energy Efficiency as a Service (EEaaS) framework to reduce ecological footprint. This ensures not only reliable operation under adverse conditions but also alignment with sustainability goals through efficient resource allocation, minimised digital footprint, and improved safety of first responders and affected communities. The planned timeline for the Greek cluster is represented in Figure 5.

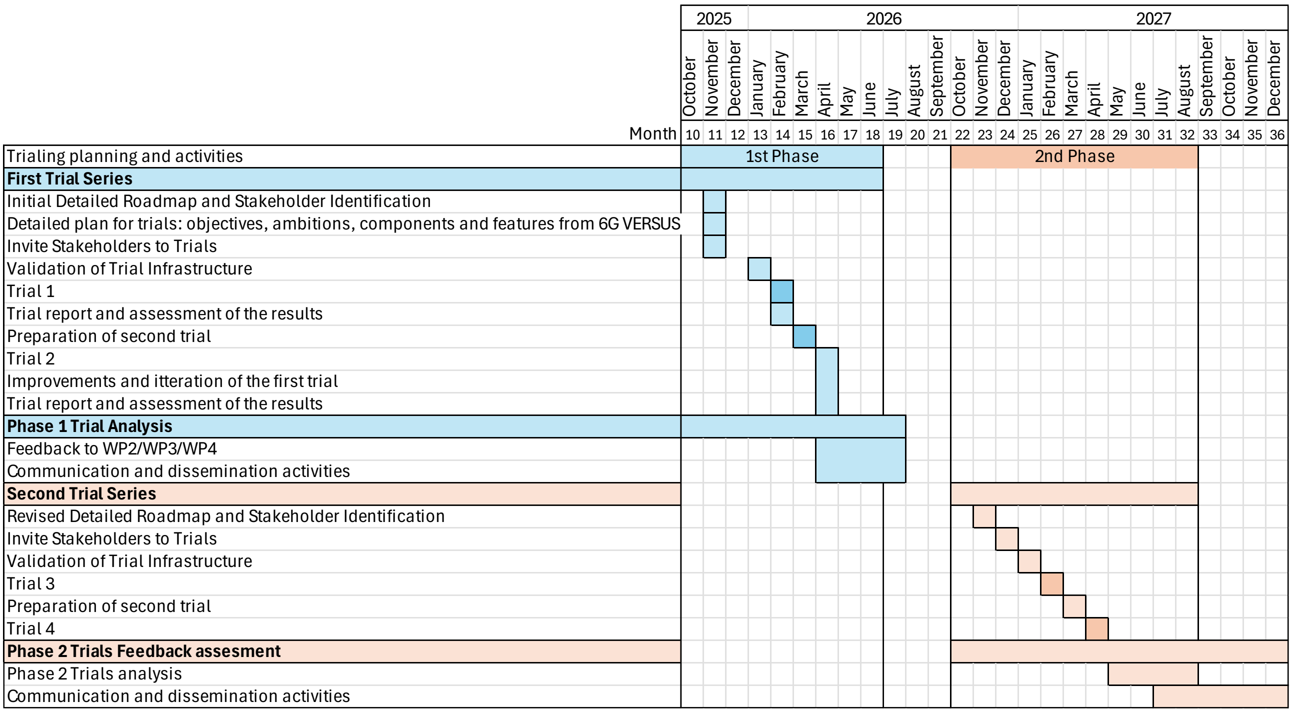

Figure 5: Greek cluster timeline.

The Greek cluster will conduct four trials (two per phase), covering three wildfire-monitoring scenarios that are repeated across the two phases, to evaluate the impact of B5G/6G-enabled technologies on early wildfire detection and response. Scenario 1 focuses on validating the N-/AI-/V-App interfaces for wildfire monitoring: NCSRD provides the N-App and NEF/CAPIF exposure, INF delivers the AI-App, OTE provides the V-App, and ICCS provides the INtelligent UxV Surveillance (INUS) Asset Management Platform (INUS Asset V-App); the business goal is early wildfire detection with reliable end-to-end telemetry and control. Scenario 2 extends this by validating the INUS drone V-App for streaming-video connectivity and on-board/edge object recognition, combining UAV feeds with fixed monitoring cameras through the same 6G-App interfaces to strengthen multi-source detection and reduce time-to-alert. Scenario 3 (trialled in both phases) complements the above with additional operational workflows and resilience under realistic field conditions, maintaining the early-detection objective while exercising the same CAPIF-mediated interfaces. Target stakeholders include the Municipalities of Agia Paraskeui and Egaleo and the Special Unit EMAK (subject to availability), alongside the research community (≥10 participants from NCSRD institutes, spin-offs, Lefkippos Tech Park, Ahedd DIH, and Egaleo’s entrepreneurship hub), local media (≥1 outlet), municipal representatives (≥1–2) and regional authorities (≥2), ensuring both technical validation and public-sector engagement. The Greek cluster trials highlights are summarised in Table 6.

Technical Goals	Validate part of 6G-App interfaces (N-App, AI-App and V App wildfire monitoring), Inus V-App (drone) in terms of connectivity and object recognition (streaming video).
Business Goals	Demonstrate early wildfire detection with multiple means, including through UAV and monitoring (camera) and V-App wildfire monitoring
Partners Involved & Roles	NCSRD: provides the N-App and NEF/CAPIF INF: provides the AI-App OTE: provides V-App and AI-App for Early Wild-Fire Detection ICCS: provides the V-App and the Grafana-based Energy Efficiency as a Service measurements dashboard; AI-Apps for Intelligent identification of casualty location, and for Virtual Fence and Prohibited Area Detection; N-Apps for Coverage extension as well as for Energy Efficiency as a Service.
Target Stakeholders	Municipality of Agia Paraskeui, Municipality of Egaleo, Special Unit EMAK (subject to availability), Regional Authorities (e.g., local police, firefighters), other Institutes of NCSRD, invitations to spin-offs, high-tech SMEs from Lefkippos Tech

	Park, external SMEs in the framework of the digital innovation hub, local Media with interest in promoting technological aspects of the trials
Expected Attendance (per trial)	Scientific Community: >15, Local Media: >5, SME/Industry: >10, Government bodies: > 6, Municipalities: >2, Infrastructure operators: >5, Environmental organizations: >2

Table 6: Greek cluster trials highlights

2.7 Portuguese Cluster

The Portuguese cluster will demonstrate the integration of B5G/6G technologies and their real-world impact in a use case located at the port of Aveiro. Utilising ultra-low-power Internet of Things (IoT) sensing and edge AI orchestration, we will enable a Digital Twin (DT) of the Port of Aveiro. Additionally, sensor-equipped drones and RC-boats with autonomous navigation capabilities will be employed to target environmental port monitoring, operational safety, and infrastructure energy efficiency. The cluster trials are structured into two phases, each comprising four trials: three scenario-specific trials addressing environmental awareness, safety awareness, and energy management, and a fourth trial combining the three scenarios.

The trials will focus on real-time monitoring and control using zero-energy and NR RedCap-enabled IoT sensors, UAVs, and RC boats for adaptive data collection and camera-based high-resolution environmental assessment. They will validate edge AI-based inference, dynamic network slicing, and context-aware resource allocation for environmental monitoring, as well as the energy-efficient operation of Electric Vehicle (EV) chargers and street lighting. The integrative trial will demonstrate the orchestration of multiple network slices and V-App/N-App interactions, leveraging NEF/CAPIF APIs and Green AI to ensure low-latency, energy-aware, and reliable operation across all verticals. On a technical perspective, these trials will evaluate sensor performance, data accuracy, system responsiveness, energy consumption, and multi-scenario coordination, while enabling end users and stakeholders to visualise, predict, and respond to environmental, safety, and energy events. The results are expected to validate regulatory compliance of port environmental monitoring, operational efficiency, and sustainability gains, showcasing the impact of advanced B5G/6G technologies.

The planned timeline for the Portuguese cluster is represented in Figure 6.

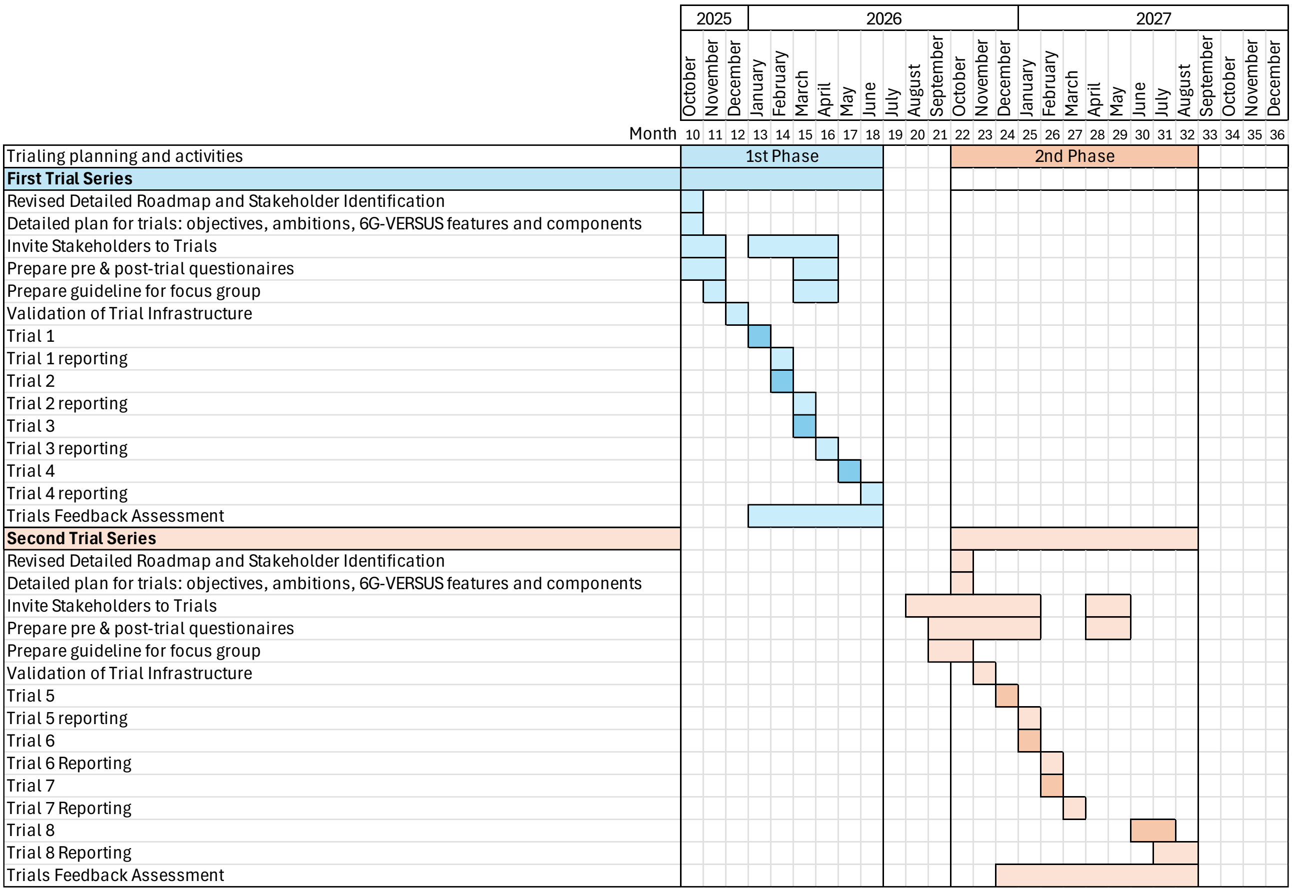

Figure 6: Portuguese cluster timeline.

The Portuguese cluster delivers eight trials (in each phase, trials 1–3 are related to three distinct scenarios and trial four merges all three scenarios). The first scenario showcases environmental monitoring in Ílhavo/Gafanha da Nazaré to assess the environmental impact of 6G-enabled technologies in the port and its surroundings. It includes two main V-Apps (Environmental Monitoring and Mobitrust), and leverages 5G NR RedCap, CAPIF, and QoS APIs to cut energy in networking and computing resources. The main roles are performed by JSIO (V-App & IoT), IPN (IoT device energy management AI-App, QoS N-App), ITAV (Network Operations & Handover N-Apps, network testbed support), APA (end-user/logistics), ONE (Mobitrust and ONEmNEF). The second scenario assesses the impact of B5G/6G technologies on port safety and logistics within the port premises. From a technical perspective, it includes the Slice Manager to reduce energy in edge AI workloads. The main roles are performed by ITAV (AI-App, V-App, N-App, Slice Manager, testbed), APA (end-user/logistics), and ONE (Mobitrust and ONEmNEF). The third scenario is related to smart city energy management. It uses device–location–aware public-lighting adaptation and AI-based prediction of energy production/consumption to support grid optimisation. The main roles are performed by ALB (6G-App triplet + testbed), ONE (OneMNEF/NEF), APA (end-user).

Finally, Trial 4 (performed in each phase) merges the three scenarios into a unified workflow (exposing capabilities via CAPIF/NEF, enforcing QoS/slices, and closing the loop with AI inference) to evaluate the impact on OPEX/CAPEX (Capital Expenditures), and to measure energy savings across environmental, port, and municipal contexts. The main target stakeholders include the Portuguese Environmental Agency, ADRA, port-based chemical firms, Port Authority and logistics operators, local municipalities (Ílhavo, Gafanha da Nazaré), telecom operator (MEO), infrastructure teams, and Universidade de Aveiro. Expected participation per trial is ≥40 researchers, ≥5 media representatives, ≥10 SMEs/industry participants, ≥6 government bodies, ≥2 port operators and solution providers, ≥2 municipal representatives, ≥5 infrastructure operators, and ≥2 representatives from environmental agencies. The Portuguese cluster trials highlights are summarised in Table 7.

Technical Goals	Validate device location acquisition (through N-App), to adapt public lightning and energy consumption in both networking and computing needs, and the possibility to predict energy production and energy consumption and adjust the power grid. Demonstrate the unifying role of CAPIF for on-demand access to different Network Functions, accounting for different applicational priorities (e.g., energy efficiency, QoS)
Business Goals	Showcase the value of B5G/6G features (5G NR RedCap, CAPIF, QoS APIs) to reduce energy consumption in both networking and computing needs. Validate the possibility of Slice Manager's ability to reduce energy consumption in networking and computing resources. Public lightning management, using AI capabilities to predict energy production and consumption for energy distribution optimisation. Demonstrate the B5G/6G technologies' impact on Port monitoring capabilities covering environmental, safety and energy dimensions.
Partners Involved & Roles	JSIO: responsible for the V-App & IoT devices provisioning IPN: responsible for IoT device energy management AI-App and the QoS N-App features relating to 5G NR RedCap configuration ITAV: responsible for the network testbed and support with N-App provisioning APA: end-user and location logistics ALB: responsible for network testbed ONE: responsible for Mobitrust and ONEmNEF provisioning
Target Stakeholders	Portuguese environmental agency, ADRA, Telecom operator (MEO) infrastructure teams, Port-located based chemical companies, Local residents' communities and associations, Local Municipalities (Aveiro, Ílhavo), Gafanha da Nazaré parish, Universidade de Aveiro (engineering & environmental research departments)
Expected Attendance (per trial)	Scientific Community: ≥ 40, Local Media: ≥ 5, SME/Industry: ≥ 10, Government bodies: ≥ 6, port operators and solution providers: ≥2, Municipalities: ≥2, Infrastructure operators: ≥5, Environmental organizations: ≥2

Table 7: Portuguese cluster trials highlights

2.8 Spanish Cluster

The Spanish cluster focuses on demonstrating how B5G/6G technologies can impact field management in agriculture and water management by enabling remote operators to improve presence in geographically distributed locations. The trials will utilise vehicles equipped with sensors and actuators (drones or Automated Guided Vehicles (AGVs)) to collect real-time environmental data, interact with local personnel, and perform operational tasks remotely. Immersive telepresence, combined with low-latency, high-throughput 6G connectivity, network slicing, edge AI, and satellite-integrated backup, allows seamless remote operation, ensures reliable decision-making, and optimises resource utilisation while minimising travel-related carbon emissions.

The trial program includes crop monitoring, irrigation optimisation, and wastewater management. These trials highlight the societal benefits for farmers, water agencies, researchers, and local communities by improving operational efficiency, environmental sustainability, and safety. Advanced AI-driven analytics, digital twin modelling, and immersive Graphical User Interfaces (GUIs) enable real-time insights, fostering collaboration, cross-disciplinary knowledge exchange, and scalable solutions for remote agricultural and industrial operations. The planned timeline for the Spanish cluster is represented in Figure 7.

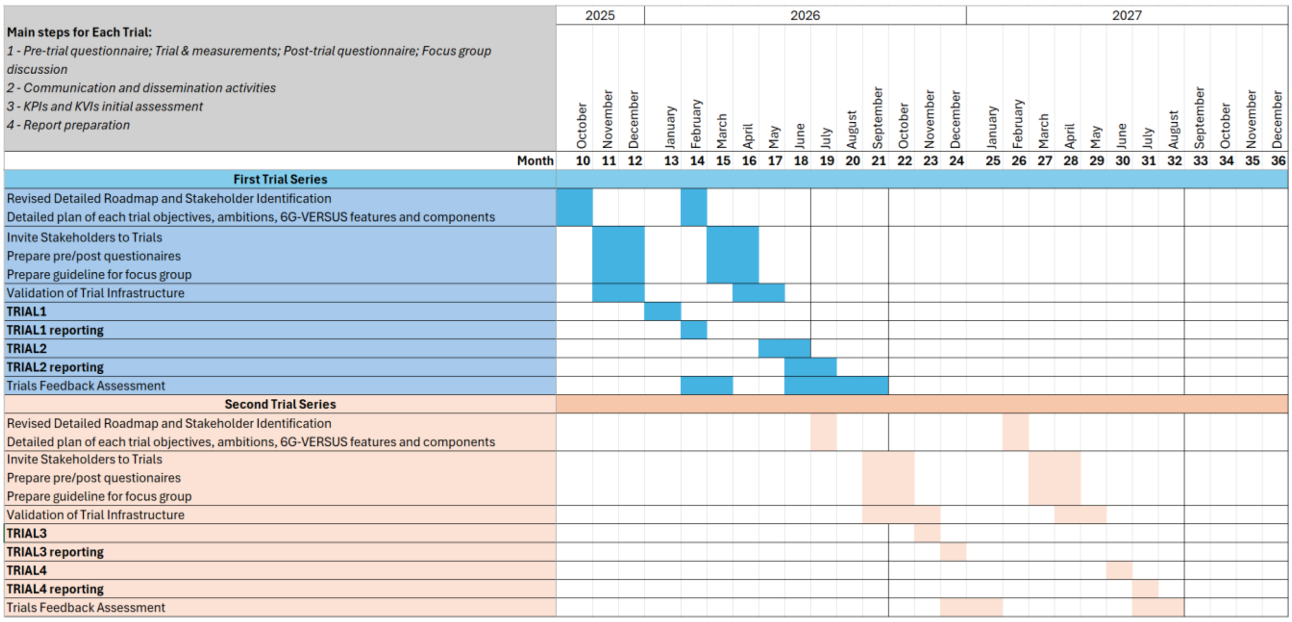

Figure 7: Spanish cluster timeline.

This cluster will run two series of trials, each with two target trial periods. Within those four total trials, five different target test scenarios or use cases across phases will be addressed to validate AI-native, B5G/6G-enabled agricultural water monitoring services. The first scenario (optimal water mixing) couples reservoir sensors (B5G modems) with drone/AGV crop scouting for continuous monitoring and daily decision-making. CET performs the main roles to provide the AI-App (requirements capture and stakeholder feedback), HPE provides support on integration of APIs, NOKIA extends the N-App to expose IPs for non-commercial sensors, CSIC drives requirements/evaluation and sensor integration, and KEYS provides measurement equipment. The second scenario (assisted sampling) equips a drone with robotic arms for sample collection, with UMA building the operator UI, CSIC handling requirements/evaluation/sensors, and TID/KEYS supporting communications and measurement. The third scenario (AGVs with immersive tele-drive & autonomous missions) is led by NOKIA (remote-driving app at La Mayora, autonomy and hyperspectral imaging), with CSIC responsible for evaluation, TID/UMA supporting integration, and KEYS delivering a DT for propagation-loss estimation. The fourth scenario (automatic surveillance & detection) combines drones/robots/Joint Communication And Sensing (JCAS) for leakage and anomaly sensing, with CSIC overseeing requirements/evaluation, UMA/TID handling system integration, and KEYS responsible for measurement. The fifth scenario (water-quality plant operations) involves HIDRA for validation and stakeholder engagement, CET for the AI-App (co-designed with users), and TID for commercial network and OpenGateway API integration (Quality-on-Demand (QoD), Dynamic Airspace Connectivity Data). Target stakeholders for these trials include government bodies, environmental organisations, vertical SMEs/industry, academia and research communities, local media, regional government agencies (ASAJA, Ecologistas en Acción/GENA-Axarquía, Consorcio del Poniente, Delegación de Agricultura y Agua, Fundación Cajamar, and Diputación de Almería). For each trial phase, expected attendees in La Mayora and UMA include the scientific community (network, agriculture, etc.): >15, regional government: >5, ASAJA: farmers association: >13, Ecologistas en acción GENA Axarquia: >5, local media: >5, and municipalities: >2. As for the EDAR Roquetas de Mar pilot site in Almeria, expected attendees are innovation centers (IFAPA, Cajamar foundation, etc.) ≥ 5, regional government (municipalities, provincial councils) ≥ 4, and farmers associations (ASAJA/irrigation communities) ≥4. The Spanish cluster trials highlights are summarised in Table 8.

Technical Goals	Deploy and integrate 5G-connected sensors, drones, and AGVs for monitoring, immersive tele-driving, autonomous missions, and robotic-arm sampling. Fuse multi-source data (sensors, drones, AGVs) for daily decision support in water use and crop management.

	Enable low-latency video, object recognition, hyperspectral imaging, and Digital Twin-based radio-propagation estimation. Implement N-App-based device onboarding, operator frontends, and AI-assisted control with predictive maintenance. Support JCAS-based surveillance and anomaly detection and integrate QoS/OpenGateway APIs for network-aware performance.
Business Goals	Optimise water use and crop yield, reducing operational costs and manual labour through automation. Increase safety by reducing human exposure to repetitive or hazardous tasks. Improve efficiency, throughput, and data reliability in agricultural monitoring, sampling, and farm operations. Enable early leak/fault detection to cut losses, downtime, and insurance costs. Ensure regulatory compliance in environmental and water-quality standards while reducing chemical/energy use. Avoid fines and service interruptions while providing transparent reporting to regulators, funding bodies, and farmer associations.
Partners Involved & Roles	**CET**: Developer of the AI-App - presents how the tool will work once developed and collects feedback from stakeholders regarding the AI-App **HPE**: Provide support for integration of APIs **NOKIA**: WP4 N-App to expose IP address for non-commercial sensors **CSIC**: requirements/evaluation/ sensors **HIDRA:** Supports use case validation and stakeholder engagement **TID:** Provides commercial network and OpenGaetway service integration for APIs: API OGW Quality On Demand, API OGW Dynamic Airspace Connectivity Data **UMA**: Developer of the user-facing front end for interaction **KEYS**: Provision of measurement tools
Target Stakeholders	Regional government representative, ASAJA - Farmer collective organisation, Ecologistas en Accion, Gena-Axarquia
Expected Attendance (per trial phase)	**La Mayora and UMA**: more than 45 from scientific community, regional government, farmers association, Ecologistas en acción GENA Axarquia, local media, and municipalities. **EDAR Roquetas de Mar**: more than 13 from innovation centers, regional government, and farmers associations

Table 8: Spanish cluster trials highlights

3 6G-VERSUS KVI Validation Framework

3.1 KVI Identification Framework

KVI (Key Value Indicator) identification framework describes a structured approach applied in assessing the broader values generated beyond the technical performance of the 6G-VERSUS project. This framework assists the 6G-VERSUS project to align its sustainability outcomes with the global societal goals, such as United Nations Sustainability Goals, European Commission Green Deal, Strategic Research and Innovation Agenda of the Smart Networks and Services Joint Undertaking (SRIA of SNS JU), and European Climate law. It enables the projects to plan and execute their activities with clearly expressed key values for respective Use Cases or Targeted Verticals.

3.1.1 Background

Next-Gen Network development projects have significantly influenced society, businesses, and the environment, creating both positive and negative impacts. To address these challenges, recent technological developments have shifted toward a value-driven approach rather than focusing solely on performance [WBM+24]. This approach assesses the environmental, societal, and economic impacts of designing, deploying, and operating new technologies while striving to amplify positive outcomes and minimise adverse effects. The value-driven approach differs from traditional KPI-focused models by promoting a more holistic assessment of impact across the environmental, societal, and economic dimensions of technology design, deployment, and operation.

