The EXTREMA Simulation Hub: A facility for early V&V and community uptake

G. Di Domenico[1], F. Ornati[2], D. Perico[3], and F. Topputo[4]

Dept. of Aerospace Science and Technology, Politecnico di Milano
Via Giuseppe La Masa 34, 20156, Milano, Italy

ABSTRACT

The increasing pace of space missions and the growing adoption of small spacecraft require higher levels of onboard autonomy, particularly in Guidance, Navigation, and Control (GNC). Spacecraft must operate with limited ground contact, navigate uncertain environments, and make time-critical decisions independently. In this context, early and continuous Verification and Validation (V&V) of autonomous systems become essential.

The ERC-funded EXTREMA project addresses this challenge through the EXTREMA Simulation Hub (ESH), a distributed hardware-in-the-loop (HIL) environment where autonomous GNC software interacts in real time with representative sensors, actuators, onboard computing, and high-fidelity dynamics. Facilities such as RETINA, ETHILE, STASIS, ELAPSE, and the SPESI real-time propagator create a digital-physical twin for realistic system-level testing.

Building on this infrastructure, AXESS (Accelerated X-in-the-loop Environment for Spacecraft Systems Testing) extends ESH into a Testing-as-a-Service platform, providing secure remote access to configurable HIL scenarios and space-tailored CI/CD workflows. AXESS enables continuous, repeatable, and data-driven V&V for researchers, startups, industry, and agencies, reducing development risks, accelerating innovation, and fostering a shared ecosystem for autonomous spacecraft validation.

1 INTRODUCTION

The space sector is undergoing a structural transformation, shifting from a low-volume, institution-driven paradigm to a high-frequency, commercially driven ecosystem. The increasing deployment of small satellites and the emergence of deep-space CubeSat missions are redefining expectations in terms of mission cadence, responsiveness, and reliability. Market analyses indicate that the bottleneck in the space value chain has transitioned from launch availability to system validation and reliability.

In this context, spacecraft autonomy, particularly in Guidance, Navigation, and Control (GNC), is becoming essential [1]. Autonomous systems must operate under limited ground contact, handle uncertain environments, and execute time-critical decisions onboard. This is especially relevant for deep-space missions, where communication delays inherently preclude real-time human intervention [2]. Despite these needs, current Validation and Verification (V&V) practices remain largely incompatible with agile development cycles. Traditional testing approaches rely on late-stage integration and expensive, centralized facilities, limiting iteration speed and increasing risk exposure.

---

[1] Post-Doc Researcher, gianfranco.didomenico@polimi.it
[2] PhD Candidate, fabio.ornati@polimi.it
[3] PhD Candidate, davide.perico@polimi.it
[4] Full Professor, francesco.topputo@polimi.it

The cost of late defect detection can increase by an order of magnitude compared to early-stage identification.

The EXTREMA project addresses this gap of testing paradigm by combining algorithmic research with experimental validation. Central to this effort is the EXTREMA Simulation Hub (ESH)**,** a distributed hardware-in-the-loop (HIL) environment designed to enable early, continuous, and integrated V&V of autonomous spacecraft systems.

1.1 The EXTREMA FRAMEWORK

The EXTREMA project is structured around three research pillars that collectively address the problem of achieving fully autonomous deep-space navigation and control. These pillars are not conceived as independent research tracks, but rather as complementary components of a unified autonomy stack. In fact, their outputs are progressively integrated and validated within a common experimental framework.

The first pillar focuses on autonomous navigation, targeting the development of algorithms capable of enabling a spacecraft to estimate its state in deep space without reliance on ground-based measurements. This represents a fundamental departure from traditional mission operations, where navigation is largely performed on Earth through radiometric tracking [3]. In EXTREMA, the emphasis is instead placed on onboard sensing, with particular attention to vision-based navigation techniques suitable for CubeSat-class platforms. The objective is to allow the spacecraft to determine its position and velocity by processing optical observations of celestial bodies, thus enabling fully self-navigation capabilities.

Building upon the navigation outputs, the second pillar addresses autonomous guidance and control. In current mission architectures, trajectory design and maneuver planning are typically performed on the ground due to computational limitations onboard. This leads to significant operational overhead and delays, especially in deep-space missions where communication windows are sparse. EXTREMA aims to overcome this limitation by developing lightweight, onboard guidance algorithms capable of computing and updating thrust profiles in real time. These algorithms leverage the state estimates provided by the navigation module to autonomously determine optimal trajectories and correction maneuvers, thus closing the loop between perception and action directly onboard the spacecraft.

The third pillar introduces the concept of autonomous ballistic capture as an alternative to traditional orbit insertion strategies. Conventional capture maneuvers require precise timing and significant propulsive effort, which are challenging to achieve with low-thrust propulsion systems typically employed by small spacecraft. EXTREMA exploits the natural dynamics of multi-body systems to identify regions in the state space, referred to as capture corridors, where a spacecraft can be temporarily captured by a celestial body with minimal propulsion. By autonomously navigating within these corridors, the spacecraft can achieve stable or quasi-stable orbits with reduced energy