Every technology development project must outline the KPIs to measure the technical performance and the KVIs to assess the impact on the environment, society, and economy. KPIs are metrics that measure the performance and the effectiveness of the developed technology, and KPIs are inherently quantitative in nature. In contrast, KVIs are aimed at capturing the broad value contributed beyond the technical aspect of the new technology to society. KVIs have traditionally been considered softer metrics, often qualitative in nature. However, they can also be quantified depending on the context. For example, in the smart agriculture vertical, 'inclusiveness' may be identified as a Key Value. This can be translated into a measurable KVI, such as the number of people living in remote areas who access smart agriculture services, thus turning a qualitative value into a quantitative indicator.

The key terminologies involved in the KVI framework, such as Key Value (KV), Key Value Indicator (KVI), Key Value Enabler (KVE), and Key Performance Indicator (KPI), are defined as follows. (Wikstrom, et al., 2024)

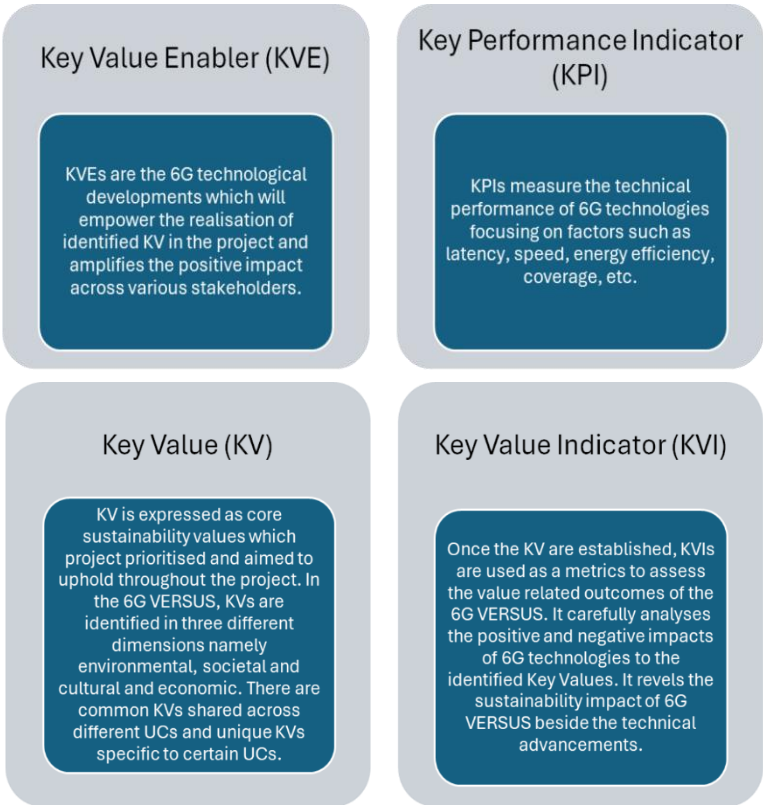

Figure 8: Terminologies related to KVI framework.

In the 6G-VERSUS project, there are six clusters, and the following example is from the Bulgarian cluster.

Sustainability Dimension	Key Value	Key Value Indicator	Description	Key Value Enabler	Key Performance Indicator
Societal & Cultural	Digital inclusion	Increase in access to 6G network for remote communities	As the 6G mobile network is extended into remote areas to connect RES to the grid, this will enable communities in those remote areas to access 6G connections.	Distributed Control System (DCS) architectures and 6G communication	Improved grid stability with URLL Communication
Environment	Environmental sustainability and viability	Increased Energy Efficiency improvements	How do advanced features of 6G contribute to increasing the energy efficiency improvements?	Facilitation of grid modernisation and RES integration	To be identified
Economic	Economic sustainability	Optimising revenue management	How do advanced features of 6G contribute to optimising revenue management?	Enhanced production monitoring and forecasting capabilities	To be identified

Table 9: Example of identification of KVs, KVIs, KVEs and KPIs of a UC

3.2 KVI Identification Process

3.2.1 Step 01: Key Value Identification

In the 6G-VERSUS project, each use case or the targeted vertical must develop its own KVIs to assess the societal impact of 6G technology. The preliminary step in developing the KVIs involves identifying KVs relevant to the respective use cases or targeted verticals. [WBM+24], [HEXII_D11], [HEX_D11]. This begins by benchmarking against global frameworks such as the United Nations' 17 Sustainable Development Goals (SDGs) [UN_15] and European Green Deal targets. [EC_19] For instance, SDG 13 (Climate Action) may guide KVs related to reducing carbon emissions in telecommunications infrastructure. *Identified KVs are validated using existing literature (e.g., peer-reviewed articles, conference papers, white papers)* as well as SNS working groups' and other SNS projects outputs, and aligned with 6G-VERSUS project objectives.

3.2.2 Step 02: Aligning Key Values with Use Cases

Once KVs are identified, they are aligned with use cases or targeted verticals to understand common KVs across all the use cases and targeted verticals or specific KVs unique to each use case and target verticals [HEX_D11], [IMB_D11], [6GS_D51], [HEXII_D11]. The methodology applied in this process includes workshops with respective stakeholders (the cluster partners) and available literature. Further, the details of the use cases are retrieved from the cluster blueprint developed by the cluster's partners, which will provide a comprehensive understanding of the application of 6G technologies in each respective use case. This step is important in positioning KVs with different natures of use or the targeted verticals. In the 6G-VERSUS project, targeted verticals expand to include renewable energy systems, life-saving environmental preservation, environmental monitoring & safety, smart agriculture and water management. The identified key values will assess the core benefits and sustainability impacts that 6G technology advancements are planned to achieve in the mentioned verticals.

3.2.3 Step 03: Defining KVIs

Once KVs are finalised with cluster partners, KVIs are defined to measure each key value effectively. The KVIs are the tools used to measure the broader environmental, societal and economic impacts of 6G-VERSUS solutions. Appropriately, several dimensions are considered with respect to the given use case or target vertical, such as Events, Resources, Processes and Coverage [IMB_D11].

- **Events**: The number of specific events related to the use case (decrease the occurrence or severity of adverse events or enhance the frequency and impact of positive events).
- **Resources**: The quantity or state quality of specific resources relevant to the use case.
- **Processes**: Assess the efficiency of a process associated with the use case that influences the realisation of a key societal value.
  Coverage: Measuring how a solution impacts a key societal value in terms of geographic or population coverage.

Formalized KVIs will identify KV Enablers (KVEs). KVEs explain how technological development contributes to generating the KV. Further, KVE will directly link to the KPIs of the use case or the targeted vertical. In the KVI formulation stage, data collection methodologies such as questionnaires, surveys, interviews, and focused group discussions are required to be finalised. The above process may require continuous iterations, interactions with use case or targeted vertical leaders and stakeholders and co-creation with use case and targeted vertical partners to refine and establish KVIs.

At the end of this step, key values are linked to the technical performance of the project through KVE and KPIs. The objective of identifying the connection is to translate the qualitative nature of key values into quantifiable measures. This connection explains how the technical advancement of 6G-VERSUS contributes to delivering the identified key values of each use case.

3.2.4 Step 04: Validation Process

The KVIs validation occurs during planned trials in collaboration with stakeholders. These trials test whether defined KVIs are feasible and practical for real-world application. Feedback from these trials informs iterative refinements, ensuring that KVIs remain relevant and actionable. Established KVIs are evaluated with the initial trials, and information is recorded about the feasibility and practicability of existing KVIs.

3.2.5 Step 05: Execution Process

The establishment of the KVI framework provides the foundation for the execution process. The framework is flexible to allow upcoming refinements required in the execution stage. At the execution stage, refined KVIs are applied and measured objectively to determine the impact of 6G technologies on the environment, society and economy aspects of each Use Case or Targeted Verticals.

In this process, preplanned data collection methodologies and evaluation techniques will be applied to assess the KVIs. As the project conducts multiple trials, suggested data collection methodologies and evaluation techniques will be changed depending on the effectiveness of the method and feedback from stakeholders and cluster partners.

3.2.6 Step 06: Monitoring and Refinement

This step monitors and recognises the refinements required in identified KVs, KVIs and evaluation methodologies in the framework. In the real-world use case testing, the framework will be revisited and updated to reflect the insights and findings of experiments. The continuous refinements will be adjusted to the framework to ensure precise measurement of sustainability impacts. The figure below illustrates the step-by-step process for identifying KVs, defining KVIs, validating them during trials, and measuring their impact during execution.

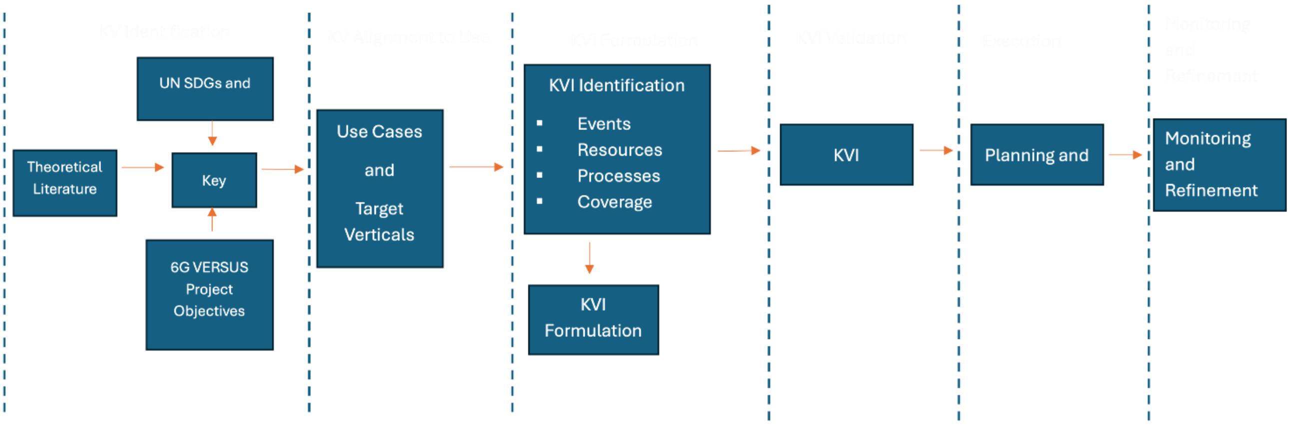

Figure 9: Process for identifying, defining, and validating the KVIs.

3.3 Procedures Before, During and After Trials

The table below outlines the procedures and verification process to be followed during and after the trials.

	Before the Trials	During the Trials	After the Trials
Responsibility	Respective WP leader (CES), Tasks Leaders (UBRAD, INC), Cluster Partners (all six clusters)	Respective WP Leaders, Task Leaders, Cluster Partners	Respective WP Leaders, Task Leaders, Cluster Partners
Objective	Setting the foundation for effective KVI assessment	Monitor and evaluation of KVIs	Assess the overall impact, draw conclusions and plan for improvement
Procedures	▪ Identification of KVs: Define the key sustainability values of the project. ▪ Formulation of KVIs: Develop qualitative or quantitative measures for each KV. Set targets and baseline metrics for KVIs. ▪ Stakeholder Identification and Cluster discussions: Refineme nt of KVs and KVIs with the consultation of stakeholders and cluster partners.	▪ Monitor and collection of KV and KVI data: Collect the required data related to the KVIs and monitor for any refinements of KVIs or KVI measurements. ▪ Engaging with clusters and stakeholders. ▪ Collect feedback from different stakeholders and cluster partners. ▪ Identification of refinements: Discuss the required changes to the KVIs and measurement of KVIs.	▪ Analyse the KVI results: Compare and contrast the outcomes against the baseline or expectations. ▪ Cross-reference with KPIs: Determination of technical performance contribution to the accomplishment of KVIs. ▪ Documentation of results: Report the results of KVIs (positive and negative impacts on sustainability dimensions) ▪ Recommend enhancements: Suggest

	▪ Design and develop the data collection methods and evaluation methods: Development of stakeholder-specific data collection and evaluation strategies. ▪ Trial designing and planning ▪ Designing the methodologies, timeframes, data collection methods, etc. ▪ Ethical and risk assessment: Ensuring all ethical clearance for the data collection and comprehensive risk assessment.		improvements for the KVI framework, technology and policy.

Table 10: Procedures for KVI identification and verification process

3.4 Guideline for UC Partners to Develop their KVI Strategy

The role of the use case partners in designing, developing, and accessing KVIs is crucial to the effective measurement of the sustainability impact of 6G technology. Their involvement from the inception to the final stage of KVI evaluation will ensure that the indicators are contextually relevant, stakeholder-informed, and accurately reflect the real-world value outcomes. Since most UC partners are familiar with the project's technical aspects but less familiar with sustainability and societal impact aspects, the following guidelines will help them approach the KVIs with clarity, relevance, and purpose. It narrows the gap between technical assessment and value-driven assessment, ensuring UC partners meaningfully contribute to the sustainability assessment of the 6G-VERSUS.

Identification and validation of KVs: UC partners possess the contextual details of the use case, such as objectives and ambitions, motivations, pain points of the cluster, stakeholder requirements and potential technological developments. This foundational understanding can be leveraged to identify the KVs of the project.
Transforming KVs into the measurable KVIs: UC partners translate KVs into measurable KVIs by focusing on sustainability impacts. Crafting effective KVIs relies heavily on a fundamental understanding of the potential sustainability impacts, both positive and negative, of the technology being deployed. This ensures that the value-driven goals of the project are not only well-defined but also empirically verifiable.
Linking of KVIs to UC activities and expected Outcomes: The linkage between KVIs, UC activities, and expected outcomes is established through a collaborative process involving all stakeholders. Drawing on their contextual expertise, UC partners play a key role in identifying meaningful and achievable indicators that align with specific technological objectives and reflect the realistic value impact of the interventions within each use case.

3.5 KVIs per Cluster

In D2.1, identified KVIs for each cluster in 6G-VERSUS across all the sustainability dimensions are presented in section 4.1.2. For completeness and ease of reference, the same tables are reproduced here [6GV-D21].

KVI IDs	KVs	KVIs	Description
KVI_001 Environmental	Reduction of Carbon footprint	Increase in utilisation of RES (lead by avoiding of RES curtailment	The Use Case will avoid RES curtailment by 15% which will increase the use of RES for energy generation.
KVI_002 Environmental	Reduction of Carbon footprint	Increase in use of RES for energy generation.	Due to the Faster and efficient integration of renewable energy to the grid, its capacity increase to absorb more renewable energy. Therefore, more RES will be used in energy generation.
KVI_003 Environmental	Reduced disruptions for ecosystems and landscapes.	Reduced Land Disturbance and Habitat Disruptions.	The use case replaces the need for wired communication with mobile communication by connecting the monitoring and control functionality of renewable energy resources to the control centre of the grid via 5G. Therefore, the investment on wired connection is avoided.
KVI_004_ Societal & Cultural	Quality of life	Increase in the grid stability (Reduced reaction time for grid frequency recovery)	How do the advanced features of 6G technology contribute to increase grid stability.
KVI_005_ Societal & Cultural	Digital inclusion	Increase in access to 6G network for remote communities.	As 6G mobile network is extended into remote areas to connect RES to grid, this will enable communities in those remote areas to access 6G connections.
KVI_006_ Societal & Cultural	Quality of life	Increase in affordable and stable Renewable Energy supply.	Increase in the annual operational uptime percentage of RES support affordable and stable energy supply to the community.
KVI_007_ Societal & Cultural	Market participation	Increased participation in energy generations. (Wind, Solar and Microgrids)	Due to the Faster and efficient integration of renewable energy to the grid, its capacity increase to absorb more renewable energy which will provide an opportunity for new RES owners to participate in energy generation.
KVI_008 Economic	Cost Reduction	Percentage reduction of Operational Cost (OPEX).	As the Use Case develops a controller which studies to how to use the assets in best possible where will lead to reduction in operation cost.
KVI_009 Economic	Financial viability	Optimising revenue management of RES Owners	How do the advanced features of 6G technology contribute to optimise revenue management of RES Owners?
KVI_010 Economic	Cost Reduction	Reduction of Investment cost due to shifting for cable connection to mobile connection.	The Use case replace cable networks with mobile network by to connect renewable energy resources to the

			grid. Therefore, the investment on cable connection is avoided.

Table 11: KVIs for Bulgarian cluster

KVI IDs	KVs	KVIs	Description
KVI_001 Environmental	Reduced Ecological footprint	Efficient resource allocation (optimal use of resources)	How do advanced features of 6G technology contribute to improve Energy efficient application integration? (Achieve a 25% improvement in network energy efficiency)
KVI_002 Environmental	Improved ecosystem conservation	Prevention of wildfire spread.	How do advanced features of 6G technology contribute to prevention of wildfire spread?
KVI_003 Environmental	Reduced Ecological footprint	Reduce of the disruptive use of the technology	How do advanced features of 6G technology contribute to reduce traditional means (e.g. Helicopters) for incident detection?
KVI_004 Societal & Cultural	Quality of Life	Reduced emergency response times	The decrease in the time taken to respond to emergencies, which is a critical metric for assessing the effectiveness of improvements in emergency services
KVI_005 Societal & Cultural	Quality of Life	Reduction of risk of injuries and casualties sustained in rescue operations for affected civilians and first respondents.	How do the new 6G-enabled tools (e.g., robotics, AI detection) helped reduce risks or injuries for responders and civilians?
KVI_006 Societal & Cultural	Quality of Life	Success rate in operational efficiency	How the extendable coverage of 6G via relay UAVs and the deployments of on-site network/compute resources contributes to a faster and more efficient (i) surveillance of the wildfire disaster and (ii) reduced health risks for first responders and civilians
KVI_007 Societal & Cultural	Quality of Life	Better emergency preparedness	Advanced features of 6G contribute to improved emergency preparedness by enabling flexible and rapidly deployable network to support dynamic, on-demand responses during an emergency.
KVI_008 Economic	Cost Reduction	Operational Cost saving	How do advanced features of 6G technology contribute to reduce operational cost of rescue operations? (reducing fully manual surveillance ultimately lowering overall operational expenses)

Table 12: KVIs for Greek cluster

KVI IDs	KVs	KVIs	Description
KVI_001 Environmental	Reduction of Carbon footprint	Increase in Renewable energy sources.	Transition of energy systems towards renewable energy sources.
KVI_002 Environmental	Reduction of Carbon footprint	Reduction of the energy consumption in remote sites.	How do advance features of 6G technology contributes to reduce energy consumption in remote sites.
KVI_003 Environmental	Energy efficiency	Reducing the energy waste	How do advanced features of 6G technology contribute to reducing energy waste?
KVI_004_ Societal & Cultural	Digital inclusion	Availability of 6G connectivity in remote areas.	How do the advanced features of 6G technology expand the 6G connectivity to remote areas?
KVI_005_ Societal & Cultural	Digital inequality	Reduced digital inequality	How does the service provide all with digital connectivity regardless of living area, economic status, gender or ethnicity?
KVI_006_ Societal & Cultural	Quality of life	Enhance the resilience of 4G/5G/6G networks	Reliable continuous 6G connectivity during electricity distribution disruptions.
KVI_007_ Societal & Cultural	Quality of life	Safety in hard-to-reach areas	How to provide access in remote areas e.g., in the cases of emergencies?
KVI_008 Economic	Cost Reduction	Reduction of Operating cost	How do advanced features of 6G technology contribute to reduce operational cost?
KVI_009 Economic	Cost Reduction	Reduction of Capital expenditure.	How do the advanced features of 6G technology contribute to reduction of capital expenditure?
KVI_010 Economic	Financial viability	Increase in revenue of businesses in remote areas.	How do advanced features of 6G technology contribute to Increase in revenue of businesses in remote areas? (reindeer owners, tourism related business in arctic region)

Table 13: KVIs for Finnish cluster

KVI IDs	KVs	KVIs	Description
KVI_001 Environmental	Resource efficiency	Increase in Water Saving	How do advanced 6G features (real-time IoT sensing, AI-driven analytics) enable smarter irrigation for greater water savings in cities facing climate and demand stress?
KVI_002 Environmental	Decarbonizatio n	Carbon Emissions Avoided	How can next-gen 6G connectivity and automation help us reduce the CO_2 footprint in green infrastructure by

			optimising resource use and operations?
KVI_003 Environmental	Cooling	Cooling through Greenspace Watering	In what ways can ultra-low-latency 6G-powered monitoring optimise plant watering to reduce urban heatwaves and improve resilience?
KVI_004 Environmental	Circularity and hardware footprint reduction	Electronic Waste Footprint Reduction	How does virtualizing infrastructure with advanced 6G and AI reduce the need for physical devices, preventing e-waste and supporting a circular economy?
KVI_005 Environmental	Longevity/ Durability	Battery life extension	How can energy-efficient 6G architectures extend the operational life of IoT sensors, reducing waste and maintenance needs?
KVI_006 Environmental	Community safety	Reduction in pollution incidents	How can integrated 6G data streams and stakeholder collaboration help us intercept urban pollutants before they harm waterways?
KVI_007 Societal & Cultural	Transparency	Trust in AI alert management	How do secure, transparent 6G-enabled AI alert systems foster greater trust for users managing critical infrastructure?
KVI_008 Societal & Cultural	Quality of life	Preservation of greenspaces	How do 6G-driven digital engagement and environmental monitoring help safeguard greenspaces and support community stewardship?
KVI_009 Economic	Value Creation	ROI from sustainability investments	In what ways do 6G-enabled efficiencies transform environmental actions into measurable financial returns for urban services?

Table 14: KVIs for French cluster

KVI ID	KV	KVIs	KVI Description
KVI_01 Environmental	Lower operational footprint	Energy consumption of infrastructure	How do the advanced energy management features of 6G technology help optimise the energy consumption of critical port infrastructure like smart lighting, contributing to sustainability and operational efficiency?
KVI_02 Environmental	Sustainable innovation	Proportion of renewable energy use (EV charging)	How does 6G-enabled monitoring and management enable increased use of renewable energy in EV charging, reducing the port's carbon footprint and supporting green transport initiatives?
KVI_03 Environmental	Sustainable innovation	Proportion of renewable energy use (smart lightning)	How can precise energy source tracking through 6G technology demonstrate the contribution of renewables powering smart lighting systems at the port, reinforcing clean energy adoption?

KVI_04 Environmental	Reduced pollution impact	Pollution & source identification time	In what ways does 6G technology enable advanced real-time detection and prevention of abnormal events at the port to enhance environmental protection?
KVI_05 Societal	Improved safety and well-being	Safety incident detection & response time	In what ways does 6G technology enable advanced real-time detection and prevention of abnormal events at the port to enhance safety, environmental protection, and operational resilience?
KVI_06 Societal	Inclusive environmental protection	Environmental monitoring coverage	How does the deployment of 6G-enabled sensors across the operational land port area improve environmental monitoring and inform decision-making for sustainability and compliance?
KVI_07 Societal	Workforce resilience and inclusion	Workforce upskilling rate	How does 6G-driven digital training and upskilling of port administration staff foster a competent workforce capable of managing advanced technologies and supporting digital transformation?
KVI_08 Cultural	Community trust and engagement	Stakeholder engagement activities	How do 6G-enabled digital platforms and communication tools facilitate inclusive, diverse, and effective stakeholder engagement across port operations and community interactions?

Table 15: KVIs for Portuguese cluster

KVI IDs	KVs	KVIs	Description
KVI_001 Environmental	Reduction in irrigation use	Water Usage Efficiency	How can advanced water management enabled by 6G connectivity support farmers in using less water while maintaining or increasing crop yields?
KVI_002 Societal & Cultural	Increase in high-tech rural jobs	Rural Workforce Retention	How might real-time information and support, delivered via 6G help make rural agricultural work more attractive and sustainable for local populations?
KVI_003 Societal & Cultural	Inclusivity	Inclusiveness in Access	In what ways can 6G-enabled solutions reduce barriers and open up access to digital tools and markets for smallholder farmers and rural communities?
KVI_004 Societal & Cultural	Increased trust/ empowerment	Community Health Engagement	How does 6G-facilitated information sharing and community engagement enhance awareness and action on health issues among agricultural communities?
KVI_005 Societal & Cultural	User trust	Cultural Acceptance Rate	How do advanced 6G features foster trust and acceptance of digital innovations within traditional farming communities?
KVI_006 Economic	Reduction in operational costs	Operational cost savings	In what ways can data-driven decision-making supported by 6G reduce operational costs for growers?

KVI_007 Economic	Increase in yield	Productivity Increase	How do real-time analytics and automation enabled by 6G contribute to increased crop productivity?
KVI_008 Societal & Cultural	Trust	Technology adoption rate	How does 6G-powered connectivity accelerate the adoption of modern farming technologies among local communities?

Table 16: KVIs for Spanish cluster

4 6G-VERSUS KPI Validation Framework

4.1 KPI Evaluation Frameworks and Terminology

In the context of KPI validation, the 6G-VERSUS project will leverage best practices from other reference EU-funded projects, such as 6G-SANDBOX and 6G-PATH, among others. However, it must be noted that a less tightly coupled approach will be followed in certain aspects because of the specific focus on large-scale trials within 6G-VERSUS and of the heterogeneity of the different testbeds and vertical UCs. For that reason, the definition of the KPI validation framework will frequently not be stated as mandatory but as recommended where applicable. The 5G!DRONES project [5GD-D44] has been a key reference for the KPI validation framework definition.

In the following subsections, several concepts are inherited from related projects; however, definitions will be reproduced with minor adaptations to make the present deliverable as self-contained as possible.

4.1.1 Experiment and Testing Terminology

Inheriting from the 6G-SANDBOX [6GS-D51] project terminology, throughout the document, the concepts of testing and experimentation will be extensively used. Therefore, it is crucial to explicitly define the precise meanings of these terms in this section.

The goal of an experiment is to explore the performance or functionality of an entity. A specific target is not formally established at the outset, but the goal is to explore the limits and capabilities of the target entity.

The goal of a test is to validate the performance or functionality of an entity. There are specific goals or thresholds that the test seeks to reach or prove that performance exceeds or falls short of. Subsequently, a test can be evaluated as passed or failed based on its outcome relative to the targets or thresholds. From this, an experiment can be seen as a more loosely defined methodology than a test.

As part of the 6G-VERSUS project, each cluster will develop and conduct trials exercising different use cases, which will be validated under a series of tests and experiments. The validation and evaluation framework will define the applicable measurements and test conditions in a clear, detailed, and extendable manner, allowing different work packages to predictably infer, design, and develop appropriate smaller-scale and unit test environments. This will ensure readiness to meet KPI targets with confidence when conducting larger-scale trial experimentation.

4.1.2 Testing Methodology

The methodology recommended by default in 6G-VERSUS is inherited from 5G!Drones project, but also encouraging, where applicable, the use of repeatable automation practices from the 6G-SANDBOX project, which in turn also had contributions from earlier work from the 5GENESIS and 5G-VINNI projects. Overall verification objectives are stated and then related test cases are defined to specify how those objectives will be assessed.

- **Test Cases (TCs)** have thus an associated objective, expected outputs, and procedures for evaluating specific KPIs. They include the metrics to be measured, the calculation methods, and the test sequence steps. Test Cases are written in a reusable format to support replication across different platforms.
- **Scenarios** define the radio conditions under which the experiment is to be run (e.g., mobility, interference, congestion). These are meant to be generic and adaptable to the configuration interfaces of different radio equipment, avoiding platform-specific details.

The combination of a Test Case and a Scenario defines a unique **Experiment**.

The methodology supports extensibility and allows test execution using different tools or platforms, provided that KPIs/KVIs are collected and stored in a consistent format, typically using time-series databases with associated metadata.

As earlier stated, within 6G-VERSUS, there will be additional flexibility to accommodate deviations from this methodology in aspects like data collection and format for internal usage within the testbeds and clusters. However, it is expected to have a unified format when it turns to data storage for external exposure.

4.1.3 Test and Experiment Automation from 6G-SANDBOX

Those parties within 6G-VERSUS that have decided to leverage the automation from 6G-SANDBOX are benefiting from the execution of the trials. Where needed, they may benefit from several features and capabilities, which are summarised next and described in greater detail in section 3 within [6GS-D51].

In the 6G-SANDBOX project, test and experiment automation is supported through an optional, yet powerful, framework built around OpenTAP and the Keysight Test Automation Cloud (TAC). This automation infrastructure is designed to assist experimenters in defining, managing, executing, and replicating test cases in a consistent and streamlined manner across distributed testbeds. However, its use is not mandatory and should be viewed as a facilitation tool rather than a requirement, i.e., something especially relevant for future projects where independent testbeds may adopt different approaches.

At the core of this automation system is OpenTAP, a modular, plugin-based test engine that allows the sequencing of discrete actions (test steps) to interact with devices, tools, or software under test. These test steps are organised into test plans, which represent concrete implementations of test cases as defined in the 6G-SANDBOX methodology. Test plans can include configuration steps, test traffic generation, statistics extraction, and even post-processing if supported by the connected tools.

To scale and manage the execution of these plans, the TAC offers features such as parameter presets, scheduling, real-time monitoring, and integration with CI/CD pipelines via a northbound REST API. Test campaigns, which are collections of one or more test plans, allow for more complex execution flows, including conditional logic and parameter sweeps. Campaigns can be triggered manually through the TAC interface, scheduled periodically, or invoked remotely through the API, enabling flexible and repeatable testing workflows.

The integration of the automation system with the Trial Network Life Cycle Manager (TNLCM) enables a seamless connection between the testing logic and the dynamic configuration of the underlying test infrastructure. The OpenTAP runner component acts as a lightweight bridge between the cloud orchestration and the physical or virtual components being tested.

Despite these advanced capabilities, the 6G-SANDBOX project emphasises that this framework is optional. Future projects building on this methodology, especially those involving heterogeneous testbeds across different organisations or regions, may choose to implement their own automation systems or manage test execution manually. As long as the core principles of the 6G-SANDBOX testing methodology are respected (i.e., the separation of Test Cases, Scenarios, and Slices), alternative execution strategies are fully compatible. This ensures that experimentation remains flexible and adaptable, while preserving the ability to generate reproducible, comparable KPI/KVI results across diverse environments.

Alternative test and experiment automation capabilities are also acceptable if they benefit different clusters and use cases. However, note that automation is not a goal but a potential resource within the scope of the 6G-VERSUS project.

4.2 Measurement and Monitoring Tools

4.2.1 Bulgarian Pilot Measurement Tools

The data collection and monitoring tools used in the Bulgarian testbed of A1 are:

- **VIAVI NITRO GEO platform**[1]: Supports multiple generations of mobile technologies from Nokia RAN vendor, providing continuous monitoring and analysis of network events.
- **Nokia's NetAct platform**[2]: Serves as both a network management and element management system, offering comprehensive operation and maintenance capabilities for core, radio, and transport networks.
- **VIAVI xSight platform**[3]: Real-time intelligence platform offers advanced analytics capabilities to capture and analyse relationships between customers, applications, services, and networks in real-time.
- **Amdocs' Helix Performance Management platform**[4]: Used for monitoring, optimisation and troubleshooting of network and service performance. It collects and processes data for use in performance reporting and resource planning in all network domains (Radio, Core and Transport), helping to increase and enhance the performance.

KPI_004 and KPI_008 will be measured using the NetAct platform, while KPI_010 and the performance data for KPI_011 will be obtained from the Helix platform. KPI_009 will be measured using the TEMS drive test tool.

4.2.2 Finnish Pilot Measurement Tools

The measurement tools to be utilised in the Finnish cluster include the following:

- **Keysight IxChariot**[5]: Measures network performance for different types of application-level traffic patterns.
- **Keysight Hawkeye**[6]: Network performance measurement system (e.g., latency, optional if required measurements can't be done with iXChariot or Qosium).
- **Keysight RICtest**[7]: Solution to test near real-time RIC exercising its interfaces to multiple network elements and to non-real-time RIC.
- **Keysight NEMO**[8]: Radio Frequency (RF) related measurements done with software (SW) installed in a User Equipment (UE).
- **Kaitotek Qosium**[9]**:** Passive QoS measurement system for network performance (for measurements like one-way latency, jitter, data bandwidth, etc.).
- **Energy analyzer EM111**[10]: Carlo-Gavazzi, 1-phase load monitoring with MQTT-broker Mosquitto (software).

[1] https://www.viavisolutions.com/en-us/solutions/nitro-location-intelligence

[2] https://customer.nokia.com/support/s/product2/netact/01t41000004gAuNAAU

[3] https://www.viavisolutions.com/en-us/news-releases/customer-experience-assurance-age-unpredictable-apps-privacy-protection-and-virtualized-networks

[4] https://www.amdocs.com/products-services/performance-management

[5] https://www.keysight.com/gb/en/products/network-test/ixchariot.html

[6] https://www.keysight.com/us/en/products/network-visibility/hawkeye-software-suite.html

[7] https://www.keysight.com/gb/en/library/demos/demo/near-real-time-ric-testing-with-keysight-s-rictest.html

[8] https://www.keysight.com/us/en/products/nemo-wireless-network-solutions.html

[9] https://www.kaitotek.com/qosium

[10] https://www.gavazziautomation.com/en-us/product/EM111DINAV81XS1PFB

- **Netio-PowerBOX**[11]**:** 4KF with MQTT-broker Mosquitto (software).

4.2.3 French Pilot Measurement Tools

The measurement tools to be utilised in the French cluster include:

- **LoadCore**[12]**:** Tests the 5G core through standard interfaces by emulating millions of stateful UE connections. In the French testbed at EURECOM, the scalability and robustness of the 5G core are validated under realistic traffic and mobility scenarios. This enables E2E testing of core functions (Access and Mobility Management Function (AMF), Session Management Function (SMF), User Plane Function (UPF), etc.) and assessment of interoperability with CNFs, xApps, and vertical applications. This tool will be used to benchmark the testbed and will support scenarios 2 and 3.
- **Keysight IxChariot:** Characterises network performance for different application-level traffic patterns. Integrated into the French testbed, it simulates real-world traffic to measure latency, throughput, jitter, and packet loss across 5G service types (Enhanced Mobile Broadband (eMBB), URLLC, Massive Machine Type Communications (mMTC)). It supports evaluation of CNFs, xApps, and verticals under varying network conditions, ensuring service quality and infrastructure reliability. This tool will be used to benchmark the testbed and will support scenarios 2 and 3.
- **Athom 16A Power Monitoring Plug (PG01EU16A)**[13]**:** A reliable solution for monitoring the energy consumption of RedCap devices in 5G and 6G networks. It provides real-time measurements of voltage, current, and power, enabling continuous energy profiling of RedCap devices optimised for low-power IoT applications. Using its integration with Kafka, the plug streams JSON-formatted energy data, which is then processed by Logstash and stored in Elasticsearch for analysis. Kibana provides intuitive graphical dashboards for visualising power consumption trends and patterns. This setup supports energy optimisation and helps extend device battery life in resource-constrained environments (scenario 3).
- **RAN Amarisoft:** The **Amarisoft Web Interface**[14] is a GUI designed for managing network components. Through the web interface, users can directly configure specific RAN parameters. It provides real-time displays of traffic, logs, and overall system status, facilitating efficient troubleshooting. In our context, the most relevant features are the graphical views of connected users, cells, and traffic activity across specific bandwidths and Physical Resource Blocks (PRBs). This tool will support Scenario 3, specifically for RAN configuration aimed at extending the battery life of sensors.
- **Temperature sensor:** Used to assess the cooling effect of green space watering. This device will be used in scenario 3.
- **Waste level sensor:** Used to assess the volume of waste collection. This device will be used in scenario 1.

4.2.4 Greek Pilot Measurement Tools

The following set of tools will be deployed and utilised within the Greek Pilot cluster for measurement and monitoring purposes:

- **Keysight LoadCore**: Solution to test a network core through the standard interfaces, emulating millions of stateful UE connections.
- **Keysight IxChariot**: Solution to characterise network performance for different types of application-level traffic patterns.

[11] https://www.netio-products.com/en/device/powerbox-4kx

[12] https://www.keysight.com/gb/en/product/P8900S/loadcore-core-network-solutions.html

[13] https://www.athom.tech/blank-1/EU-plug

[14] https://tech-academy.amarisoft.com/ltewww.doc

- **Keysight Cloud Peak**[15] [16]: Solution to verify virtualisation infrastructures. Cloud Peak is a web application used to validate the functionality and benchmark the performance of virtualised network infrastructures. By deploying real workloads on top of the Network Functions Virtualisation Infrastructure (NFVI), Cloud Peak provides key insights into the capability to sustain the required Virtual Network Function (VNF) workloads.
- **Keysight Hawkeye**: Monitoring of the network performance tool.
- **Keysight Nemo Handy**[17]**:** Application for KPI measurement on mobile devices.
- **Jetson Stats**[18]: Solution that allows the energy efficiency monitoring, as well as holistic performance monitoring of NVIDIA Jetson devices. For power measurements in particular, it employs the incorporated INA power monitors.
- **pyJoules**[19]: A software toolkit that facilitates energy measurements of INTEL CPUs. It utilises the internal Running Average Power Limit (RAPL) counters of INTEL devices.
- **Scaphandre**[20]**:** Monitors power consumption of Virtual Machines (VMs).
- **Kepler**[21]: Monitors power consumption on VMs.
- **Iperf3**[22]: TCP, UDP IP measurements for bandwidth, loss, and other parameters.
- **Speedtest**[23]: Uplink (UL) and Downlink (DL) throughput/latency measurements of the internet connection towards a publicly available application server.
- **COSMOTE-tools**: A custom Android Application developed by OTE measuring RAN Network performance (Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), Signal to Interference plus Noise Ratio (SINR)) measurements, latency and network throughput (UL/DL traffic).
- **System logs**
- **Application logs**

4.2.5 Portuguese Pilot Measurement Tools

The following tools are available:

- **Keysight LoadCore**: Solution to test a network core through the standard interfaces, emulating millions of stateful UE connections.
- **Keysight IxChariot**: Solution to characterise network performance for different types of application-level traffic patterns.
- **Keysight Cloud Peak**: Solution to verify virtualisation infrastructures. Cloud Peak is a web application used to validate the functionality and benchmark the performance of virtualised network infrastructures. By deploying real workloads on top of the NFVI, Cloud Peak provides key insights into the capability to sustain the required VNF workloads.
- **Keysight RICtest**: Solution to test near real-time RIC exercising its interfaces to multiple network elements and to non-real-time RIC.
- **Keysight WaveJudge**[24]: Keysight's SJ001A WaveJudge Wireless Analyzer Toolset software can be integrated with O-RAN test solutions by importing and decoding collected IQ captures for detailed troubleshooting and full-stack protocol analysis. WaveJudge also processes PCAP

[15] https://www.keysight.com/gb/en/products/network-test/protocol-load-test/cloud-peak.html

[16] https://www.youtube.com/watch?v=8hyZmAAgXRM

[17] https://www.keysight.com/us/en/product/NTH50047B/nemo-handy-handheld-measurement-solution.html

[18] https://github.com/rbonghi/jetson_stats

[19] https://github.com/powerapi-ng/pyJoules

[20] https://hubblo-org.github.io/scaphandre-documentation/index.html

[21] https://sustainable-computing.io/

[22] https://github.com/esnet/iperf

[23] https://www.speedtest.net/

[24] https://www.keysight.com/gb/en/product/SJ001A/wavejudge-5000.html

captured from the O-RAN fronthaul Enhanced Common Public Radio Interface (eCPRI) interface, enabling troubleshooting of interoperability issues between O-RAN vendors. If IQ or fronthaul PCAP traffic can be obtained in the testbeds, the WaveJudge software could be a candidate tool to be utilised.

- **Power Profiler**: Custom electronic circuit designed for the project to measure, analyse, and optimise power consumption in electronic devices. It provides real-time insights into voltage, current, and energy efficiency across hardware and firmware components.
- **Fuel Gauge**: Custom electronic circuit designed for the project to accurately monitor and report battery State-of-Charge (SoC), voltage, current, and remaining capacity in real time. It enables precise power management and battery life optimisation for portable and embedded systems.

Moreover, ITAV is equipped with a **monitoring platform** for measuring and analysing resource usage and energy consumption across physical and virtualised infrastructures, as shown in Figure 10. This is achieved through a multi-layered architecture integrating various tools and technologies for data collection, processing, storage, and visualisation.

At the core of the platform, the Data Collection Adaptation Layer gathers metrics from different sources, including Virtual Machines (VMs) and Containers (resource utilisation and energy consumption are measured using plugins tailored for other environments) and Physical Nodes (energy consumption is tracked using Energy Probes).

The data collection is facilitated by **Telegraf**[25], an agent-based tool that integrates multiple Data Source Plugins to capture information from various systems.

The collected data is streamed in near real-time using an Apache Kafka bus, ensuring scalable and fault-tolerant message distribution. The Data Processing Engine Layer, also leveraging Telegraf, applies transformation and filtering before forwarding the processed data to the storage and visualisation components.

For long-term storage and historical analysis, the platform utilises **InfluxDB**[26], a time-series database, allowing efficient retrieval of historical metrics for trend analysis and system performance evaluation.

Visualisation and monitoring dashboards are provided through **Prometheus**[27] **& Grafana**[28], which enable real-time metric visualisation, alerting, and advanced data querying. External consumers, such as orchestrators and analytics tools, can access monitoring data using:

- Kafka Client Libraries for streaming data.
- REST APIs for querying the monitoring platform.
- InfluxDB Client Libraries for direct database interaction.

The Monitoring Platform Configuration Manager is responsible for configuring and activating monitoring jobs.

[25] https://www.influxdata.com/time-series-platform/telegraf/

[26] https://www.influxdata.com/

[27] https://prometheus.io/

[28] https://grafana.com/

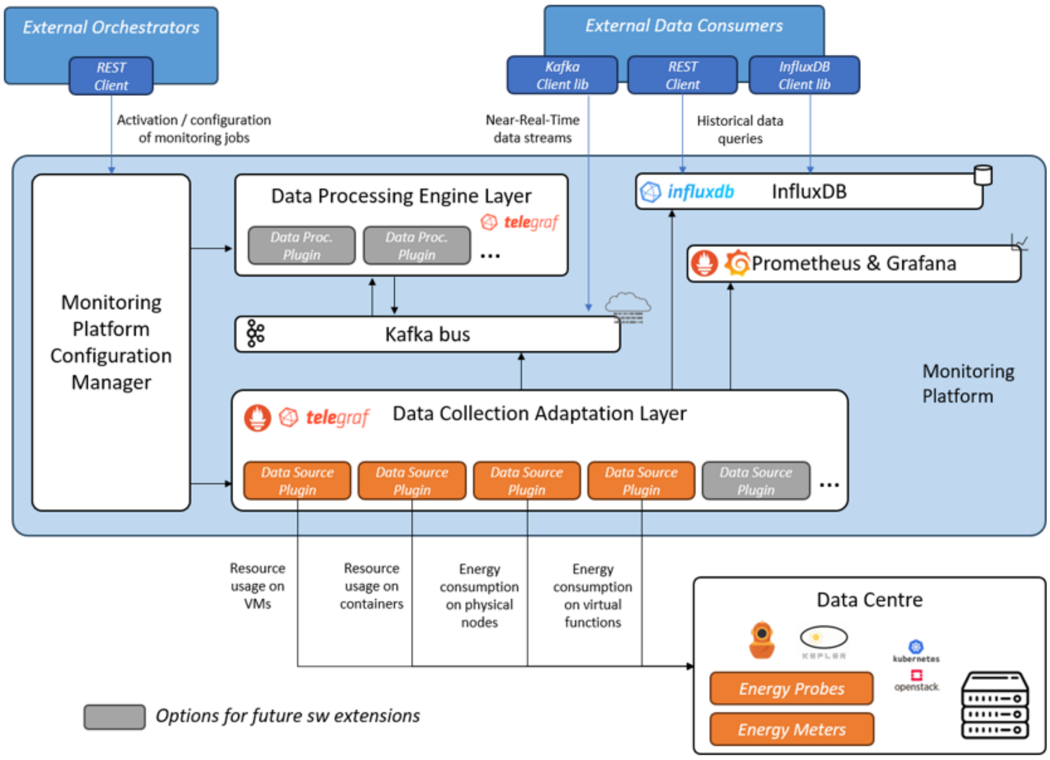

Figure 10: ITAV Monitoring Platform architecture.

4.2.6 Spanish Pilot Measurement Tools

In the Victoria Network testbed, multiple tools are utilised:

- **Network Operations Centre (NOC)**: Manages the network's basic functionalities for data collection and monitoring, which includes a Network Management System (NMS) for monitoring and managing devices, ensuring QoS through automated operations via the Operations Support System (OSS).
- **Nokia Web Element Manager:** Monitors network traces from eNB, gNB, Evolved Packet Core (EPC), and 5G Core.
- **Keysight NEMO:** Nemo Handy Handheld Measurement Solution, Nemo Outdoor 5G NR Drive Test Solution, and Nemo Analyze Drive Test Post Processing Solution facilitate air interface parameter measurement and troubleshooting in the mobile network.
- 6G-SANDBOX **Experiment Lifecycle Manager**[29]**:** Enable complete control over experiments, including replicability and resource management.

In addition to monitoring network connections, energy-related aspects will also be analysed via **Nemo Handy** and **Keysight power analysers**[30].

Testing in a lab environment using these tools will help profile and characterise the behaviour of applications, as well as benchmark software metrics available during field trials. These capabilities will be applied not only to support the Spanish trial but also to support reproducing selected experiments from other clusters for energy profiling purposes. In the Water Management Plant scenario, the Telefonica commercial network and related monitoring capabilities are in place.

[29] https://github.com/6G-SANDBOX/TNLCM

[30] https://www.keysight.com/us/en/product/PA2203A/integravision-power-analyzer-4-channels-3-phase-ac.html

4.3 Common Network KPIs and Metrics

It is expected to try to have a definition of a common set of test cases. In general terms, this will be a minimum common target and parameters, but not necessarily in implementation tools.

#	Network KPI	Description	Relevant metrics and tools
K0.1	Coverage	5G network coverage for the whole route that the connected nodes follow during the trials.	• Reference Signal Received Power (RSRP): Indicates the average power received from a single reference signal. • Reference Signal Received Quality (RSRQ): Indicates the quality of the received signal as a ratio of power and interference. • Received Signal Strength Indicator (RSSI): Indicates the power of the received radio signal. • Signal to Interference plus Noise Ratio (SINR): Indicates the theoretical upper bounds of channel capacity.
		Excellent -44 dBm > RSRP > -85 dBm Good -85 dBm ≥ RSRP > -105 dBm Fair/Bad -105 dBm ≥ RSRP > -140 dBm Excellent -50 dBm > RSSI > -70 dBm Good -70 dBm > RSSI > -80 dBm Fair/Bad -80 dBm ≥ RSSI > -100 dBm	0 dB > RSRQ > -10 dB -10 dB ≥ RSRQ > -14 dB -14 dB ≥ RSRQ > -40 dB 20 dB > SINR > 12 dB 12 dB ≥ SINR > 7 dB 7 dB ≥ SINR > 0 dB
K0.2	Throughput	Throughput is a KPI characterising the data transfer capability of a connectivity link under specific network conditions. It refers to data (payload) successfully transferred within a given time period from a data source to a destination receiver. Multiple factors may impact this parameter.	Based on the direction (uplink/downlink) and the respective payload (TCP or UDP traffic), we consider the following metrics as necessary per use case: • Upload UDP Rate (bps): Max./Avg./Min. values for UDP traffic • Download UDP Rate (bps): Max./Avg./Min. values for UDP traffic • Upload TCP Rate (bps): Max./Avg./Min. values for TCP traffic • Download TCP Rate (bps): Max./Avg./Min. values for TCP traffic
K0.3	Latency	Latency determines the time elapsed between outgoing data from a transmitting source and its receipt at the target destination endpoint.	Depending on the use case, multiple traffic patterns and measurement approaches may be adopted, including using dedicated ad hoc Internet Control Message Protocol (ICMP) traffic or deriving latency information from inline metrics within the actual application traffic. The measurement methods will be elaborated on the respective clusters.

Table 17: Common network KPIs

4.4 Common End-to-End Testing Methodologies

Given the heterogeneous nature of the 6G-VERSUS pilots, there is no strict requirement to enforce the use of the same methodologies. In this section, we will provide a more detailed description of the main tools and methodologies used for traffic and UE-related KPI collection.

To the extent possible, the experiments in the pilots will be designed and repeated to achieve statistical relevance when using these tools.

4.4.1 End-to-End Traffic Generation and Measurement Using IxChariot

IxChariot simulates real-world applications to predict device, system, and network performance under realistic load conditions. The solution consists of an IxChariot Console and Performance Endpoints that are distributed throughout multiple sites and data centres, end-user PCs, and mobile devices. It has the following features:

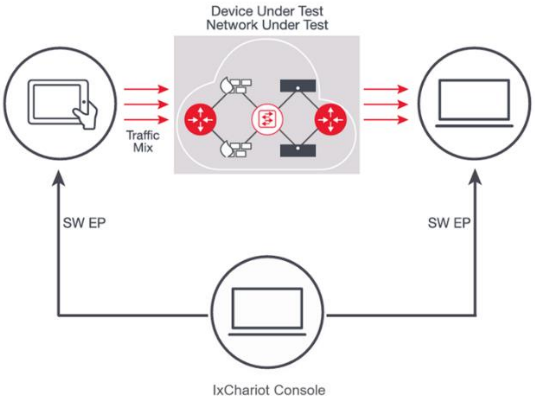

Figure 11: IxChariot deployment.

1. IxChariot Performance Endpoints run on "everything"
   a. Wi-Fi Alliance certified tests
   b. Endpoints are traffic generators – compatible with both IxChariot Server and Desktop editions
   c. Endpoints can be installed on any Windows or Linux PC/server
   d. Endpoints are available on Android and iPhone, installed from Google Play and Apple Store
   e. Endpoints are ready to run on several off-the-shelf small-board PCs
   f. Support for Raspberry Pi B+, BeagleBone Black, and others
   g. Run on virtual and physical (Windows, Mac, Linux) servers
   h. Ixia hardware endpoints available: XRPi2 and XR2000

Figure 12: Supported platforms of IxChariot.

2. Endpoints run "anywhere" – Local Access Network (LAN), Cloud, Public
   a. Supports private endpoints running on devices/servers behind enterprise firewalls
   b. Public points of presence running on any Cloud provider (AWS EC2, Rackspace, Azure)
   c. Endpoints auto-register and maintain heartbeat to the console/registration service Application Emulation

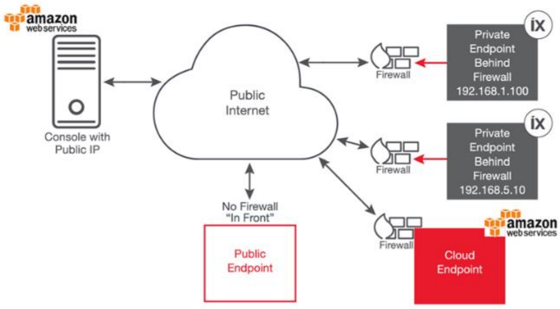

Figure 13: IxChariot endpoint deployments.

3. Interactive real-time statistics
   a. Use IxChariot Server or Desktop console to control endpoints and define test scenarios
   b. Inject real traffic into the network to measure the impact of new applications
   c. View test results in real-time
   d. Measure all critical network metrics, including up/down bitrate, loss, jitter, delay, Voice Mean Opinion Score (MOS), and Video Media Delivery Index (MDI)

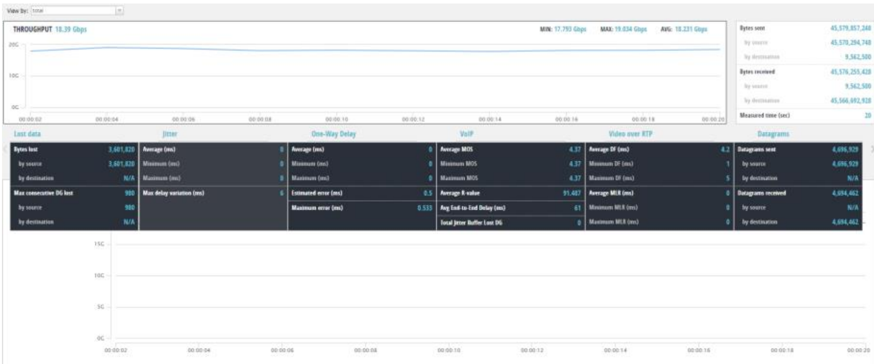

Figure 14: IxChariot statistics.

4. IxChariot Server Edition "IxChariot Anywhere": Server Edition adds the ability to run IxChariot from the cloud, with the main console hosted from one site or data centre and exposed publicly for easy access by multiple teams. Patented Ixia technology enables endpoints to traverse company firewalls, delivering a seamless and elegant solution for assessing network performance between private and cloud networks without compromising security. It has the following advantages:
   a. Cloud-capable console – Runs on EC2 and private cloud
   b. New HTML5-based web interface – Zero client installation; browser-based
   c. Multi-user system – Create and run tests simultaneously
   d. Test from behind firewalls – No configuration required
   e. 100% software endpoints – Runs on "everything"
   f. Setup your own public points – Endpoints on any Virtual Private Server (VPS)/cloud
   g. Centrally manage endpoints – Verify connectivity before use
   h. Measure throughput, jitter, packet loss, end-to-end delay – Per-user metrics
   i. Adaptive Streaming emulation – Like Netflix, Apple TV, YouTube, Youku
   j. Create sophisticated traffic patterns – AppMix, TCP, UDP, and RTP over IPv4 and IPv6
   k. Manage test resources – Built-in resource manager to re-use test scripts
   l. Download IxChariot Server – Available as an open virtual appliance (OVA)
   m. Available on Amazon EC2 marketplace – Bring-your-own-license (BYOL)
   n. Full REST API – Control and integrate with anything

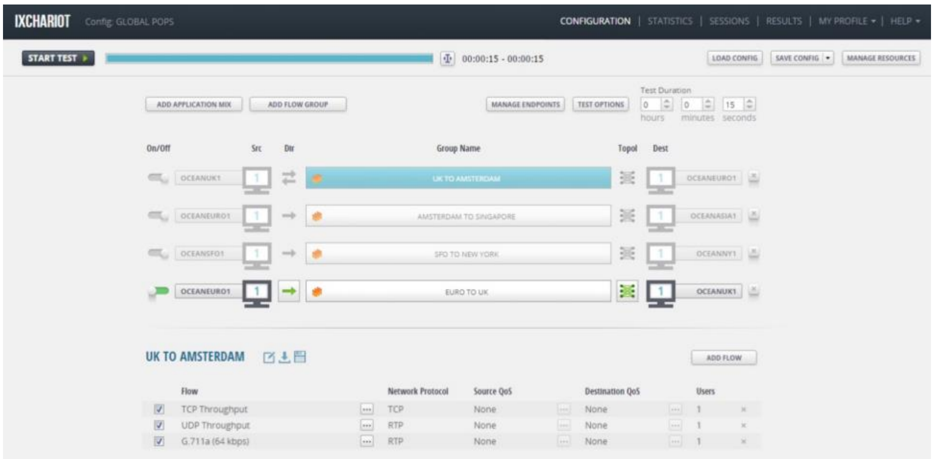

Figure 15: IxChariot traffic pattern configuration.

4.4.2 UE KPI Collection with Nemo and Commercial-Off-The-Shelf UEs

Using Nemo Solutions, it is possible to:

- Measure RF coverage and UE power consumption with commercial, unmodified Android-based devices.
- Compare the RF and power characteristics conveniently using Nemo 5G Device Analytics benchmarking capabilities and reports.
  Match end-user behaviour with test scripts that represent different kinds of traffic profiles, such as Two-Way Active Measurement Protocol (TWAMP), FTP, HTTP, and others.
- Standalone Nemo Handy with commercial Android devices collects RF KPIs.
- Using Nemo Handy and Nemo Diagnostic Module (NDM) allows for collecting chipset diagnostics for detailed troubleshooting.

Figure 16: Nemo deployment in Commercial-Off-The-Shelf (COTS) mobile devices.

4.4.3 UE KPI Collection with Nemo and Custom Firmware UEs

Using UEs with custom Nemo Firmware enables direct collection of both RF KPIs and chipset diagnostic logging, containing low-layer stack information without the need for external additional Hardware (HW) like the NDM.

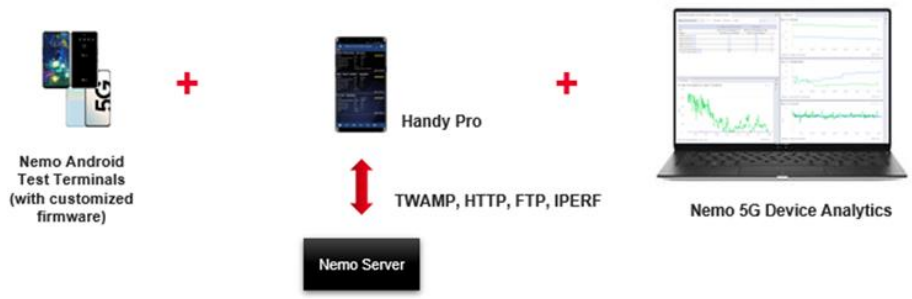

Figure 17: Nemo deployment in mobile devices with custom firmware (FW).

4.5 Bulgarian Cluster. KPI Detailed Verification Plan

4.5.1 Trial Evaluation Objectives

The following high-level objectives have been pursued during the design of the associated test cases.

Objectives
1. Evaluate the network-assisted power grid under normal conditions (Sc.1)
2. Evaluate the network-assisted frequency response under abnormal conditions (Sc.2), considering a large power disturbance
3. Characterise network conditions to enable corrective decision-making

Table 18: Bulgarian cluster trial evaluation objectives

4.5.2 Test Cases

This section identifies the different test cases, which will describe the sequence of executions that will be used for the evaluation of the objectives in the grid in 4 scenarios:

Scenarios
1. 5G communication assessment for DERs – connectivity KPIs and baseline scenario for the vertical with no frequency support from DERs
2. 5G communication for DERs and no NApp
3. 5G communication for DERs +NApp, no AIApp
4. 5G communication for DERs +Napp &AIApp

Table 19: Bulgarian cluster scenarios

The following set of test cases has been identified, whose details are provided in the Annexe.

ID	Title
1.1	Evaluate the Mobile Network Performance in remote areas of DERs
1.2	FFR service from synchronous generators and no DERs participation (baseline scenario)
2.1	PMU-to-V-APP communication over OAI 5G testbed (baseline eMBB setup)
2.2	PMU-to-V-APP communication over the 5G testbed of A1
3.1	PMU-to-V-App/N-App communication over the 5G testbed of A1
4.1	PMU-to-V-App Central (WAM-C) & ML-Ops framework for training of the ML models (AI-App)
4.2	PMU-to-6G-Versus triplet (V-App Central/Regional, N-App, AI-App) focused on the network automation for QoS delivery for a dedicated slice.

Table 20: Bulgarian cluster list of test cases

The KPIs used during the test cases conducted during the trials are defined in deliverable D2.1; however, an enumeration of the list of KPIs is provided here to provide context when mapping test cases to KPIs.

KPI ID	Title
KPI_001	Ultra-Fast Frequency response service time
KPI_002	Improved grid stability with URLL Communication
KPI_003	Network availability
KPI_004	5G SA accessibility success ratio
KPI_005	Average delay in downlink per slice
KPI_006	Average delay in uplink per slice
KPI_007	Packet loss rate
KPI_008	Coverage
KPI_009	5G SA active QoS flow drop ratio
KPI_010	PDU session establishment failure ratio per slice and DNN
KPI_011	Network latency
KPI_012	E2E latency

Table 21: Bulgarian cluster list of KPIs

The following table maps the applicable KPIs to the different test cases defined above.

	Test Case ID						
KPI ID	1.1	1.2	2.1	2.2	3.1	4.1	4.2
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							

Table 22: Mapping of test cases to KPIs

4.6 Finnish Cluster. KPI Detailed Verification Plan

4.6.1 Trial Evaluation Objectives

Adopting the framework suggested by OTE from 5G!Drones project, the definition of tests will be based on the target evaluation of specific trial objectives. Some high-level trial evaluation objectives will be used.

Objectives
1. Evaluate the energy efficiency of the energy grid-free remote arctic base station site.
2. Evaluating the impact of power saving measures at the intelligent remote RU based on energy budgeting scenarios.
3. Validating the correctness of the defined architecture and energy + network components selection for the remote base station.

Table 23: Finnish cluster trial evaluation objectives

4.6.2 Test Cases

This section identifies the different test cases that will describe the sequence of executions used for evaluating the objectives in the grid.

The test cases are listed in the table below, while descriptions of each test case are presented in the Annexe. However, more details will be provided in the future deliverable D5.2 [6GV-D52].

ID	Test Case Name

1	Reliability testing
2	Performance testing
3	Energy efficiency
4	Energy capacity
5	Energy profiling
6	Sustainability assessment
7	AI forecast accuracy assessment

Table 24: Finnish cluster list of test cases

The KPIs for the Finnish cluster are described in deliverable D2.1, but the KPIs are also enumerated next to provide a self-contained context in this document.

KPI ID	Title
KPI_001	Achievable network availability YoY with renewable energy sources
KPI_002	Maximum data rates per user with energy balance awareness in DL and UL
KPI_003	Achievable latency with microwave link in fronthaul
KPI_004	Achievable latency with Non-Terrestrial Networks (NTN) in fronthaul
KPI_005	Coverage energy efficiency
KPI_006	Data energy efficiency
KPI_007	Energy efficiency of the power supply unit
KPI_008	Battery status State-of-Health (SoH)
KPI_009	Battery status SoC
KPI_010	Cost counter/Energy savings
KPI_011	Reduction of CO2 emissions
KPI_012	AI algorithms accuracy
KPI_013	Hydrogen fuel cell usage control
KPI_014	Remote site energy sufficiency

Table 25: Finnish cluster list of KPIs

	Test Cases						
KPI ID	1	2	3	4	5	6	7
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							

Table 26: Mapping of test cases to KPIs (will be further defined in D5.2)

Finnish cluster will develop a remote, energy-grid-free Arctic base station. At the time of writing this document, simulations based on the available 2024 weather data from the Finnish Meteorological Institute are ongoing. Simulation results will be used to define the structure and components of the remote system. Before the system design is complete and the components are selected, it is not possible to fully define the TCs. Therefore, the Finnish cluster TC definitions provided in Annexe are filled for the parts that are known at this phase. TC definitions will be fully defined in deliverable D5.2.

4.7 French Cluster. KPI Detailed Verification Plan

4.7.1 Trial Evaluation Objectives

The high-level trial evaluation objectives in the French cluster are summarised below.

Objectives
1. Demonstrate the capability to provide relevant information on waste volume in the rainwater network, reducing the number of deployed sensors
2. Demonstrate the capability to detect abnormal behaviour in the sewer network
3. Demonstrate the capability to produce a real-time environmental benefit indicator
4. Demonstrate the capability to detect abnormal behaviour such as leakages
5. Demonstrate the capability to save energy by optimising RAN parameters (wake-up cycles in DRX)

Table 27: French cluster trial evaluation objectives

4.7.2 Test Cases

This section outlines the various test cases to be conducted during the trials, aiming to collect associated KPIs and facilitate the evaluation of different pilots. A number of different scenarios will be involved.

Scenarios
Scenario 1: Large-scale sensor virtualisation in the rainwater network to reduce the hardware footprint and increase the ROI
Scenario 2: Large-scale Abnormal sewer network behaviour detection to reduce inundation and pollution risk
Scenario 3: Automatic Detection, watering, and energy optimisation of abnormal watering and leakage in the irrigation network

Table 28: French cluster list of test scenarios

The following set of test cases has been identified. To recall, the French cluster includes two use cases (Use Case 1: Sewer Network Monitoring and Use Case 2: Smart Watering in Green Spaces). We adopted an ID format for the test cases consisting of three numbers x.y.z (where x refers to the use case, y to the scenario, and z to the test case). The following test cases will be described in detail later in the Annexe.

ID	Title
1.1.1	Drainage Site Clusterisation
1.1.2	FFR service from synchronous generators and no DERs participation (baseline scenario)
1.1.3	Prediction of Waste Volume
1.2.1	Detection of Abnormal Behaviour in Sewer Network
2.3.1	Energy Saving. RAN DRX Optimisation
2.3.2	Environmental Benefit Indicator. Green Space

Table 29: French cluster list of test cases

The KPIs for the French cluster are described in Deliverable D2.1, but the KPIs are also enumerated next to keep the context in the current document.

KPI ID	Title
KPI_001	Battery lifetime. eDRX
KPI_002	Battery lifetime. AI DRX
KPI_003	Battery lifetime: Optimal network resource usage
KPI_004	Latency / wake-up delay with eDRX
KPI_005	AI-enhanced RAN: performance: generalisation capabilities
KPI_006	AI-enhanced RAN: reconfiguration time
KPI_007	Alerts filtering
KPI_008	Real-time environmental benefit indicator
KPI_009	Hardware footprint reduction
KPI_010	Latency
KPI_011	Energy efficiency

KPI_012	AI clusterisation and prediction performance
KPI_013	AI clogging detection performance

Table 30: French cluster list of KPIs

	Test Cases					
KPI ID	1.1.1	1.1.2	1.1.3	1.2.1	2.3.1	2.3.2
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						

Table 31: Mapping of test cases to KPIs

4.8 Greek Cluster. KPI Detailed Verification Plan

4.8.1 Trial Evaluation Objectives

The trial objectives are formulated considering the project results and ambitions targeted by the use case, including its requirements fulfilment and KPIs evaluation. The definition of tests will be based on the target evaluation of specific trial objectives. The high-level trial evaluation objectives of the Greek cluster are as follows:

Objectives
1. Verify that the trial site's mobile network is operational and provides the performance characteristics expected, as a fundamental prerequisite for the execution of the trials.
2. Demonstrate operational efficiency gains by exploiting network intelligence.
3. Demonstrate energy reduction gains by exploiting network intelligence.
4. Evaluate UAVs as B5G/6G network relays

Table 32: Greek cluster Trial Evaluation Objectives

4.8.2 Test Cases

For each of the trial objectives, a set of test cases is defined, each test case well describing the specific target, execution scenarios, conditions and expected results. Outlined below are the test cases (details are presented in the Annexe).

ID	Title
1	Network Performance Evaluation
2	UAV relay Peak Data Rate
3	Accuracy Score AI-App Location
4	Latency (Inference Time)
5	Wildfire Detection AI-APP (WF AI-App) Inference Time Evaluation
6	Wildfire Detection -APP Accuracy in Detection and Reporting
7	Wildfire Detection -App End-User Verification
8	Wildfire Detection AI-APP (WF AI-App) Frame per Second (FPS) Impact in Energy Consumption
9	Wildfire Detection AI-APP (WF AI-App) Energy-Efficient Deployment Evaluation
10	Reliable Data transmission between UAVs, robotic dog and control station
11	Localization accuracy of victims and objects
12	Victim Detection Accuracy Evaluation

13	UAV Tracking Latency
14	Video Transmission Latency
15	Evaluate AI-App detection inference time to detect object or victim
16	Energy Efficiency Evaluation

Table 33: Greek cluster list of test cases

KPI ID	Title
KPI_001	Peak user data rate
KPI_002	C-Plane latency in the order of round-trip time
KPI_003	U-plane latency
KPI_004	Delay deviation/jitter
KPI_005	Measure coverage through radio signal parameters (RSRP/RSRQ)
KPI_006	Packet loss rate
KPI_007	Operational network / Availability
KPI_008	Energy measurements on Intel-based CPU systems
KPI_009	Energy measurements on NVIDIA Jetson accelerators
KPI_010	Energy measurements on ARM-based CPU systems
KPI_011	Percentage of accuracy by the AI App location
KPI_012	Wild-fire Detection AI (Average) confidence score
KPI_013	AI-App Location inference latency
KPI_014	Wild-fire Detection AI inference time
KPI_015	Wild-fire Detection AI time to detect
KPI_016	Wild-fire Detection AI time to alert
KPI_017	Wild-fire Detection AI false positive alerts
KPI_018	Wild-fire Detection AI verified successful alerts
KPI_019	Wild-fire Detection AI energy reduction vs processed FPS per GPU
KPI_020	Wild-fire Detection AI energy efficient deployment
KPI_021	Data transmission success rate
KPI_022	Localization error margin
KPI_023	Accuracy rate of detecting and counting victims during operational time
KPI_024	Real-time UAV tracking latency (from data capture to display on the ground control interface during flight operations)
KPI_025	Video transmission latency during live UAV operations
KPI_026	AI-App Inference Latency for Microservice-Based Detection

Table 34: Greek cluster list of KPIs

	Test Cases															
KPI ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1																
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23																
24																
25																
26																

Table 35: Mapping of test cases to KPIs

4.9 Portuguese Cluster. KPI Detailed Verification Plan

4.9.1 Trial Evaluation Objectives

The following high-level objectives have been pursued during the design of the associated test cases.

Objectives
1. Demonstrate energy-efficiency enabled by energy harvesting and intelligent network-aware verticals
2. Measure energy efficiency of the network platform
3.Validate network performance and service agility through advanced slice management
4. Demonstrate the benefits of AI for performant and efficient 6G applications
5. Optimise renewable energy integration with 6G-Apps' vertical infrastructure
6. Assess 6G-Apps' autonomous systems' reliability and precision

Table 36: Portuguese cluster trial evaluation objectives

4.9.2 Test Cases

The following set of test cases has been identified, whose details are provided in the Annexe.

S1, S2, S3: Denote scenarios 1, 2, and 3, respectively.

ID	Title
S1.1	5G NR RedCap-enabled IoT device energy and harvesting efficiency
S1.2	Energy efficiency using AI-App
S1.3	AI-based forecasting performance
S1.4	Maritime object detection

S1.5	Environmental monitoring capabilities
S2.1	Network radio energy optimisation
S2.2	Slice provisioning time
S2.3	Service provisioning time
S2.4	Throughput
S2.5	Autonomous navigation coverage
S3.1	Real-time processing with an AI/ML framework or the AI-App
S3.2	Energy usage optimisation and users' maximisation

Table 37: Portuguese cluster list of test cases

The KPIs used during the test cases conducted during the trials are defined in deliverable D2.1; however, an enumeration of the list of KPIs is provided here to provide context when mapping test cases to KPIs.

KPI ID	Title
KPI_001	Network radio energy consumption
KPI_002	Device radio energy efficiency gain using 5G NR RedCap
KPI_003	Throughput
KPI_004	Slice provisioning time
KPI_005	Service provisioning time
KPI_006	Network energy efficiency
KPI_007	5G NR RedCap IoT device transmission efficiency
KPI_008	Relative Efficiency Index (REI)
KPI_009	Datalogger energy consumption
KPI_010	Energy harvesting efficiency
KPI_011	Mean Absolute Scaled Error (MASE)
KPI_012	Coefficient of Variation RMSE (CV-RMSE)
KPI_013	Forecast Bias
KPI_014	Inferencing speed
KPI_015	Sustainable energy used in EV chargers
KPI_016	Object detection accuracy (water)
KPI_017	Autonomous boat navigation precision
KPI_018	Area Coverage Efficiency
KPI_019	Water surface pollution detection
KPI_020	Water surface minimum object detection size

Table : Portuguese cluster list of KPIs

	Test Case											
KPI ID	S1.1	S1.2	S1.3	S1.4	S1.5	S2.1	S2.2	S2.3	S2.4	S2.5	S3.1	S3.2
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												
16												
17												

18												
19												
20												

Table 38: Mapping of test cases to KPIs

4.10 Spanish Cluster. KPI Detailed Verification Plan

4.10.1 Trial Evaluation Objectives

For each scenario and use case, a set of evaluation objectives is defined to guide the assessment of activities during the trials. These objectives form the basis for the design of test cases that:

- Facilitate the evaluation of the targeted trial goals
- Enable the accurate measurement of the related KPIs
- Provide demonstrations of the relevant aspects to stakeholders in a clear and meaningful way.

The following evaluation objectives have been identified:

Trial Objectives
#1: Remote Driving in Smart Agriculture. Assess the feasibility and performance of remote driving solutions applied to agricultural operations. The objective will focus on connectivity, responsiveness, and reliability in field environments.
#2: Autonomous Driving Missions and Mode Switching. Evaluate the execution of autonomous vehicle missions in agricultural settings and the ability to switch seamlessly between autonomous and remote driving modes, ensuring continuity and safety of operations.
#3: Natural Resource Sampling with Assisted Vehicles. Demonstrate the use of drones and robots for the sampling of natural resources such as crops and water. Multiple technologies will be leveraged, including vehicle telemetry, robotic arms, and multispectral or hyperspectral cameras, to provide comprehensive environmental and crop monitoring.
#4: Field Monitoring with Sensor and Imaging Technologies. Validate the collection and integration of information from distributed field sensors, drones, and AGVs. This includes real-time video monitoring of crops and processes of the water treatment plant, detection of abnormal conditions in crops (e.g., leaf discolouration, pest or fungus presence, irrigation issues), and in the water treatment plant (abnormal operation of water processing), and the incorporation of 5G-enabled connectivity for stand-alone sensors.
#5: Joint Communication and Sensing (JCAS) Capabilities. Explore the use of JCAS to enhance both operational safety and network performance. This includes incident detection (e.g., falls or injuries) to ensure timely worker assistance, as well as network coverage improvement through the deployment of advanced techniques such as passive or active reflectors (e.g., Reconfigurable Intelligent Surfaces, where available).

Table 39: Spanish cluster trial evaluation objectives

Three different scenarios will be considered for the tests:

Trial Scenarios
Scenario 1: La Mayora. Agricultural research farm
Scenario 2: EDAR Roquetas de Mar. Water Treatment Plan
Scenario 3: Lab. Victoria Network / 6G Research Lab in UMA at Ada Byron building

Table 40: Spanish cluster list of test scenarios

4.10.2 Test Cases

The definition of scenarios and use cases will establish the operational context within which the KPIs will be assessed. Each scenario will reflect realistic conditions for deploying the targeted applications and technologies, while the use cases will address the specific functionalities and vertical requirements to be validated.

The relevance of these conditions will lie in ensuring that results are representative, reproducible, and comparable across different clusters and trials. Depending on the type of test, the applications involved, and the traffic generated, conditions will include the number of repetitions, the duration of sessions, bandwidth, packet size, and the use of different protocols. These factors will collectively define the environment in which KPI measurements will be carried out, providing transparency and consistency in the verification process.

The following list of test cases will be further introduced in more detail in the Annexe.

ID	Title
1.1	Optimal water mixing with connected sensors
1.2	Drone-assisted monitoring and sample collection with multipath protocol
1.3	Immersive remote driving and autonomous AGV missions
1.4	Automated surveillance and incident detection with JCAS
2.1	DSS for reliable water treatment
2.2	Open Gateway Quality Assurance
3.1	Lab-based validation of RedCap and energy consumption features
3.2	Energy consumption for model training
3.3	Core network energy profiling

Table 41: Spanish cluster list of test cases

The list of KPIs is initially described in Deliverable D2.1 in more detail. In the following table, though, the KPIs are just enumerated to make this document self-contained.

KPI ID	Title
KPI_01	Throughput. User Experienced Data Rate
KPI_02	Latency. Application layer
KPI_03	Coverage. RSRP and RSRQ
KPI_04	Packet Loss Rate. Application layer
KPI_05	Accessibility. Connection Success
KPI_06	Accessibility. PDU Session Success
KPI_07	Maintainability. PDU Session Drop
KPI_08	AI. Accuracy
KPI_09	AI. Training Energy
KPI_10	JCAS. Channel Estimation Accuracy
KPI_11	JCAS. Detection Accuracy
KPI_12	Sustainability. Water Analysis Accuracy
KPI_13	Sustainability. Water Saving
KPI_14	Sustainability. CO2 reduction
KPI_15	Sustainability. Time Saving
KPI_16	Sustainability. Renewable Powered Device Ratio
KPI_17	Sustainability. UE Energy
KPI_18	Sustainability. Core Network Energy

Table 42: Spanish cluster list of KPIs

Here is the mapping table between the KPIs and the use cases.

KPI ID	Test Case								
	TC_1.1 W.DSS	TC_1.2 DRONE	TC_1.3 AGV	TC_1.4 JCAS	TC_2.1 W.VISION	TC_2.2 OPEN.GW	TC_3.1 UE.EN	TC_3.2 AI.EN	TC_3.3 CORE.EN
1									
2									
3									
4									
5									
6									

7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									

Table 43: Mapping of test cases to KPIs

4.11 Extended Definition of KPIs

For each pilot, an extended definition of test cases is being prepared, including reference to applicable standards, measurement challenges and other applicable aspects.

Due to the additional complexity required to report this work under D5.1, the extended KPI context will be included in the later deliverable D5.2.

5 Conclusions

In this 6G-VERSUS deliverable D5.1, we present the initial trial planning and validation/evaluation framework, laying the foundation for future trial executions across all six clusters. This document shows the methodology for planning, preparation, execution, and reporting of trials, which is applied to each cluster. It also outlines the first iteration result of the KVI and KPI validation frameworks, which enable the assessment of sustainability impacts and technical performance, respectively. These metric definitions ensure methodological consistency across heterogeneous testbeds and use cases. Additionally, by consolidating cluster-specific trial objectives, KPIs, and test case plans, this deliverable builds a clear blueprint and baseline to align technical readiness, stakeholder engagement, and evaluation goals.

Nevertheless, as a first iteration, some content in this document remains preliminary and will be further refined through feedback loops with other work packages, as well as from different sources such as early trial execution results. Moving forward, the outputs of this deliverable will directly feed into D5.2 “6G trials validation and evaluation – Initial” (M18), where the first batch of results from trials will be reported, and into D5.3 “6G trials validation and evaluation – Final” (M36), which will consolidate the final evaluation framework and trial outcomes.

To summarise, D5.1 serves as an essential reference point for the project, ensuring alignment between technical innovations, trial execution, and evaluation objectives. It clarifies the methodology for evaluating and demonstrating the technical performance of 6G technologies, as well as their societal, environmental, and economic value, reflecting the dual ambition of 6G-VERSUS to achieve both technical excellence and sustainability impact.

References

[6GV-D21] 6G-VERSUS, D2.1, “Platform architecture and use case requirements-Initial”, October 2025.

[6GV-D31] 6G-VERSUS, D3.1, “First Integrated Testbeds and Components”, December 2025.

[6GV-D41] 6G-VERSUS, D4.1, “First Version of Integrated Applications and Experiments”, April 2026.

[6GV-D52] 6G-VERSUS, D5.2, “6G Trials Validation and Evaluation – Initial”, June 2026.

[6GS-D51] 6G-SANDBOX, D5.1, “Definition of Evaluation Approach and Test System”, December 2023.

[5GD-D44] 5G!DRONES, D4.4, “Trial Evaluation Report”, November 2022.

[HEXII_D11] HEXA-X-II. D1.1, "Environmental, social and economic drivers and goals for 6G", June 2023.

[HEX_D11] HEXA-X. D1.1, "6G Vision, Use cases and key societal values", February 2021.

[UN_15] United Nations, 2015. UN Sustainable Development Goals. [Online] Available at: https://www.undp.org/sustainable-development-goals, [Accessed 03 August 2025].

[EC_19] European Commission, 2019. European Commission. [Online], Available at: https://ec.europa.eu/commission/presscorner/detail/en/ip_19_6691, [Accessed 03 August 2025].

[IMB_D11] IMAGINE-B5G. D1.1, "Use cases, Ecosystem Analysis and KVIs/KPIs for OC1 and OC2.

[WBM+24] Wikstrom, G. et al., 2024. Key value indicators: A framework for values-driven next-generation ICT solutions. Telecommunications Policy.

Annexe: Test Case Details

This Annexe describes in detail the various test cases identified above as essential to fulfil the trial objectives. The focus is to provide technical information for their execution, such as deployment set-up, the necessary sequence of actions that will take place, specific measurement and calculation methods, conditions and expected results.

Bulgarian Cluster

Test Case #1.1: Evaluate the Mobile Network Performance in remote areas of DERs

Leading Partner: A1	Test Case #1.1: Evaluate the Mobile Network Performance in remote areas of DERs					
Test Case Objective	In this test case, it is verified that the mobile network supporting the Bulgarian use case is operational with the performance characteristics expected.					
Test Case Target Layer or Application	Network Layer – Monitoring Tools					
Test Environment	Field / Sliven region where the 5G mobile base station will be installed					
Test Deployment Setup	All Network Components are up and running Measurement Tools installed NetAct platform					
Network Setup	As described in the section “Testbed description”; 5G SA connectivity with KPI monitoring.					
Test Configuration	**End Device Density**	1	**Mobility Level**	No	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
1. 5G Connectivity is established. 2. Monitoring Tools/Servers are running.						
Test Scenario						
This is the specific set of steps to follow during the test: 1. Perform continuous monitoring of network quality parameters through the measurement tool to evaluate KPIs. 2. Measure the following performance metrics from Sliven region, where is the location of the DERs 3. Repeat measurements in gNB, covering the area within the proximity of the EE wind farm (demo field); Repeat measurements in a walk/drive route of the demo field 4. The experiment is repeated several times for statistical confidence and the Average, Min and Max over the total number of repetitions will be presented as end values.						
Test Variables						
• cell_avail time • 5G SA RRC Connection Establishment SR • 5G SA INITUEMESSAGE SR • 5G SA Initial Context Setup SR • Reference Signal Received Power (RSRP)						
Expected Behaviour / Target KPIs / Values						
Calculated KPI: • KPI_003 Network Availability • KPI_004 SA accessibility • KPI_008 Coverage						

Table 44: Bulgarian Cluster Test Case #1.1

Test Case #1.2: FFR service from synchronous generators and no DERs participation (baseline scenario)

<table>
<tr><td>Leading Partner: UCY</td><td colspan="6">Test Case #1.2: FFR service from synchronous generators and no DERs participation (baseline scenario)</td></tr>
<tr><td>Test Case Objective</td><td colspan="6">This testcase is aimed to establish the baseline scenario, similar to the current state of the art approach when DERs do not participate in the grid ancillary services such as the FFR</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">V-App (WAM-C) Application specific (baseline)</td></tr>
<tr><td>Test Environment</td><td colspan="6">Lab environment using a digital-twin of a power grid including DERs (OPAL-RT technology)</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">OAPL-RT digital twin emulator, power amplifier, PMUs (emulated), V-App Central (WAM-C) deployed as a docker container running on a PC directly connected to the OPAL-RT</td></tr>
<tr><td>Network Setup</td><td colspan="6">No 5G network setup is involved in this test-case (all communication between the power system measuring sensors and the V-App central is wired, via ethernet)</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>6</td><td>Mobility Level</td><td>No</td><td>Has Background Traffic</td><td>No</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">Stable system</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test:
1. at second 5 there is a grid event (large power generator loss)
2. WAM-C application detects the fault and calculates the power imbalance
3. Synchronous generators use their inertial response to mitigate the frequency drop
4. Measure the time it took until the frequency was dumped to normal levels</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">Several other types of grid events can be performed (e.g., loss of a large transmission line, sudden increase of load demand) and observe the same frequency damping effect when only synchronous generator provide the compensation mechanism</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">• KPI_001 Ultra-Fast Frequency Response Service Time
• KPI_002 Grid Stability - frequency time settling (service is provided by the synchronous generators only in up to 2-3 s).</td></tr>
</table>

Table 45: Bulgarian Cluster Test Case #1.2

Test Case #2.1: PMU-to-V-APP communication over OAI 5G testbed (baseline eMBB setup)

Leading Partner: UCY	Test Case: PMU-to-V-APP communication over OAI 5G testbed (baseline eMBB setup)
Test Case Objective	• Validate 5G-based connectivity for synchrophasor data delivery • Evaluate end-to-end latency and reliability of PMU data transport through the 5G testbed • Assess network availability and PDU session establishment success rate
Test Case Target Layer or Application	• End-to-End (PMU → V-APP CENTRAL). • Covers both network layer performance and application-layer data delivery.
Test Environment	Lab environment using the pre-piloting digital of UCY composed of a digital twin of a power grid, a commercial PMU and a 5G OAI-based testbed.
Test Deployment Setup	• PMUs (Model 1133A Power Sentinel).

<table>
<tr><td></td><td colspan="6">• Backhaul Network.
• 5G Gateway.
• USRP B210 SDR / gNB.
• PC running OAI 5G Stack (RAN + CN in Docker).
• PC - CN Central Cloud (V-APP-CENTRAL).</td></tr>
<tr><td>Network Setup</td><td colspan="6">Radio Access Technology (RAT): 5G NR (OAI gNB + USRP B210)
• Operating Band: n78 (3.6 GHz, Band 78)
• Subcarrier Spacing: 30 kHz
• Carrier Bandwidth: 40 MHz channel
• TDD configuration – Single Input Single Output (SISO)
• Power:
    ○ >UE pMax = 20 dBm
    ○ >gNB ssPBCH_BlockPower = -25 dBm</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>1</td><td>Mobility Level</td><td>No</td><td>Has Background Traffic</td><td>No</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">• PMUs connected to backhaul and synchronized.
• OAI 5G Stack (RAN + CN) operational and stable.
• CN Central Cloud (V-APP-CENTRAL) reachable over Ethernet (192.168.10.x).
• PDU session established successfully between UE (PMU traffic endpoint) and CN.</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test:
1. PMUs generate synchrophasor data.
2. Data transmitted via backhaul → 5G Gateway.
3. Data enters OAI 5G RAN via USRP B210 (gNB).
4. OAI 5G Core handles session setup and forwarding.
5. Data forwarded via Ethernet to CN Central Cloud (V-APP-CENTRAL).
6. V-APP receives, decodes, and logs PMU data.</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">NA.</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">• KPI_003 Network Availability
• KPI_007 packet loss
• KPI_010 PDU session establishment failure
• KPI_011 network latency
• KPI_012 e2e latency

Target values:
• Network availability: TBD e.g. ≥ 99% uptime during test period.
• Packet loss: TBD e.g. < 0.01%.
• PDU session establishment failure: TBD e.g. < 2%.
• Network latency:
    ○ Ping 5G Gateway → Central: TBD e.g. ~10–20 ms.
    ○ Ping 5G Gateway → Edge (OAI core): TBD e.g. ~5–10 ms.
• E2E latency (PMU → V-APP CENTRAL): TBD e.g. ≤ 50 ms (target).</td></tr>
</table>

Table 46: Bulgarian Cluster Test Case #2.1

Test Case #2.2: PMU-to-V-APP communication over 5G testbed of A1

Leading Partner: UCY	Test Case: PMU-to-V-APP communication over 5G testbed of A1
Test case objective	• Validate 5G-based connectivity for the MV PMU connected at EE wind farm to the V-App Central (deployed at A1 central cloud).

<table>
<tr><td></td><td colspan="6">Evaluate E2E latency and reliability of PMU data transport through the 5G testbed.
Assess network availability and PDU session establishment success rate.</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">End-to-End (PMU → V-APP CENTRAL).
Covers both network layer performance and application-layer data delivery.</td></tr>
<tr><td>Test Environment</td><td colspan="6">EE plant and A1 5G testbed</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">PMU (Model STERPMU - Rx complete) – 5G ready.
Backhaul Network.
5G Gateway (optional).
5G network of A1 (CN and edge)
CN Central Cloud (V-App-Central).</td></tr>
<tr><td>Network Setup</td><td colspan="6">As described in the main text, more details can be found in Deliverables D2.1 and D3.1</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>1</td><td>Mobility Level</td><td>No</td><td>Has Background Traffic</td><td>No</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">PMUs connected to backhaul and synchronized.
RAN + CN operational and stable.
CN Central Cloud (V-App-Central) is reachable
Uplink traffic from PMU to RAN and CN</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test:
PMUs generate synchrophasor data.
Data transmitted via backhaul → 5G Gateway.
Data enters OAI 5G RAN via USRP B210 (gNB).
OAI 5G Core handles session setup and forwarding.
Data forwarded via Ethernet to CN Central Cloud (V-APP-CENTRAL).
V-APP receives, decodes, and logs PMU data.</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">NA.</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">KPI_003 Network Availability
KPI_004 SA accessibility
KPI_007 packet loss
KPI_010 PDU session establishment failure
KPI_011 network latency
KPI_012 e2e latency.
Target values:
Network availability: TBD e.g. ≥ 99% uptime during test period.
SA accessibility: TBD (e.g.,99.9%)
Packet loss: TBD (e.g. < 1%).
PDU session establishment failure: TBD (e.g. < 1%).
Network latency:
Ping 5G Gateway → Central: TBD e.g. ~10-20 ms.
Ping 5G Gateway → Edge: TBD e.g. ~5-10 ms.
E2E latency (PMU → V-App Central): TBD e.g. ≤ 50 ms (target).</td></tr>
</table>

Table 47: Bulgarian Cluster Test Case #2.2

Test Case #3.1: PMU-to-V-App/N-App communication over 5G testbed of A1

<table>
<tr><td>Leading Partner: UCY</td><td colspan="6">Test Case #3.1: PMU-to-V-App/N-App communication over 5G testbed of A1</td></tr>
<tr><td>Test Case Objective</td><td colspan="6">• Validate the joint monitoring of power grid electrical measurements and the 5G-network related KPIs for the QoS requested by the vertical using a single web-interface.
• Assess N-App capability to communicate with CN NEF via CAPIF.</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">• End-to-End (PMU → V-App Central &N-App).
• Covers both network layer performance and application-layer data delivery.</td></tr>
<tr><td>Test Environment</td><td colspan="6">EE plant and A1 5G testbed</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">TBD</td></tr>
<tr><td>Network Setup</td><td colspan="6">TBD</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>1</td><td>Mobility Level</td><td>No</td><td>Has Background Traffic</td><td>No</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">• MV feeder PMU of EE connected to backhaul of A1 and synchronized.
• RAN + CN operational and stable.
• CN Central Cloud (V-App-Central) reachable.
• CAPIF installed and integrated (hosted at testbed of A1)</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test:
1. PMUs generate synchrophasor data.
2. Data transmitted via backhaul → 5G Gateway.
3. V-App receives, decodes, logs and shows the PMU data in real-time via the web-interface.
4. V-App continuously monitors the real time end-to-end latency and sends adjustment requests to the N-App (in case of deviations from target values).
5. N-app and NEF/CAPIF exposure (performs adjustments if possible).</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">NA.</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">• KPI_003 Network Availability
• KPI_004 SA accessibility
• KPIs_005-6 Avg DL/UL PDCP delay/slice
• KPI_007 packet loss
• KPI_010 PDU session establishment failure
• KPI_011 network latency
• KPI_012 e2e latency

Target values:
• Network availability: TBD e.g. ≥ 99% uptime during test period.
• SA accessibility: TBD (e.g., 99.9%)
• Avg DL PDCP delay/(e.g., <15 ms)
• Avg UL PDCP delay/ (e.g., <15 ms)
• Packet loss: TBD (e.g. < 1%)
• PDU session establishment failure: TBD (e.g. < 1%).
• Network latency (e.g., <50ms)
• E2E latency (PMU → V-App Central): TBD e.g. ≤ 50 ms (target)</td></tr>
</table>

Table 48: Bulgarian Cluster Test Case #3.1

Test Case #4.1: PMU-to-V-App Central (WAM-C) & ML-Ops framework for training of the ML models (AI-App)

<table>
<tr><td>Leading Partner: UCY</td><td colspan="6">Test Case#4.1: PMU-to-V-App Central (WAM-C) & ML-Ops framework for training of the ML models (AI-App).
Testcase focused on the vertical data driven WAM-C service based on emulated network conditions.</td></tr>
<tr><td>Test Case Objective</td><td colspan="6">• Generate a large set of grid conditions (e.g., RES integration level) and grid events (e.g., sudden loss of one or more generation units, sudden load change of different levels etc) for learning the grid frequency behavior when different control actions are taken
• Emulate different network conditions as offline scenarios to be used for the training phase of the AI-assisted QoS management.
• Evaluate the performance of the data-driven WAM-C assisted by the trained AI-Apps.</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">• End-to-End (PMU → V-App Central &AI-App).
• Covers both network layer performance and application-layer data delivery.</td></tr>
<tr><td>Test Environment</td><td colspan="6">UCY pre-piloting testbed (digital twin of the power grid, 5G - OAI-based testbed and ML-Ops)</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">OPAL-RT running the digital-twin model of the power grid including emulated BESS and RES, physical and virtual PMUs, 5G-OAI testbed (CN, radio and MEC), 5G-gateway, PC emulating central cloud, V-App Central (beta version – WAC component) - running on the central cloud, PC-emulating MEC, V-App Regional (beta version, RC component) running on the MEC, ML-Ops running on the central cloud.</td></tr>
<tr><td>Network Setup</td><td colspan="6">TBD</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>6</td><td>Mobility Level</td><td>No</td><td>Has Background Traffic</td><td>No</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">• MV feeder PMU connected in the control hardware in the loop (C-HIL) setup with the OPAL-RT and to the backhaul of the 5G OAI testbed of UCY and synchronized.
• RAN + CN operational and stable.
• CN Central Cloud (V-App-Central) reachable.
• MEC (V-App-Regional) reachable.
• PDU session established successfully between UEs (PMU traffic endpoint) and CN.
• ML-Ops installed and endpoint connections established.</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test:
1. PMUs generate synchrophasor data.
2. Data transmitted via backhaul → 5G Gateway.
3. V-App receives, decodes, logs the PMU data in real-time and sends a subset of measurements (e.g., frequency and Rate of Change of Frequency (RoCoF) only) to the ML-Ops pipeline for further analysis, ML model selection and training.
4. Export pre-trained ML model for the AI-App integration and testing.
5. Test performance of WAC component of the V-App Central, through the V-App Central and & AI-App integration</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">Changing grid dynamics in the digital-twin model by performing and exhaustive set of possible grid events.</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">• KPI_001 Ultra-Fast Frequency Response Service Time
• KPI_002 Grid Stability – frequency time settling
• KPI_007 packet loss*</td></tr>
</table>

• KPI_011 network latency* • KPI_012 e2e latency* • AI-App – KPIs** * These KPIs will be derived from results collected during test case #3.1 and used as input for emulating the network conditions. **Performance metrics (KPIs) for the AI-App will be defined based on the selection of the AI-models. This selection will be based on the analysis of the training data to be generated as part of this testcase and hence cannot be done a-priori.

Table 49: Bulgarian Cluster Test Case #4.1

Test Case #4.2: PMU-to-6G-Versus triplet (V-App Central/Regional, N-App, AI-App) focused on the network automation for QoS delivery for a dedicated slice

<table>
<tr><td>Leading Partner: UCY</td><td colspan="6">Test Case #4.2: PMU-to-6G-Versus triplet (V-App Central/Regional, N-App, AI-App) focused on the network automation for QoS delivery for a dedicated slice</td></tr>
<tr><td>Test Case Objective</td><td colspan="6">• test a dynamic network slice management based on the AI-App advisory which adapts to the real-time monitoring and assessment of the joint network and grid conditions.
• test the E2E dataflow of the integrated triplet (V-Apps, N-App and AI-App)</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">• End-to-End (PMU → V-App Central & N-App & AI-App).
• Covers both network layer performance and application-layer data delivery.</td></tr>
<tr><td>Test Environment</td><td colspan="6">EE and A1 testbed</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">TBD</td></tr>
<tr><td>Network Setup</td><td colspan="6">TBD</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>1</td><td>Mobility Level</td><td>No</td><td>Has Background Traffic</td><td>No</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">• MV feeder PMU connected to the backhaul of the A1 testbed and synchronized.
• RAN + CN operational and stable.
• CN Central Cloud (V-App-Central) reachable.
• MEC (V-App-Regional) reachable.
• PDU session established successfully between UE (PMU traffic endpoint) and CN.
• AI-App installed and integrated with the other two components of the triplet, all running in the central cloud.</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test: TBD</td></tr>
<tr><td colspan="7">Test variables</td></tr>
<tr><td colspan="7">NA</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">KPI_005-6 Avg. DL/UL PDCP delay/slice
KPI_009 Active QoS flow drop
KPI_011 network latency
KPI_012 e2e latency</td></tr>
</table>

Table 50: Bulgarian Cluster Test Case #4.2

Finnish Cluster

Finnish pilot hardware components, namely the energy assets are being simulated and measured for the use case at the time of the D5.1 submission, therefore the final test configurations, target values and setup details will be reported in D5.2.

Test Case #1: Reliability testing

Leading Partner: UOULU	Test Case #1: Reliability testing					
Test Case Objective	Measure the achievable 5G network availability YoY					
Test Case Target Layer or Application	Full stack					
Test Environment	Field					
Test Deployment Setup	Full system					
Network Setup	Will be defined in D5.2					
Test Configuration	**End Device Density**	Will be defined in D5.2	**Mobility Level**	Will be defined in D5.2	**Has Background Traffic**	Will be defined in D5.2
Initial Conditions / Prerequisites						
Will be defined in D5.2						
Test Scenario						
Will be defined in D5.2						
Test Variables						
Will be defined in D5.2						
Expected Behaviour / Target KPIs / Values						
Will be defined in D5.2						

Table 51: Finnish Cluster Test Case #1

Test Case #2: Performance testing

Leading Partner: UOULU	Test Case #2: Performance testing					
Test Case Objective	Measure the technical 5G performance parameters of the End-to-End System					
Test Case Target Layer or Application	End to End					
Test Environment	Field					
Test Deployment Setup	Full system					
Network Setup	Will be defined in D5.2					
Test Configuration	**End Device Density**	Will be defined in D5.2	**Mobility Level**	Will be defined in D5.2	**Has Background Traffic**	Will be defined in D5.2
Initial Conditions / Prerequisites						
Will be defined in D5.2						
Test Scenario						
Will be defined in D5.2						
Test Variables						
Will be defined in D5.2						

Expected Behaviour / Target KPIs / Values
Will be defined in D5.2

Table 52: Finnish Cluster Test Case #2

Test Case #3: Energy efficiency

Leading Partner: UOULU	Test Case #3: Energy efficiency					
Test Case Objective	Measures the energy efficiency of the system of three different KPIs, includes coverage energy efficiency, data energy efficiency and energy efficiency of power supply unit.					
Test Case Target Layer or Application	End to End					
Test Environment	Field and Lab					
Test Deployment Setup	Full system					
Network Setup	Will be defined in D5.2					
Test Configuration	End Device Density	Will be defined in D5.2	Mobility Level	Will be defined in D5.2	Has Background Traffic	Will be defined in D5.2
Initial Conditions / Prerequisites						
Will be defined in D5.2						
Test Scenario						
Will be defined in D5.2						
Test Variables						
Will be defined in D5.2						
Expected Behaviour / Target KPIs / Values						
Will be defined in D5.2						

Table 53: Finnish Cluster Test Case #3

Test Case #4: Energy capacity

Leading Partner: UOULU	Test Case #4: Energy capacity					
Test Case Objective	Measures the energy production, storage and charging/discharging capability, includes KPIs battery status, hydrogen fuel cell usage control and remote site energy sufficiency.					
Test Case Target Layer or Application	End to End and all Applications					
Test Environment	Field					
Test Deployment Setup	Full system					
Network Setup	Will be defined in D5.2					
Test Configuration	End Device Density	Will be defined in D5.2	Mobility Level	Will be defined in D5.2	Has Background Traffic	Will be defined in D5.2
Initial Conditions / Prerequisites						
Will be defined in D5.2						
Test Scenario						
Will be defined in D5.2						
Test Variables						

Will be defined in D5.2
Expected Behaviour / Target KPIs / Values
Will be defined in D5.2

Table 54: Finnish Cluster Test Case #4

Test Case #5: Energy profiling

Leading Partner: UOULU	Test Case #5: Energy profiling					
Test Case Objective	1-year operation under dynamic photovoltaic (PV)-WIND- DESS and Hydrogen fuel cell vs baseline cost of electricity using nordpool (simulated RU), includes KPI: Cost counter/Energy savings (€)					
Test Case Target Layer or Application	End to End and all Applications.					
Test Environment	Field and Lab (simulations)					
Test Deployment Setup	Full system					
Network Setup	Will be defined in D5.2					
Test Configuration	**End Device Density**	Will be defined in D5.2	**Mobility Level**	Will be defined in D5.2	**Has Background Traffic**	Will be defined in D5.2
Initial Conditions / Prerequisites						
Will be defined in D5.2						
Test Scenario						
Will be defined in D5.2						
Test Variables						
Will be defined in D5.2						
Expected Behaviour / Target KPIs / Values						
Will be defined in D5.2						

Table 55: Finnish Cluster Test Case #5

Test Case #6: Sustainability assessment

Leading Partner: UOULU	Test Case #6: Sustainability assessment					
Test Case Objective	This test measures how much carbon emissions can be avoided compared to the legacy commercial solutions.					
Test Case Target Layer or Application	End to End and all applications					
Test Environment	Field and Lab					
Test Deployment Setup	Full system					
Network Setup	Will be defined in D5.2					
Test Configuration	**End Device Density**	Will be defined in D5.2	**Mobility Level**	Will be defined in D5.2	**Has Background Traffic**	Will be defined in D5.2
Initial Conditions / Prerequisites						
Will be defined in D5.2						
Test Scenario						
Will be defined in D5.2						
Test Variables						

Will be defined in D5.2
Expected Behaviour / Target KPIs / Values
Will be defined in D5.2

Table 56: Finnish Cluster Test Case #6

Test Case #7: AI forecast accuracy assessment

Leading Partner: UOULU	Test Case #7: AI forecast accuracy assessment
Test Case Objective	This test case measures the accuracy of the deployed AI models used for energy weather forecasting and traffic forecasting. It provides the percentage based analysis of model accuracy.
Test Case Target Layer or Application	End to End and all applications
Test Environment	Field / Lab
Test Deployment Setup	Full system
Network Setup	Will be defined in D5.2
Test Configuration	**End Device Density**: Will be defined in D5.2 **Mobility Level**: Will be defined in D5.2 **Has Background Traffic**: Will be defined in D5.2
Initial Conditions / Prerequisites	
Will be defined in D5.2	
Test Scenario	
Will be defined in D5.2	
Test Variables	
Will be defined in D5.2	
Expected Behaviour / Target KPIs / Values	
Will be defined in D5.2	

Table 57: Finnish Cluster Test Case #7

French Cluster

Test Case #1.1.1: Drainage Site Clusterization

Leading Partner: GREEN	Test Case #1.1.1: Drainage Site Clusterization
Test Case Objective	Demonstrate the capability to clusterize drainage sites using meta data inputs such as Global Positioning System (GPS) position, address, altitude, and other context information such as nearby waste production site, criticality regarding flooding or pollution
Test Case Target Layer or Application	ML and vertical application layer
Test Environment	Field deployment in the city of Toulon
Test Deployment Setup	200 referenced gullies in the application with contextual information, 50 instrumented gullies with connected sensors and legacy connectivity, operational agent equipped and trained on vertical application.
Network Setup	NA. The test is based on meta data collected from referencing campaign
Initial Conditions / Prerequisites	
Getting previous cleaning frequency per drainage site	

<table>
<tr><th>Test Scenario</th></tr>
<tr><td>This is the specific set of steps to follow during the test:
1. Catching following information during referencing:
o GPS location
o operating criticality
o pollution criticality
o flooding criticality
o altitude
o network link regarding water flow (previous and next drainage)
o scoring of nearby production of waste
o identification of potential source of wastes: trees, fastfood, market, bar, …
2. Run the model to define the drainage bunches and the sampling to install sensors
3. Catching cleaning events and assessing clustering/prediction relevancy
4. Adjust model or its parameters to increase relevancy</td></tr>
<tr><th>Test Variables</th></tr>
<tr><td>Test with different machine learning models</td></tr>
<tr><th>Expected Behaviour / Target KPIs / Values</th></tr>
<tr><td>KPIs:
#9 (Hardware footprint reduction)
#12 (AI clusterisation and prediction performance)</td></tr>
</table>

Table 58: French Cluster Test Case #1.1.1

Test Case #1.1.2: Prediction of Waste Volume

<table>
<tr><th>Leading Partner: GREEN</th><td colspan="6">Test Case #1.1.2: Prediction of Waste Volume</td></tr>
<tr><th>Test Case Objective</th><td colspan="6">Description: Demonstrate the capability to predict the volume of waste in each bunch for proper waste collection

Use case 1: Sewer Network monitoring
Scenario 1: Large scale sensor virtualization in the rainwater network to reduce the hardware footprint and increase the ROI</td></tr>
<tr><th>Test Case Target Layer or Application</th><td colspan="6">ML and vertical application layer</td></tr>
<tr><th>Test Environment</th><td colspan="6">Field deployment in the city of Toulon</td></tr>
<tr><th>Test Deployment Setup</th><td colspan="6">200 referenced gullies in the application with contextual information, 50 instrumented gullies with connected sensors and legacy connectivity, operational agent equipped and trained on vertical application.</td></tr>
<tr><th>Network Setup</th><td colspan="6">NA. The test is based on meta data collected from referencing campaign</td></tr>
<tr><th>Test Configuration</th><th>End Device Density</th><td>NA</td><th>Mobility Level</th><td>NA</td><th>Has Background Traffic</th><td>NA</td></tr>
<tr><th colspan="7">Initial Conditions / Prerequisites</th></tr>
<tr><td colspan="7">Getting operating usages from the operator team</td></tr>
<tr><th colspan="7">Test Scenario</th></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test:
1. Catching following information during referencing:
o GPS location
o operating criticality
o pollution criticality
o flooding criticality
o altitude
o network link regarding water flow (previous and next drainage)
o scoring of nearby production of waste
o identification of potential source of wastes: trees, fastfood, market, bar, ...</td></tr>
</table>

2. Run the model to define the drainage bunches and the sampling to install sensors 3. Catching cleaning events and assess clustering/prediction relevancy 4. Adjust model or its parameters to increase relevancy
Test Variables
Test different machine learning models
Expected Behaviour / Target KPIs / Values
KPIs: #9 (Hardware footprint reduction) #12 (AI clusterization and prediction performance)

Table 59: French Cluster Test Case #1.1.2

Test Case #1.1.3: Environmental Benefit Indicator. Waste Collection

Leading Partner: GREEN	Test Case #1.1.3: Environmental Benefit Indicator. Waste collection
Test Case Objective	Description: Demonstrate the capability to report environmental benefit indicator Use case 1: Sewer Network monitoring Scenario 1: Large scale sensor virtualization in the rainwater network to reduce the hardware footprint and increase the ROI
Test Case Target Layer or Application	Application layer
Test Environment	Field deployment in the city of Toulon
Test Deployment Setup	200 referenced gullies in the application with contextual information, 50 instrumented gullies with connected sensors and legacy connectivity, operational agent equipped and trained on vertical application.
Network Setup	LPWAN Legacy network is used for this test that need large scale deployment
Initial Conditions / Prerequisites	
Organization of a workshop with the organization Citeo to define a method for evaluating the quantity of additional collected waste compared to legacy operation	
Test Scenario	
1. Definition of the method for evaluating the quantity of additional waste collected 2. Compute and Record collected waste volume for each intervention using the waste level change before and after the intervention during 3 months over the 200 monitored gullies. 3. Use the method defined in 1 to produce the indicator 4. Present the result to stakeholder (municipality and CITEO) in order to agree on results	
Test Variables	
NA	
Expected Behaviour / Target KPIs / Values	
KPIs #8 (Environmental benefit report). The result evaluation is qualitative using stakeholders feedback.	

Table 60: French Cluster Test Case #1.1.3

Test Case #1.2.1: Detection of Abnormal Behavior in Sewer Network

Leading Partner: GREEN	Test Case #1.2.1: Detection of Abnormal Behavior in Sewer Network
Test Case Objective	Description: Demonstrate the capability to detect abnormal behavior in the sewer network. Use case 1: Sewer Network monitoring

	Scenario 2: Large scale Abnormal sewer network behaviour detection to reduce inundation and pollution risk
Test Case Target Layer or Application	AI and application layer
Test Environment	Lab testing and Field Test in the city of Sete
Test Deployment Setup	30 instrumented black spots with connected sensors and legacy connectivity, operational agent equipped and trained on vertical application.
Network Setup	LPWAN Legacy network is used for this test that need large scale deployment
Initial Conditions / Prerequisites	
Deployment of the new AI service on the environment	
Test Scenario	
1. Collect historical data (waste water level and rain to train the model with normal behaviour 2. Run the model offline on normal chroniques and abnormal chroniques. Evaluate the probability of detection and probability of false alarm 3. Present the solution together with the alert priorization and filtering to the stakeholders 4. Deploy the model in production for the city of sete for 1 month 5. Interview the customer about the quality and benefit of the solution	
Test Variables	
Test the threshold and persistence of detection	
Expected Behaviour / Target KPIs / Values	
#7 (Alert filtering and prioritization) #13 (AI clogging detection performance)	

Table 61: French Cluster Test Case #1.2.1

Test Case #2.3.1: Energy Saving. RAN DRX Optimization

Leading Partner: TSGF, EUR	Test Case #2.3.1: Energy Saving. RAN DRX Optimization					
Test Case Objective	Demonstrate the capability of saving energy, by dynamically adjusting the RAN wake up cycles in DRX, tailored to the sensors traffic					
Test Case Target Layer or Application	RAN and data-fabric layer. AI-App and N-app					
Test Environment	Lab					
Test Deployment Setup	RedCap device, 5G RAN, 5G CN					
Network Setup	Amarisoft RAN connected to EURECOM platform and ML-Ops framework. RedCap compliant UE connected to the RAN.					
Test Configuration	End Device Density	1	Mobility Level	No	Has Background Traffic	
Initial Conditions / Prerequisites						
5G connectivity between RedCap user and Amarisoft RAN Prerequisite: AI model trained						
Test Scenario						
This is the specific set of steps to follow during the test: 1. Uplink traffic from 5G sensors to RAN 2. Link establishment between AI-App, N-App and RAN 3. Continuous collection of RAN measurements and AI-App inference						
Test Variables						
Different traffic profiles can impact RAN configuration						
Expected Behaviour / Target KPIs / Values						

Calculated KPI: KPIs 1-6 and 10, 11

Table 62: French Cluster Test Case #2.3.1

Test Case #2.3.2: Environmental Benefit Indicator. Green Space

Leading Partner: GREEN	Test Case #2.3.2: Environmental Benefit Indicator. Green Space
Test Case Objective	Description: Demonstrate the capability to report environmental benefit indicator Use case 1: Preserving green space with sustainable watering Scenario 3: Automatic Detection of abnormal watering and leakage in the irrigation network KPIs: #8 (Environmental benefit report)
Test Case Target Layer or Application	Vertical application layer
Test Environment	Field Test in the city of Marseille: Parc Longchamp
Test Deployment Setup	Smart watering solution was deployed in the Parc Longchamp in the city of Marseille with soil moistures connected sensors, connected meteo station and connected Electrovalve controleur. A meteo station will also be deployed on adjacent area with no green space in order to evaluate the cooling effect of vegetal.
Network Setup	LPWAN Legacy network is used for this test that need large scale deployment
Initial Conditions / Prerequisites	
Available team to run the survey. Obtain agreement from the city of Marseille	
Test Scenario	
1. Prepare survey questionnaire for habitant and prepare interview of stakeholders 2. Present the solution with the environmental benefit indicator to the stakeholders 3. Deploy the model in production for the city of Marseille for 1 month 4. Interview the customer about the perceived quality and benefit of the solution 5. Run survey questionnaires over 50 citizens regarding perceived benefits	
Test Variables	
Adjust the questionnaires regarding the first feedback	
Expected Behaviour / Target KPIs / Values	
KPI_008 Environmental and societal benefit indicators	

Table 63: French Cluster Test Case #2.3.2

Greek Cluster

Test Case #1: Network Performance Evaluation

Leading Partner: NCSRD	Evaluate the Mobile Network Performance
Test Case Objective	To verify that the mobile network supporting the Greek trials is fully operational and meets the expected performance benchmarks under real-world conditions. The focus is on evaluating end-to-end network performance metrics—including throughput, latency, jitter, packet loss, and signal quality—to ensure alignment with predefined KPIs.
Test Case Target Layer or Application	Network Layer - Monitoring Tools
Test Environment	Field / NCSRD Campus
Test Deployment Setup	All Network Components are up and running

	Measurement Tools are pre-installed Iperf3/Speedtest servers available in NCSRD premises and towards public server respectively					
Network Setup	- Amarisoft 5G SA services up and running - UE/Monitoring probe configuration in the NCSRD campus accessing NCSRD network with Public Land Mobile Network (PLMN) ID 00101 through the outdoor Amarisoft RAN using Access Point Name (APN) internet.					
Test Configuration	**End Device Density**	1-2	**Mobility Level**	Yes	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
- 5G Connectivity is established, - Monitoring Tools/Servers are running.						
Test Scenario						
This is the specific set of steps to follow during the test: 1. Measure the following performance metrics from location X - Downlink Throughput (DL) using iPerf3 and Speedtest - Uplink Throughput (UL) using iPerf3 and Speedtest - Latency (Round Trip Time, RTT) using ping or Speedtest latency values - Jitter using iPerf3 - Packet Loss using iPerf3 - Signal Strength & Quality (RSRP, RSRQ, SINR) from the UE 2. Repeat measurements in a walk/drive route of the demo field 3. The experiment is repeated Z times for statistical confidence and the Average, Min and Max over the Z repetitions is presented as end values.						
Test variables						
Ping interval, Maximum Transmission Unit (MTU) size, UDP/TCP Protocol, RTT						
Expected Behaviour / Target KPIs / Values						
Calculated KPI: KPIs 1-7						

Table 64: Greek Cluster Test case #1

Test Case #2: UAV relay Peak Data Rate

Leading Partner: ICCS	UAV relay peak data rate					
Test Case Objective	To verify the sustained peak data rate achievable when a UAV relay is used under near-ideal radio conditions. The test isolates the UAV relay's maximum capacity in order to provide a benchmark and validate that its operation does not introduce significant bottlenecks compared to direct device to device communication (e.g., between UAV and robotic dog through PC5).					
Test Case Target Layer or Application	Network Layer – Throughput Measurement Tools					
Test Environment	Field / NCSRD Campus					
Test Deployment Setup	All 5G SA core and RAN components are operational UAV relay airborne and stable at predefined altitude/position Iperf3/Speedtest servers available in NCSRD premises and towards public server respectively					
Network Setup	Amarisoft 5G SA services up and running UAV relay integrated with Amarisoft outdoor gNB setup					
Test Configuration	**End Device Density**	1	**Mobility Level**	Yes	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
• 5G connectivity established via UAV relay						

• UAV relay stable at nominal test position (altitude and location logged) • Monitoring tools/servers running • Network configured to allocate maximum available bandwidth to single UE • UAV start moving away from the gNB
Test Scenario
1. Establish stable 5G connection via UAV relay 2. Measure Downlink (same applies for Uplink) Throughput using: iPerf3 (TCP and/or UDP modes) 3. Log radio signal strength and quality metrics from UE (RSRP, RSRQ, SINR). 4. Repeat measurement X times to ensure statistical reliability. 5. Report Average, Min, Max throughput values
Test Variables
UAV relay altitude/position, UDP/TCP Protocol
Expected Behaviour / Target KPIs / Values
KPI_001 Peak user data rate (Sustained peak DL throughput close to theoretical gNB capacity for allocated bandwidth) KPI_005 Measure coverage through radio signal parameters (RSRP/RSRQ) (Stable RSRP, RSRQ, SINR values consistent with near-ideal conditions)

Table 65: Greek Cluster Test case #2

Test Case #3: Accuracy Score AI-App Location

Leading Partner: INF	Evaluate AI-App Location mean accuracy score					
Test Case Objective	To measure the mean accuracy metric of the AI-App Location within the Greek cluster.					
Test Case Target Layer or Application	Accuracy of AI-App Location					
Test Environment	Trial site infrastructure					
Test Deployment Setup	As described in the main text, more details can be found in Deliverables D2.1 and D3.1					
Network Setup	As described in the main text, more details can be found in Deliverables D2.1 and D3.1					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	N/A
Initial Conditions / Prerequisites						
• Network services are up and running • AI-Apps and ML-Ops framework are up and running • Corresponding V-Apps are up and running						
Test Scenario						
This is the specific set of steps to follow during the test: 1. Deploy the AI-App in the designated testbed environment as per the deployment setup. 2. Load the labelled test dataset corresponding to the Greek UC into the AI-App. 3. Initiate the inference pipeline and run predictions on the full test set. 4. Collect prediction results and compare them to ground truth labels to determine correct predictions. 5. Calculate the accuracy as the ratio of correct predictions to total predictions. 6. Repeat the inference process at least three times to capture variability and compute the mean and standard deviation of the accuracy. 7. Log the measured accuracy and check whether it meets or exceeds the predefined threshold (≥ 80%).						
Test Variables						
• Model Version: Different model iterations or retrained versions may yield varying accuracy levels.						

• Dataset Composition: Variations in the test dataset (e.g., balanced vs. imbalanced classes, domain-specific subsets). • Data Quality: Level of noise, label correctness, or resolution in the test inputs. • Repeatability Runs: Number of times the test is repeated to capture mean and standard deviation of results.
Expected Behaviour / Target KPIs / Values
#KPI_011 Reliability (Accuracy) ≥ 80%

Table 66: Greek Cluster Test Case #3

Test Case #4: Latency (Inference Time)

Leading Partner: INF	Evaluate AI-App Location inference time to detect abnormal activity					
Test Case Objective	To measure the mean latency for inference of the AI-App location.					
Test Case Target Layer or Application	Latency of AI-App Location					
Test Environment	Trial site infrastructure					
Test Deployment Setup	As described in the main text, more details can be found in Deliverables D2.1 and D3.1					
Network Setup	As described in the main text, more details can be found in Deliverables D2.1 and D3.1					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	N/A
Initial Conditions / Prerequisites						
• Network services are up and running • AI-App and ML-Ops framework are up and running • Corresponding V-Apps are up and running						
Test Scenario						
This is the specific set of steps to follow during the test: 1. Deploy the AI-App on the designated test hardware. 2. Trigger input data and ensure proper timestamping of input delivery. 3. Measure input preprocessing time and data transmission time if applicable. 4. Record the model inference time using profiling tools. 5. Capture the time taken to deliver the output response. 6. Sum all time segments to compute total latency: 7. Repeat the test at least 5 times to ensure consistency and compute average latency. 8. Validate if the latency is under the 300 ms target range for real-time responsiveness.						
Test Variables						
• **Hardware Type:** Latency may vary across Raspberry Pi, cloud VMs, or other edge devices. • **Model Complexity:** Size and architecture of the AI model. • **Input Data Size:** Signal length affects input processing time.						
Expected Behaviour / Target KPIs / Values						
#KPI_013 Latency (Inference Time) ≤ 300ms						

Table 67: Greek Cluster Test Case #4

Test Case #5: Wild-fire Detection AI-APP (WF AI-App) Inference Time Evaluation

Leading Partner: OTE	Evaluate Wild-fire Detection (WF) AI-APP Inference Time to Detect Smoke/Fire					
Test Case Objective	In this test case, it is verified that the time taken for the WF AI model to analyse a frame and detect Smoke/Fire is instant and according to first responder needs					
Test Case Target Layer or Application	Application Layer - WildFire Monitoring AI App					
Test Environment	Field / NCSRD Campus					
Test Deployment Setup	One (1) CAMERA at NCSRD campus connected over the 5G SA sending continuous video streams to the AI-APP in OTE Academy premises					
Network Setup	UE (Camera) located in the NCSRD campus accessing NCSRD network with PLMN ID 00101 through the outdoor Amarisoft RAN using APN internet generating video traffic to the WF AI-App installed at OTE edge site.					
Test Configuration	**End Device Density**	1	**Mobility Level**	No	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
• 5G Connectivity is established, • AI model successfully works with NCSRD Camera feed.						
Test Scenario						
This is the specific set of steps to follow during the test: 1. The WF AI model is running with NCSRD camera feeds for X minutes 2. The AI model logs, generated per frame, get post-processed: (log per frame example: Speed: 4.2ms preprocess, 50.3ms inference, 0.5ms postprocess per image at shape (1, 3, 1088, 1920)) 3. Average, Min and Max inference time is calculated 4. The experiment is repeated Z times for statistical confidence and the Average, Min and Max over the Z repetitions is presented as end values.						
Test Variables						
X minutes of running the AI model, Z repetition times for statistical confidence						
Expected Behaviour / Target KPIs / Values						
The expected behaviour is that the max inference time calculated will be less than 100 ms. Calculated KPI: KPI_014 WF AI Inference Time						

Table 68: Greek Cluster Test Case #5

Test Case #6: Wild-fire Detection -APP Accuracy in Detection and Reporting

Leading Partner: OTE	Evaluate the AI Performance in terms of (Average) Confidence Score, time to detect and time to alert for a smoke/fire event
Test Case Objective	To verify the appropriate AI model confidence for a prediction provided and measure the cumulative time taken to detect and alert for a detection event.
Test Case Target Layer or Application	Application Layer - WildFire Monitoring AI App + V App
Test Environment	Field / NCSRD Campus
Test Deployment Setup	One (1) CAMERA at NCSRD campus connected over the 5G SA sending continuous video streams to the AI-APP in OTE Academy premises
Network Setup	UE (Camera) located in the NCSRD campus accessing NCSRD network with PLMN ID 00101 through the outdoor Amarisoft RAN using APN internet generating video traffic to the WF AI-App installed at OTE edge site.

Test Configuration	End Device Density	1 Camera	Mobility Level	N/A	Has Background Traffic	No
Initial Conditions / Prerequisites						
• 5G Connectivity is established • AI model successfully works with NCSRD Camera feed • Sunny day with clear sky (no clouds). • The smoke/fire shall be put in a distance < 2000m. • No reflection from direct sunlight @camera.						
Test Scenario						
This is the specific set of steps to follow during the test: 1. The WF AI model starts running with NCSRD camera feeds. 2. Wait for a detection event either (a) through a real smoke/fire and/or use of a smoke agent or (b) feed the system with a recorded video of a real wild-fire starting 3. For each detection event (either (a) or (b) the following metrics get collected from the system: (a) the confidence score per frame processed (c_score), collected from the AI model logs for all frames (b) the timestamp of the first detection (fd_time) (c) the timestamp of the first alert notification (fa_time) (d) the earliest time that the smoke (s_time) is visible in the video stream analysed The following values are then calculated: • the average value of the confidence score for these frames (Sum of confidence scores per frame / the total number of frames) • the detection time difference (fd_time-s_time) • the alert time difference (fa_time-stime) The experiment is repeated Z times for statistical confidence and the Average, Min and Max over the Z repetitions is presented as end values for all calculated metrics.						
Test Variables						
1. Wild-fire event triggering 2. Z repetition times for statistical confidence						
Expected Behaviour / Target KPIs / Values						
The target is to detect smoke with acceptable confidence at its early stages within acceptable time to detect and alert intervals. Calculated KPIs: 1. KPI_012: Average Confidence Score > 60% 2. KPI_015: Time to Detect < 10s 3. KPI_016: Time to Alert < 60s						

Table 69: Greek Cluster Test Case #6

Test Case #7: Wild-fire Detection -App End-User Verification

Leading Partner: OTE	Evaluate the WF AI-App performance in early wildfire monitoring considering the end-user (first responders) interpretation of the detected versus real events.
Test case objective	To measure the percentage of detected events successfully confirmed by the first responders and the false positives (detected fires that were not true).
Test Case Target Layer or Application	Application Layer - WildFire Monitoring AI App + V App
Test Environment	Field / NCSRD Campus
Test Deployment Setup	One (1) CAMERA at NCSRD campus connected over the 5G SA sending continuous video streams to the AI-APP in OTE Academy premises

Network Setup	UE (Camera) located in the NCSRD campus accessing NCSRD network with PLMN ID 00101 through the outdoor Amarisoft RAN using APN internet generating video traffic to the WF AI-App installed at OTE edge site.					
Test Configuration	**End Device Density**	1 Camera	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
1. 5G Connectivity is established 2. AI model successfully works with NCSRD Camera feed						
Test Scenario						
This is the specific set of steps to follow during the test: 1. The WF AI model is running with NCSRD camera feeds for a period of time P (P >= month) Each event generated (#detected) is characterised by the first responders as correct (#true_detected) or not (#false_detected) 2. The sum of #false_detected in period P is calculated 3. The sum of real events #real in period P is reported by first responders 4. The #true_detected/#real in period P is calculated.						
Test Variables						
1. Wild-fire event 2. P period of running the test for statistical confidence						
Expected Behaviour/ Target KPIs / Values						
The target is to detect smoke satisfying the end-user experience reducing the number of false positives (due to other elements in nature such as such as clouds, fog, trees with red leaves on them, or sun burning on the background that can confuse the model) and concluding to a high percentage of true/verified notifications Calculated KPIs: 1. KPI_017: Number of False Positives < 7 per week 2. KPI_018: Verified Events confirmed > 90%						

Table 70: Greek Cluster Test Case #7

Test Case #8: Wild-fire Detection AI-APP (WF AI-App) Frame per Second (FPS) Impact in Energy Consumption

Leading Partner: OTE	Evaluate the impact of the camera's FPS in the WF AI-App Energy Consumption					
Test case objective	To measure the energy reduction at GPU upon reduction of the camera's FPS, progressively lowering FPS so that to reach 30% energy consumption reduction					
Test Case Target Layer or Application	Application Layer - WildFire Monitoring AI App					
Test Environment	Field / NCSRD Campus					
Test Deployment Setup	One (1) CAMERA at NCSRD campus connected over the 5G SA sending continuous video streams to the AI-APP in OTE Academy premises					
Network Setup	N/A – measurement in the AI infrastructure					
Test Configuration	**End Device Density**	1 Camera	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
1. AI model successfully works						
Test Scenario						

This is the specific set of steps to follow during the test: 1. The WF AI model is running with NCSRD camera feeds set to max FPS allowed (F = 20 FPS). 2. GPU's energy consumption data (Watt) are collected from NVIDIA-SMI every X seconds for Y intervals to calculate baseline min (minEC), max (maxEC) and average (aGC) consumption per GPU. 3. The camera's FPS is reconfigured to F = F-5 and step.2 is repeated. 4. aGC values collected are compared against the baseline measurement. 5. Steps 3 and 4 are repeated until an energy reduction of 30% is observed
Test Variables
1. Camera's FPS 2. GPU consumption 3. X minutes of running the AI model, Z repetition times for statistical confidence
Expected Behaviour / Target KPIs / Values
The target is to reduce the energy consumption of the WF-AI App per GPU to 30% from the baseline adapting the FPS processing Calculated KPIs: 1. KPI_019: Energy reduction in watts > 30%

Table 71: Greek Cluster Test Case #8

Test Case #9: Wild-fire Detection AI-APP (WF AI-App) Energy-Efficient Deployment Evaluation

Leading Partner: OTE	Identify the impact of GPU sharing between AI-App instances for Energy Consumption					
Test Case Objective	To measure the energy consumption at GPU SoC level and impact on the AI-App FPS when collocating more than one AI-App instance in the same GPU.					
Test Case Target Layer or Application	Application Layer - WildFire Monitoring AI App					
Test Environment	Field / NCSRD Campus					
Test Deployment Setup	One (1) CAMERA at NCSRD campus connected over the 5G SA sending continuous video streams to the AI-APP in OTE Academy premises Two (2) extra CAMERAs that perform wildfire monitoring in other locations					
Network Setup	N/A - measurement in the AI infrastructure					
Test Configuration	**End Device Density**	3 Cameras	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
1. The AI model is functioning as expected and producing accurate results						
Test Scenario						
This is the specific set of steps to follow during the test: 1. The WF AI model is running with NCSRD camera feeds set to max FPS allowed (F = 20 FPS). 2. Following metrics are collected every X seconds for Y intervals: a. GPU's energy consumption data (Watt), collected from NVIDIA-SMI to calculate baseline min (minEC), max (maxEC) and average (aGC) consumption per GPU. b. the average FPS processed by the AI model, measuring how many frames get processed per second 3. A second camera/AI model is associated to the same GPU and measurements of Step.2 are collected. 4. A third camera/AI model is associated to the same GPU and measurements of Step.2 are collected. 5. aGC values collected are compared against the baseline measurement. 6. The GPU consumption, and the average FPS of all AI models running at Steps 1,3,4 are compared to deduce optimal, energy efficient deployment. 7. The experiment is repeated Z times for statistical confidence and the Average, Min and Max over the Z repetitions is presented as end values for all calculates metrics.						

Test Variables
1. WF-AI App FPS 2. GPU and ARM SoC consumption (and Intel server consumption) 3. Z repetition times for statistical confidence
Expected Behaviour / Target KPIs / Values
The target is to deduce the optimal deployment by allocating more than one camera's/AI models per GPU considering the impact (drop) in the FPS processed due to the colocation. Calculated KPIs: 2. KPI_020: Energy Efficient Deployment > 2 AI models per GPU

Table 73: Greek Cluster Test Case #9

Test Case #10: Reliable Data Transmission Evaluation

Leading Partner: ICCS	Reliable Data Transmission Evaluation					
Test Case Objective	To verify the reliability and stability of data transmission between UAVs, robotic dog, and control stations within the 5G/6G infrastructure during AI/V-App operation.					
Test Case Target Layer or Application	End-to-End / V-App (Video & Data Stream Reliability)					
Test Environment	Field					
Test Deployment Setup	UAV, robotic dog and Edge Server hosting AI/V-Apps. Data transmission monitored.					
Network Setup	Edge server connected via fiber.					
Test Configuration	**End Device Density**	2-3 devices (UAVs, robotic dog, server)	**Mobility Level**	UAVs and robotic dog moving within 100 m radius	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
Network operational, devices registered, AI/V-Apps running, monitoring tools active.						
Test Scenario						
• Establish continuous data stream (telemetry and video) between devices. • Run transmission for 10 min intervals. • Record packet loss, retransmissions, throughput.						
Test Variables						
UAV distance, mobility, network slice configuration.						
Expected Behaviour / Target KPIs / Values						
KPI_021: Reliable Data Transmission ≥ 95 % success rate.						

Table 74: Greek Cluster Test Case #10

Test Case #11: Localization accuracy of victims and objects

Leading Partner: ICCS	Localization accuracy of victims and objects
Test Case Objective	To measure the positioning accuracy of the AI/V-App system that fuses UAVs, Robotic dog, GPS and visual data for objects or victims localization.
Test Case Target Layer or Application	AI-App / V-App
Test Environment	Field and Lab

<table>
<tr><td>Test Deployment Setup</td><td colspan="6">UAV with camera and GPS sensors; robotic dog with positioning unit; AI-App performing localization on edge GPU.</td></tr>
<tr><td>Network Setup</td><td colspan="6">Testbed</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>2-4 devices</td><td>Mobility Level</td><td>moving targets</td><td>Has Background Traffic</td><td>No</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">AI-App loaded with trained model, synchronization of GPS clocks.</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">Execute UAV patrol over predefined coordinates.
Compare detected object positions vs ground-truth GPS.
Compute localization error in meters.
Or /and
Execute UAV patrol over predefined coordinates.
Compare detected object positions vs ground-truth GPS.
Compute localization error in meters.</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">Altitude, frame rate</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">KPI_022: Localization Accuracy ≤ 5 m.</td></tr>
</table>

Table 72: Greek Cluster Test Case #11

Test Case #12: Victim Detection Accuracy Evaluation

<table>
<tr><td>Leading Partner: ICCS</td><td colspan="6">Victim Detection Accuracy Evaluation</td></tr>
<tr><td>Test Case Objective</td><td colspan="6">To assess the AI-App’s ability to detect human victims or abnormal activity from UAV and robotic dog video feeds.</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">AI-App</td></tr>
<tr><td>Test Environment</td><td colspan="6">Field - Simulated rescue area</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">UAV and robotic dog with RGB cameras streaming to edge GPU for inference; labeled dataset for verification.</td></tr>
<tr><td>Network Setup</td><td colspan="6">Testbed</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>UAV and robotic dog</td><td>Mobility Level</td><td>-</td><td>Has Background Traffic</td><td>-</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">Trained AI-App model deployed on edge node.</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">Run UAV scan missions detecting dummy victims.
Record true/false positives, recall, and precision.
Or/and
Run robotic dog scan missions detecting victims.
Record true/false positives, recall, and precision.</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">Camera type, lighting conditions, altitude.</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">KPI_023: Victim Detection Accuracy ≥ 80 %.</td></tr>
</table>

Table 73: Greek Cluster Test Case #12

Test Case #13: UAV Tracking Latency

Leading Partner: ICCS	UAV Tracking Latency						
Test case objective	To measure the delay between UAV capture and its visualization on the operator dashboard.						
Test Case Target Layer or Application	V-app						
Test Environment	Lab and Field						
Test Deployment Setup	UAV transmitting via 5G to Edge Server; dashboard application displays live position data.						
Network Setup	Testbed						
Test Configuration	**End Device Density**	1 UAV	**Mobility Level**	-	**Has Background Traffic**	-	
Initial Conditions / Prerequisites							
All systems synchronized.							
Test Scenario							
• Start UAV telemetry transmission. • Record timestamps at capture and dashboard update. • Compute end-to-end delay.							
Test Variables							
Packet size, update frequency, distance.							
Expected Behaviour / Target KPIs / Values							
KPI_024: UAV Tracking Latency < 2 s.							

Table 77: Greek Cluster Test Case #13

Test Case #14: Video Transmission Latency

Leading Partner: ICCS	Video Transmission Latency						
Test case objective	To measure the delay between video capture by UAV and its visualization at the control center or V-App.						
Test Case Target Layer or Application	V-App						
Test Environment	Field - outdoor UAV mission area						
Test Deployment Setup	UAV with camera streaming live to edge node; AI-App optionally performing frame analysis.						
Network Setup	Testbed						
Test Configuration	**End Device Density**	1 UAV	**Mobility Level**	-	**Has Background Traffic**	-	
Initial Conditions / Prerequisites							
Streaming encoder and player synchronized; clock alignment.							
Test Scenario							
• Capture live video stream over 5G link. • Timestamp frames at source and reception. • Calculate average end-to-end video latency.							
Test Variables							
Resolution, bitrate, distance, codec type.							
Expected Behaviour / Target KPIs / Values							

KPI_025: Video Transmission Latency < 3 s.

Table 78: Greek Cluster Test Case #14

Test Case #15: Evaluate AI-App detection inference time to detect object or victim

Leading Partner: ICCS	Evaluate AI-App detection inference time to detect object or victim					
Test case objective	To evaluate the inference latency of the AI-App deployed as a set of distributed microservices running on edge GPU infrastructure. The objective is to verify that the detection pipeline performs within real-time limits (<100 ms per frame), enabling timely identification.					
Test Case Target Layer or Application	Latency of AI-App Location					
Test Environment	Field or lab					
Test Deployment Setup	Testbed					
Network Setup	Testbed					
Test Configuration	**End Device Density**	2-3 edge-connected devices (UAVs/robots)	**Mobility Level**	N/A	**Has Background Traffic**	N/A
Initial Conditions / Prerequisites						
• All AI microservices are deployed, containerized, and reachable via the service mesh. • The edge GPU node is operational, with the trained detection model preloaded and optimized. • Message broker and data bus (e.g., MQTT / Kafka) are configured for event communication between services. • Time synchronization is ensured between edge, UAVs, and control station. • Logging and monitoring services are running for performance data collection.						
Test Scenario						
• Start video and telemetry streams from UAV and ground robot to edge node. • Process incoming frames through the distributed AI-App microservices chain. • Record inference timestamps at each processing stage (input → detection output). • Calculate end-to-end inference latency for each frame. Repeat measurements for different network loads and model versions.						
Test Variables						
Frame rate, model size, container orchestration latency, GPU load, and message queue delay.						
Expected Behaviour / Target KPIs / Values						
KPI_001, KPI026						

Table 79: Greek Cluster Test Case #15

Test Case #16: Energy Efficiency Evaluation

Leading Partner: ICCS	Evaluate the energy efficiency
Test Case Objective	To evaluate the energy efficiency of compute and network components or applications used in the 6G-VERSUS experimentation environment within the Greek Cluster. The goal is to quantify the energy consumption per task across different hardware platforms (Intel-based CPU, ARM-based CPU, NVIDIA Jetson), supporting energy-aware orchestration and alignment with sustainability KPIs (KPI_008-KPI_010).
Test Case Target Layer or Application	Network and Application Layers – Energy Monitoring and Profiling.

Test Environment	ICCS Laboratory / Testbed Premises; NCSRD Trial Premises.					
Test Deployment Setup	Compute nodes equipped with Intel and ARM CPUs, and NVIDIA Jetson accelerators. Power measurement tools (e.g. RAPL, tegrastats, Smart Power Meter) are pre-installed and synchronised with task execution logs.					
Network Setup	Local ICCS compute cluster connected to the 6G-VERSUS network. Energy monitoring agents report metrics to the orchestration layer.					
Test Configuration	**End Device Density**	2-5	**Mobility Level**	Yes	**Has Background Traffic**	Yes

Initial Conditions / Prerequisites
- Compute and monitoring tools are operational. - Synchronisation between task start/stop timestamps and power meters is verified. - Benchmark workloads (e.g. inference tasks, synthetic loops) pre-deployed.
Test Scenario
1. Run benchmark tasks on Intel, ARM, and NVIDIA systems sequentially. 2. Record instantaneous and average power consumption during each run. 3. Calculate energy per task (J/task) and average power (Watt). 4. Repeat for varying workloads and record task duration, completion rate, and CPU/GPU utilisation. 5. Aggregate data to compute KPI008–KPI010 values.
Test Variables
Defined after - Task type (AI inference, network or compute control loop, etc.); - CPU/GPU architecture; - Power measurement interval (in ms); - Task size and duration; Values (J/task, Watt) expected within defined efficiency thresholds and to be compared against baseline reference measurements.
Expected Behaviour / Target KPIs / Values
Calculated KPI: KPI_008; KPI_009; KPI_010

Table 74: Greek Cluster Test Case #16

Portuguese Cluster

Scenario 1 – Test Case #S1.1

Leading Partner: JSIO	Test Case #S1.1 - 5G NR RedCap-enabled IoT device energy and harvesting efficiency
Test Case Objective	TO #1: Demonstrate energy-efficiency enabled by energy harvesting and intelligent network-aware verticals
Test Case Target Layer or Application	Port environmental monitoring V-App
Test Environment	Trial site infrastructure
Test Deployment Setup	IoT Device, V-App
Network Setup	As described in section 3.5.1 Testbed Overview in D2.1. OAI 5G RAN is assumed for 5G NR RedCap support.

Test Configuration	End Device Density	1	Mobility Level	None	Has Background Traffic	No
Initial Conditions / Prerequisites						
The IoT device must be pre-registered in the 5G Network and powered on.						
Test Scenario						
This is the specific set of steps to follow during the test: 1. Initialize energy measurement at the IoT device. 2. Start and configure the 5G NR RedCap IoT device, connecting it to the network infrastructure, initializing the sensors and its datalogger. 3. Keeping the device in idle mode, measure the energy balance after a given time - register IoT Device energy consumption in idle mode. 4. Measure available battery power before transmission. Transmit N standardized data packets to V-App server. Measure available battery power immediately after each transmission, and calculate energy consumed per successful data transmission. Register 5G NR RedCap IoT device transmission efficiency. 5. Replace the connectivity module from 5G NR RedCap to standard 5G NR. Repeat steps 1 to 4. Register IoT device transmission efficiency and calculate the difference to 5G NR RedCap IoT device transmission efficiency. 6. Switch from battery power to energy harvesting mode (solar, kinetic or both). 7. Register available power at start and after a given time. Calculate energy harvesting efficiency						
Test Variables						
Uplink communication rate / frequency. Energy harvesting mode.						
Expected Behaviour / Target KPIs / Values						
KPI_002: Device Radio energy efficiency gain using 5G NR RedCap (> 25%) KPI_007: 5G NR RedCap IoT device transmission energy efficiency (target value under study) KPI_009: Datalogger energy consumption in idle mode (< 50 mW/h) KPI_010: Energy harvesting efficiency (> 80%)						

Table 75: Portuguese Cluster Test Case #S1.1

Scenario 1 – Test Case #S1.2

Leading Partner: IPN	Test Case #S1.2 - Energy efficiency using AI-App					
Test Case Objective	**TO #1:** Demonstrate energy-efficiency enabled by energy harvesting and intelligent network-aware verticals					
Test Case Target Layer or Application	Port environmental monitoring V-App					
Test Environment	Trial site infrastructure					
Test Deployment Setup	IoT Device, V-App					
Network Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	1	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						

• The IoT device must be pre-registered in the 5G Network and powered on. • In the testing environment external factors that could skew results must be minimized. • All test devices must be properly configured, and the collection system must be operational to accurately record the relevant metrics. • The user-defined T max value must be established and configured for the test, according with the context.
Test Scenario
This is the specific set of steps to follow during the test: 1. Initialize energy measurement at the IoT device. 2. Start and configure the 5G NR RedCap IoT device, connecting it to the network infrastructure, initializing the sensors and its datalogger. 3. Define the baseline energy consumption by applying a naive, rule- or threshold- based policy to a defined group of devices over a set period. During this time, for each transmission register the following: data points per batch, cost, and the time interval. 4. On the same group devices, repeat the process with the AI-app for a similar period, collecting the same data/measurements. 5. Compute the TWE final score for each approach: Baseline and using AI-app. 6. Compute the Relative Efficiency Index
Test Variables
• Independent Variable: Policy Generation Method (baseline rules/polices vs AI-driven) • Dependent Variable: Time-Weighted Efficiency score • Controlled Variables: Device hardware and configurations, testing environment, evaluation time period, device readings, KPI’s T_max parameter.
Expected Behaviour / Target KPIs / Values
KPI_008: Relative Efficiency Index (REI) (> 1.1)

Table 76: Portuguese Cluster Test Case #S1.2

Scenario 1 – Test Case #S1.3

Leading Partner: IPN	Test Case #S1.3 - AI-based forecasting performance					
Test Case Objective	**TO #4:** Demonstrate the benefits of AI for performant and efficient 6G applications					
Test Case Target Layer or Application	IoT device energy management AI-App					
Test Environment	Lab					
Test Deployment Setup	IoT Device, V-App, AI-App, AI/ML Ops Framework					
Network Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	0	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
Available data: Multivariable temporal series collected from sensors, completed, clean and continuous. With a regular and consistent sampling interval across all variables, with timestamps. Stored in a database or properly structured file to ensure reproducibility. Forecasting AI approach developed.						

Selected baseline.
Test Scenario
This is the specific set of steps to follow during the test: 1. Retrieve multivariate time series data from sensors with regular sampling intervals 2. Split the data into training, validation, and test sets, with temporal ordering preserved. 3. Train the model on the training data. 4. Apply trained model to test data to generate predictions. 5. Calculate baseline forecasts for comparison and dataset-wide means for each variable using all available historical data. 6. Compute: a. Mean Absolute Error (MAE) for both model predictions and naive baseline across all forecasts. b. Root Mean Square Error (RMSE) for each variable. 7. Calculate: a. MASE and aggregate the results. b. CV-RMSE and aggregate results across variables. c. Forecast Bias and aggregate results for each variable. 8. Validate: a. MASE against the target value. b. CV-RMSE against the target value. c. Forecast Bias against the target value. 9. Data characteristics: • Time-stamped, continuous metric readings from sensors, sampled at regular intervals. • All data is stored in a structured database for reproducibility and traceability.
Test Variables
Independent variables: Sensor data (multivariable time series data). Dependent variables: Model predictions of future sensor reading/device states. Control variables: • Data split used for training, validation, and testing. • Temporal order of the data. • The naive forecasting baseline used for comparison.
Expected Behaviour / Target KPIs / Values
KPI_011: Mean Absolute Scaled Error (MASE) (≤ 90 %) **KPI_012:** Coefficient of Variation RMSE (CV-RMSE) (≤ 25 %) **KPI_013:** Forecast Bias (≤ 5 %)

Table 77: Portuguese Cluster Test Case #S1.3

Scenario 1 – Test Case #S1.4

Leading Partner: ONE	Test Case #S1.4: Maritime Object Detection
Test Case Objective	**TO #3**: Validate network performance and service agility through advanced slice management **TO #4**: Demonstrate the benefits of AI for performant and efficient 6G applications
Test Case Target Layer or Application	End to End / Layer or Application specific – V-App (Mobitrust)
Test Environment	Trial site infrastructure (Port of Aveiro / Ílhavo)
Test Deployment Setup	RC boat with autonomous navigation, shoreline beacon or marker buoys, Mobitrust dashboard, data logging server.
Network Setup	ALB/ITAV testbed, dedicated slice for control/telemetry where applicable

Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
• Devices registered and powered • Safety checks completed • Mobitrust UI reachable • Evaluation area geo-fenced.						
Test Scenario						
1. Define patrol path (waypoints) and sensing cadence. 2. Start test, stream video to Mobitrust V-App, record ground-truth buoys/targets. 3. Run detection pipeline; log detections with timestamps and positions. 4. Repeat with varied lighting/sea state (if feasible). 5. Store and export logs/data.						
Test Variables						
Patrol speed; camera sampling rate; altitude of UAV (if used); target size/type.						
Expected Behaviour / Target KPIs / Values						
KPI_016: Object detection accuracy (Water) (target ≥ 80%)						

Table 78: Portuguese Cluster Test Case #S1.4

Scenario 1 – Test Case #S1.5

Leading Partner: ONE	Test Case #S1.5: Environmental Monitoring Capabilities					
Test Case Objective	**TO #3:** Validate network performance and service agility through advanced slice management **TO #4:** Demonstrate the benefits of AI for performant and efficient 6G applications					
Test Case Target Layer or Application	End to End / Layer or Application specific – V-App (Mobitrust)					
Test Environment	Trial site infrastructure (Port of Aveiro / Ílhavo waterways)					
Test Deployment Setup	RC boat and/or UAV with RGB (Red, Green, and Blue) camera; known surface targets (e.g., controlled markers, benign floaters); Mobitrust dashboard; ground-truth team.					
Network Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
• Calibrate cameras • Load target classes • Place/verify ground-truth markers • Verify time sync and storage.						
Test Scenario						
1. Define evaluation area, place surface markers of varying sizes/shapes. 2. Execute navigation path; stream imagery to Mobitrust; log detections. 3. Repeat under varied illumination/angles; capture control set with no targets. 4. Store and export logs/data.						
Test Variables						

Target size (cm²), color/contrast, viewing angle/altitude, frame rate.
Expected Behaviour / Target KPIs / Values
KPI_019: Water surface pollution detection (target ≥ 80% detection rate) **KPI_020:** Water surface minimum object detection size (target ≤ 10 cm² detectable)

Table 79: Portuguese Cluster Test Case #S1.5

Scenario 2 – Test Case #S2.1

Leading Partner: ITAV	Test Case #S2.1 Network radio energy optimization					
Test Case Objective	**TO #2:** Measure energy efficiency of the network platform					
Test Case Target Layer or Application	End to End / Layer or Application specific					
Test Environment	Trial site infrastructure					
Test Deployment Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Network Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
The full network platform deployment including a 5G UE (or UE emulator), a 5G NR gNB and 5G core is up and running. A traffic generator capable of producing UDP or TCP traffic is installed behind the N6 interface of 5G system and a traffic sink that collects traffic is used at the UE side. Equipment able to measure energy consumption over a certain period of time is in place.						
Test Scenario						
The test case is composed of two parts. The first part is intended to measure the platform energy efficiency under normal (non-optimized) conditions. The second part is similar to the first, except that dynamic resource management and intelligent activation/deactivation of network components are performed, with a view to evaluating energy consumption optimization. In part 1, multiple test iterations are executed as follows: 1. Test iterations consisting of generating traffic composed of full-length IP packets (i.e. 1500 bytes) are performed during the baseline measurement duration of 24 hours. 2. Step 1 may be repeated for multiple days, with different values of traffic throughput, test iteration duration and periodicity of iterations, according to multiple target scenarios. 3. The volume of data traffic and respective energy consumption over a certain time period is collected via appropriate measurement tools. 4. Energy Efficiency in a certain time interval (KPI_006) is calculated as follows: $EE = DV / EC$ in which: EE: Energy Efficiency (in bit/Joule) DV: Total Data Volume in the network (in bit) EC: Energy Consumption (in Joule) The second part is similar to the first, except that advanced resource management techniques, including dynamic activation/deactivation of network components, or smart network slice management, are used to optimize the energy efficiency of the network platform. Energy efficiency is measured following the same methodology described above, under the same conditions. The ratio between energy efficiencies measured in parts #1 and #2 under similar conditions, which quantifies the energy efficiency optimization enabled by dynamic resource management (KPI_001), is calculated as follows: $EE_{op} = (EE_1 - EE_2) / EE_1$						

in which: EE_1: Energy Efficiency measured in phase 1 (in bit/Joule) EE_2: Energy Efficiency measured in phase 2 (in bit/Joule) EE_{op}: Energy Efficiency Optimization (in %)
Test Variables
Number of traffic iterations in 24-hour period, number of days, volume and characteristics of the data injected in the network, netwok resource management techniques used in the second part of the test.
Expected Behaviour / Target KPIs / Values
KPI_001: Network radio energy consumption (< 10 %) **KPI_006:** Network energy efficiency (target value under study)

Table 80: Portuguese Cluster Test Case #S2.1

Scenario 2 – Test Case #S2.2

Leading Partner: ITAV	Test Case #S2.2: Slice Provisioning Time					
Test Case Objective	**TO #3**: Validate network performance and service agility through advanced slice management					
Test Case Target Layer or Application	End to End / Layer or Application specific					
Test Environment	Trial site infrastructure					
Test Deployment Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Network Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
Network services are up and running. A VM hosting a packet capture service (to be started during test execution) is created.						
Test Scenario						
This test case is performed using a VM and the methodology is composed of the following steps: 1. A UE is turned on and configured to use the slice which is still non-existent. This UE will persistently try to connect to the created VM by sending a ping-like packet. 2. A packet capture is started in the VM 3. The VM sends a request to create the slice to the Slice Manager 4. The VM will receive the UE's attempt to connect as soon as the slice has been provisioned 5. The packet capture is then used to obtain the slice provisioning time by analysing the timestamps of the slice creation request packet and the first successful UE packet.						
Test Variables						
Type of network slice to be created, type of VM to be used.						
Expected Behaviour / Target KPIs / Values						
KPI_004: Slice Provisioning time (< 180 s)						

Table 81: Portuguese Cluster Test Case #S2.2

Scenario 2 – Test Case #S2.3

Leading Partner: ITAV	Test Case #S2.3: Service Provisioning Time

Test Case Objective	**TO #3:** Validate network performance and service agility through advanced slice management					
Test Case Target Layer or Application	End to End / Layer or Application specific					
Test Environment	Trial site infrastructure					
Test Deployment Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Network Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
Network services are up and running. A VM hosting a packet capture service (to be started during test execution) is created.						
Test Scenario						
The measurement of Service Provisioning Time is performed using a VM and adheres to the following methodology (considering that the service is packed in container format): 1. A packet capture is started in the VM. 2. The VM is used to send a service creation command towards the platform's orchestrator. The service to be deployed contains an add-on which signals an agent, via an "alive" message in the requesting VM when it is ready to start accepting requests. 3. The orchestrator will instantiate the necessary container(s) for the service. 4. When the service is running, it sends an "alive" message to the VM which sent the deployment request, informing that it is ready to accept requests. 5. The packet capture is then analysed to obtain the service provisioning time by subtracting the service's "alive" packet timestamp to the timestamp of the service provisioning packet.						
Test Variables						
Type of container to be instantiated.						
Expected Behaviour / Target KPIs / Values						
KPI_005: Service provisioning time (< 100 ms)						

Table 82: Portuguese Cluster Test Case #S2.3

Scenario 2 – Test Case #S2.4

Leading Partner: ITAV	Test Case #S2.4: Throughput					
Test Case Objective	**TO #3:** Validate network performance and service agility through advanced slice management					
Test Case Target Layer or Application	End to End					
Test Environment	Field					
Test Deployment Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Network Setup	As described in section 3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	Optional
Initial Conditions / Prerequisites						

<table>
<tr><td>Prior to the test execution, two network slices are established, “base tier” and “top tier”. Two monitoring probes running iPerf client and server, are connected to 5G network N1 and N6 interfaces, respectively</td></tr>
<tr><th>Test Scenario</th></tr>
<tr><td>This is the specific set of steps to follow during the test:

For TCP throughput

1. Start monitoring probes.
2. Through the slice manager, steer the traffic between probes to the “base tier” slice.
3. Start transmitting from client probe to server probe (bandwidth limitation set to unlimited, duration = 90 s, skip the first 20 seconds).
4. Through the slice manager API, switch the traffic between probes to the “top tier” slice;
5. Repeat step 3.
6. Record the throughput for each time interval.
7. Stop the traffic generator.
8. Calculate and record the KPIs per iteration.
9. Repeat steps 2 to 8 for 25 iterations
10. Stop monitoring probes.
11. Compute the KPIs

For UDP throughput, the process is similar to TCP, except:

- Bandwidth limitation set to the maximum theoretical throughput available in the network
- 4 simultaneously transmitting probes</td></tr>
<tr><th>Test Variables</th></tr>
<tr><td>Volume and characteristics of background traffic in the “base tier” slice when measuring throughput in the “top tier” slice.</td></tr>
<tr><th>Expected Behaviour / Target KPIs / Values</th></tr>
<tr><td>KPI_003: Throughput (> 25 Mbps)

Throughput should be enough to guarantee the expected quality required by video streaming in the two different situations.</td></tr>
</table>

Table 83: Portuguese Cluster Test Case #S2.4

Scenario 2 – Test Case #S2.5

<table>
<tr><th>Leading Partner: ONE</th><td colspan="6">Test Case #S2.5: Autonomous Navigation Coverage</td></tr>
<tr><th>Test Case Objective</th><td colspan="6">TO #3: Validate network performance and service agility through advanced slice management
TO #4: Demonstrate the benefits of AI for performant and efficient 6G applications
TO #6: Assess 6G-Apps' autonomous systems’ reliability and precision</td></tr>
<tr><th>Test Case Target Layer or Application</th><td colspan="6">End to End / Layer or Application specific – V-App (Mobitrust)</td></tr>
<tr><th>Test Environment</th><td colspan="6">Trial site infrastructure (Port of Aveiro / Ílhavo)</td></tr>
<tr><th>Test Deployment Setup</th><td colspan="6">RC boat with autonomous navigation, shoreline beacon or marker buoys, Mobitrust dashboard, data logging server.</td></tr>
<tr><th>Network Setup</th><td colspan="6">ALB/ITAV testbed, dedicated slice for control/telemetry when applicable</td></tr>
<tr><th>Test Configuration</th><th>End Device Density</th><td>N/A</td><th>Mobility Level</th><td>N/A</td><th>Has Background Traffic</th><td>No</td></tr>
</table>

Initial Conditions / Prerequisites
Navigation waypoints uploaded; geofence loaded; failsafe/RTL configured; time sync verified; Mobitrust session open
Test Scenario
1. Route/risk briefing, prepare RC vehicle and Mobitrust app setup, system checks (Global Navigation Satellite System (GNSS) lock, battery). 2. Run test run to check favourable water conditions (e.g., current strength, agitation) 3. Register route waypoints in Mobitrust. 4. Request low-latency control slice via N-App; confirm RTT/jitter values. 5. Start autonomous navigation along predefined route path. 6. Log planned vs. actual track, record control/telemetry latency. 7. Finish navigation and return to shore, collect and export telemetry.
Test Variables
Waypoint spacing/location; navigation speed; path deviation; environmental conditions (wind/current).
Expected Behaviour / Target KPIs / Values
KPI_017: Autonomous boat navigation precision (≥ 95% route-adherence) **KPI_018:** Area Coverage Efficiency (≥ 90% coverage)

Table 84: Portuguese Cluster Test Case #S2.5

Scenario 3 – Test Case #S3.1

Leading Partner: ALB	Test Case #S3.1: Real-time processing with AI/ML framework or the AI-App					
Test Case Objective	**TO #4:** Demonstrate the benefits of AI for performant and efficient 6G applications					
Test Case Target Layer or Application	IoT device energy management AI-App					
Test Environment	Field					
Test Deployment Setup	IoT Device, V-App, AI-App, AI/ML Ops Framework					
Network Setup	As described in section3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
The N-App/V-App are collecting data and passing it to the AI-App, and are ready to process the model results.						
Test Scenario						
The following steps are considered: 1. Process the incoming data 2. Generate the features dataset 3. Get the model to execute 4. Execute the model 5. Return the results to the N-App/V-App						
Test Variables						
Response time for model inference						
Expected Behaviour / Target KPIs / Values						

KPI_014: Inferencing speed (< 100 ms)

Table 85: Portuguese Cluster Test Case #S3.1

Scenario 3 – Test Case #S3.2

Leading Partner: ALB	Test Case #S3.2: Energy usage optimization and users' maximization					
Test Case Objective	**TO #5:** Optimize renewable energy integration with 6G-Apps' vertical infrastructure					
Test Case Target Layer or Application	IoT device energy management AI-App or AI/ML Framework					
Test Environment	Field					
Test Deployment Setup	IoT Device, V-App, AI-App, AI/ML Ops Framework					
Network Setup	As described in section3.5.1 Testbed Overview in D2.1.					
Test Configuration	**End Device Density**	N/A	**Mobility Level**	N/A	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
The N-App/V-App are collecting data and passing it to the AI-App, and are ready to process the model results.						
Test Scenario						
The following steps are considered: 1. Process the incoming data 2. Generate the features dataset 3. Get the model to execute 4. Execute the model 5. Return the results to the N-App/V-App 6. Calculate the amount of energy used to charge EV that was produced by solar panels						
Test Variables						
Amount of clean energy used to charge the vehicles						
Expected Behaviour / Target KPIs / Values						
KPI_015: Sustainable energy used in EV chargers (> 50 %) The results should maximize the amount of energy used that is generated by the solar panels and maximize the number of EV charged.						

Table 86: Portuguese Cluster Test Case #S3.2

Spanish Cluster

La Mayora. Test Case #1.1: Optimal water mixing with connected sensors

Leading Partner: CET	Test Case #1.1: Optimal water mixing with connected sensors
Test Case Objective	Confirm correct End-to-End performance
Test Case Target Layer or Application	End to End
Test Environment	Field (La Mayora)
Test Deployment Setup	• Water quality sensors with 5G modem connectivity in reservoirs.

	• App "DSS for water quality" ready (or a preliminary version). V-App for alert notification; AI-App for Decission System Support (DSS); N-App for sensors connectivity. • Integrated with 6G-SANDBOX toolkit, OpenCAPIF APIs, HPE core.					
Network Setup	• 5G SA coverage (FR1-B40) • HPE N-App provisioning • KPI monitoring via 6G-SANDBOX and other means.					
Test Configuration	**End Device Density**	Low	**Mobility Level**	No	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
• Connectivity validation						
Test Scenario						
Alarm reporting (1) Activate sensor reporting (2) AI processes information from the sensors (3) If the water quality is below the acceptable quality threshold for irrigation, it will trigger events to the V-App (4) V-App will notify final users by means of alarms Periodic quality reporting (1) At a predefined reporting period: (2) AI-App process information to the sensors and report to V-App (3) V-App reports a notification to the user, including water quality metrics						
Test Variables						
• Sensor related variables: Reporting frequency of sensors • Water quality thresholds: Acceptable threshold values for irrigation; Variation of threshold settings • Reporting variables: Format of report (numeric values, graphical trends, summary statistics) • Environmental / contextual variables: Environmental conditions (temperature, humidity, etc)						
Expected Behaviour / Target KPIs / Values						
• KPI_01 User Available Data Rate • KPI_03 Coverage • KPI_08 AI Accuracy • KPI_12 Sustainability. Water Analysis Accuracy • KPI_13 Sustainability. Water saving • KPI_14 Sustainability. CO2 Reduction • KPI_16 Sustainability. Renewable Powered Device Ratio						

Table 87: Spanish Cluster Test Case #1.1

La Mayora. Test Case #1.2: Drone-assisted monitoring and sample collection with multipath protocol

Leading Partner: UMA	Test Case #1.2: Drone-assisted monitoring and sample collection with multipath protocol
Test Case Objective	Confirm correct End-to-End performance in drone remote control
Test Case Target Layer or Application	End to End
Test Environment	Field (La Mayora)

Test Deployment Setup	• Low device density (1 drone) • High mobility (flight operations), with background traffic					
Network Setup	• 5G SA connectivity with multipath protocols enabled • KPI measurement through monitoring tools.					
Test Configuration	**End Device Density**	Low	**Mobility Level**	Drone Path	**Has Background Traffic**	Low / No
Initial Conditions / Prerequisites						
• Connectivity validated • Drone pre-flight checks and calibration • Multipath protocol						
Test Scenario						
This is the specific set of steps to follow during the test: 1. Deploy sensors and connect via 5G. 2. Launch drone for monitoring. 3. Remote control the drone through multiple paths. 4. Collect sample data.						
Test Variables						
• The environment where the drone is evaluated (e.g., emulated or real) • The configuration of the underlying multipath protocol						
Expected Behaviour / Target KPIs / Values						
• KPI_01 Throughput • KPI_02 Latency • KPI_03 Coverage • KPI_04 Packet Loss Rate						

Table 88: Spanish Cluster Test Case #1.2

La Mayora. Test Case #1.3: Immersive remote driving and autonomous AGV missions

Leading Partner: NOKIA	Test Case #1.3: Immersive remote driving and autonomous AGV missions					
Test Case Objective	Confirm correct End-to-End performance					
Test Case Target Layer or Application	End to End					
Test Environment	Field (La Mayora)					
Test Deployment Setup	• AGV with three streaming cameras • PC with immersive driving front-end • Hyperspectral vehicle camera for crop samples collection • Predefined mission routes					
Network Setup	• 5G SA coverage • KPI instrumentation • Immersive video uplink					
Test Configuration	**End Device Density**	Medium	**Mobility Level**	AGV Path	**Has Background Traffic**	No
Initial Conditions / Prerequisites						
• AGV functional check						

<table>
<tr><td>• Connectivity validation
• Vehicle cameras verification
• Open area for GPS signal
• Good environmental conditions (no rain essential)
• Pre-check coverage prediction with Network Digital Twin if available</td></tr>
<tr><th>Test Scenario</th></tr>
<tr><td>Interactive Remote Driving
1. Switch between autonomous and remote modes.
2. Start immersive remote driving session.
3. Monitor camera images.
4. Evaluate responsiveness and stability.
5. KPIs will be collected on parallel
Predefined missions:
1. Execute AGV predefined autonomous mission
2. AGV will navigate through the predefined mission path
3. Monitor camera images
4. KPIs will be collected on parallel</td></tr>
<tr><th>Test Variables</th></tr>
<tr><td>• Field physical area
• Missions route
• AGV environment cameras resolution</td></tr>
<tr><th>Expected Behaviour / Target KPIs / Values</th></tr>
<tr><td>• KPI_01 User Experienced Data Rate
• KPI_02 Latency
• KPI_03 Coverage</td></tr>
</table>

Table 89: Spanish Cluster Test Case #1.3

La Mayora. Test Case #1.4: Automated surveillance and incident detection with JCAS

<table>
<tr><th>Leading Partner: UMA</th><td colspan="6">Test Case #1.4: Automated surveillance and incident detection with JCAS</td></tr>
<tr><th>Test Case Objective</th><td colspan="6">Confirm correct End-to-End performance</td></tr>
<tr><th>Test Case Target Layer or Application</th><td colspan="6">End to End</td></tr>
<tr><th>Test Environment</th><td colspan="6">Field (La Mayora)</td></tr>
<tr><th>Test Deployment Setup</th><td colspan="6">Robots capturing visual data, JCAS signals (Sounding Reference Signals, SRS), AI-App detection, V-App notifications.</td></tr>
<tr><th>Network Setup</th><td colspan="6">5G SA connectivity with KPI monitoring.</td></tr>
<tr><th>Test Configuration</th><th>End Device Density</th><td>Medium</td><th>Mobility Level</th><td>Low</td><th>Has Background Traffic</th><td>Low</td></tr>
<tr><th colspan="7">Initial Conditions / Prerequisites</th></tr>
<tr><td colspan="7">Robots calibrated, AI model pre-trained, connectivity validated.</td></tr>
<tr><th colspan="7">Test Scenario</th></tr>
<tr><td colspan="7">1. Deploy robots in the field.
2. Capture visual and JCAS data.
3. Process in AI-App to detect incidents.
4. Trigger alarms and notifications.</td></tr>
</table>

5. Validate detection accuracy.	
Test Variables	
• The environment where detection is evaluated (lab environment and La Mayora). • SRS reporting parameters (period, PRBs)	
Expected Behaviour / Target KPIs / Values	
• KPI_03 Coverage • KPI_10 JCAS. Channel Estimation Accuracy • KPI_11 JCAS. Detection Accuracy	

Table 90: Spanish Cluster Test Case #1.4

EDAR Roquetas de Mar. Test case #2.1: DSS for reliable water treatment

Leading Partner: HIDRA	Test Case #2.1: DSS for reliable water treatment					
Test Case Objective	Confirm correct End-to-End performance					
Test Case Target Layer or Application	End to End					
Test Environment	Field (EDAR Roquetas de Mar)					
Test Deployment Setup	• Water quality sensors installed • Drone equipped with camera monitoring • App "DSS for reliable water treatment" ready. V-App for alert notification; AI-App for DSS; N-App for sensors connectivity. • AI-App analysing images/data • V-App visualization • Open Gateway API for network reliability					
Network Setup	• 5G SA with OpenGateway APIs (Quality on Demand, Dynamic Airspace Connectivity) and network slicing.					
Test Configuration	**End Device Density**	Medium	**Mobility Level**	Medium	**Has Background Traffic**	Yes
Initial Conditions / Prerequisites						
• Sensor calibration • Drone validation • Minimum quality images • External/internal cameras if needed • Open Gateway API setup.						
Test Scenario						
Alarm reporting 1. Activate sensor reporting 2. AI processes information from the sensors 3. If a process at the wastewater treatment plant is found to be malfunctioning, the AI-App will trigger an event to the V-App 4. V-App will notify final users by means of alarms, including the data from the sensor that made the alarm trigger (for final human decision)						
Test Variables						
• AI-App variables: AI model version/algorithm: baseline vs. updated model; Decision threshold sensitivity: strict vs. relaxed alarm thresholds.						

• V-App / user notification: Notification channel: push notification, Short Message Service (SMS), email. • Environmental / operational: Time of day / lighting conditions (for drone camera): daylight vs. low light.
Expected Behaviour / Target KPIs/ Values
• KPI_03 Coverage • KPI_05 Accessibility. Connection Success • KPI_08 AI Accuracy • KPI_12 Sustainability. Water Analysis Accuracy • KPI_15 Sustainability. Time Saving

Table 91: Spanish Cluster Test Case #2.1

EDAR Roquetas de Mar. Test case #2.2: Open Gateway Quality Performance

Leading Partner: TID	Test Case #2.2: Open Gateway Quality Assurance					
Test Case Objective	Confirm correct performance of Open Gateway APIs Activation					
Test Case Target Layer or Application	Application					
Test Environment	Lab and Field (EDAR Roquetas de Mar)					
Test Deployment Setup	5G SA Telefonica commercial network 5G Device SIM Card provisioning					
Network Setup	5G SA with OpenGateway APIs (Quality on Demand, Dynamic Airspace Connectivity) and network slicing.					
Test Configuration	**End Device Density**	Medium/High	**Mobility Level**	Low	**Has Background Traffic**	Yes
Initial Conditions / Prerequisites						
• Commercial nodes with 5G SA configuration • Network Slicing available • SIMs provisioned for Open Gateway activation						
Test Scenario						
Identify traffic conditions under which Open Gateway services need to be activated to assure the quality of the AGV flight over the commercial network. 1. Prior to initiating the drone flight, perform a check of air quality conditions using onboard sensors and ensure compliance with safety thresholds. 2. Continuously monitor the network quality parameters during the test. If congestion is detected, use Open Gateway mechanisms to prioritize the communication flows of the components involved in the test case, and then continue using the measurement tool to evaluate the KPIs						
Test Variables						
• Network Slice • 5G Quality of Service Identifier (5QI)						
Expected Behaviour / Target KPIs / Values						
• KPI_01 User Experienced Data Rate • KPI_02 Latency • KPI_05 Packet Loss Rate • KPI_06 PDU Session Success						

- KPI_07 PDU Session Reject

During the test, the conditions should be complemented with information on radio conditions and network quality, ensuring that the flight is supported by reliable connectivity parameters. In addition, a prioritization of quality for the SIM devices associated with the project will be applied through the activation of QoD mechanisms. In cases of network congestion or degraded throughput, the UE with QoD activated will be prioritised, ensuring continuity of the AGV/drone operation and demonstrating the effectiveness of Open Gateway services for critical communications.

Table 92: Spanish Cluster Test Case #2.2

Lab. Test Case #3.1: Trial/Lab-based validation modem energy consumption features

<table>
<tr><td>Leading Partner: KEYS</td><td colspan="6">Test Case #3.1: Lab-based validation of RedCap and energy consumption features</td></tr>
<tr><td>Test Case Objective</td><td colspan="6">Confirm correct End-to-End performance</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">End to End</td></tr>
<tr><td>Test Environment</td><td colspan="6">Lab</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">• Core network instance
• RedCap/NTN features (if adopted)
• Measurement tools for energy consumption</td></tr>
<tr><td>Network Setup</td><td colspan="6">• Lab-based 5G SA instance
• Instrumentation for energy monitoring</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>Low</td><td>Mobility Level</td><td>No</td><td>Has Background Traffic</td><td>No</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">• Lab setup configured
• 5G/RedCap/NTN features available
• Measurement tools calibrated</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">1. Activate core and RAN test instances with 5G/RedCap/NTN capabilities.
2. Establish controlled radio configuration parameters.
3. Generate controlled traffic loads.
4. Measure modem energy consumption.</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">• UE transmission power
• UE traffic pattern: transmission interval and packet size
• Technology under test. 5G/RedCap/NTN

At the time of reporting this deliverable it is deemed unlikely that RedCap will be evaluated in the Spanish cluster provided dependencies with the deployed test network. However the testing capabilities in the KEYS UXM network emulator would allow this characterization if finally planned to be adopted.</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">• KPI_017 UE Energy Consumption</td></tr>
</table>

Table 93: Spanish Cluster Test Case #3.1

Lab. Test Case #3.2: AI training and inference energy profiling.

Leading Partner: KEYS/CET	Test Case #3.2: Energy consumption for model training

<table>
<tr><td>Test Case Objective</td><td colspan="6">In the lab, this test could be used to resemble an equivalent workload to that used during the actual AI training and inference for the trials. This test would be used to gather relevant information to better assess the energy footprint of the trials.</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">AI</td></tr>
<tr><td>Test Environment</td><td colspan="6">Lab</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">Water treatment plant process monitoring (necessary to construct the dataset)

DSS for water quality/DSS for reliable water treatment Application, running on test servers at UMA facilities.

JCAS application training and inference, in lab.</td></tr>
<tr><td>Network Setup</td><td colspan="6">N/A, this setup aims to characterise the AI processing sustainability.</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>N/A</td><td>Mobility Level</td><td>N/A</td><td>Has Background Traffic</td><td>N/A</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">A dataset of labelled images will be generated in advance from the cameras deployed in the water treatment plant at Roquetas de Mar.
The dataset will have been modified to remove any potential images that could contain privacy sensitive information from water plant workers or other persons.
The AI algorithm must be previously implemented and deployed in the test station by Cetaqua for training in this test
Power measurement equipment will be deployed and connected to the training station</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test. In a first step:
Power monitoring equipment will be used to obtain an energy floor baseline without actual ongoing training
In a second phase:<ol start="2">
Consumed power will be sampled during the training phase for each applicable training process
Training will be conducted while the measurements continue
Collected data will be stored and compared with the original baseline to estimate the energy consumption of the model training</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">Model and training configuration: Image resolution; Training regime: from scratch vs. transfer learning vs. fine-tuning
Dataset and pipeline: Dataset size: number of images / samples used</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">KPI_09 AI.Training Energy</td></tr>
</table>

Table 94: Spanish Cluster Test Case #3.2

Lab. Test Case #3.3: Core network energy profiling

In the lab, we would include this activity to help better define the energy footprint of some network components such as the core network which are not typically addressed. This will not make claims on generic performance (as indicated by HPE), but will collect selected data obtained under specific working conditions and hardware choices, resembling what could be experienced during the trials.

<table>
<tr><th>Leading Partner: KEYS/HPE</th><th colspan="6">Test Case #3.3: Core network energy profiling</th></tr>
<tr><td>Test Case Objective</td><td colspan="6">Assess the power consumption of the core network under controlled test conditions in Lab environment</td></tr>
<tr><td>Test Case Target Layer or Application</td><td colspan="6">Network Core</td></tr>
<tr><td>Test Environment</td><td colspan="6">Lab</td></tr>
<tr><td>Test Deployment Setup</td><td colspan="6">LoadCore / Core Network Under Test / AC Power Analyzer</td></tr>
<tr><td>Network Setup</td><td colspan="6">Victoria Network</td></tr>
<tr><td>Test Configuration</td><td>End Device Density</td><td>N/A</td><td>Mobility Level</td><td>N/A</td><td>Has Background Traffic</td><td>NOT</td></tr>
<tr><td colspan="7">Initial Conditions / Prerequisites</td></tr>
<tr><td colspan="7">• Core network deployed in lab environment
• LoadCore solution is deployed and configured
• AC power analyser is physically connected to the server hosting the core network</td></tr>
<tr><td colspan="7">Test Scenario</td></tr>
<tr><td colspan="7">This is the specific set of steps to follow during the test:
1. Core network is successfully initialized
2. LoadCore configuration is activated to emulate RAN and connect to core network
3. Power Analyzer sample acquisition is started
4. For different traffic profiles (throughput and number of connected users)
A) LoadCore emulated users attach to the core network
B) Activate traffic generation at both ends of the core network resembling traffic patterns utilized during the trials
C) Measure traffic in both DL and UL and derive metrics using LoadCore
D) Core network metrics are collected
E) Power samples are collected from Power Analyzer and associated to the configuration under test.</td></tr>
<tr><td colspan="7">Test Variables</td></tr>
<tr><td colspan="7">• Number of connected UEs
• Total Aggregated Throughput</td></tr>
<tr><td colspan="7">Expected Behaviour / Target KPIs / Values</td></tr>
<tr><td colspan="7">It is expected that the core network will serve all users.
Collected KPIs:
• KPI_018: Sustainability. Core Network Energy</td></tr>
</table>

Table 95: Spanish Cluster Test Case #3.3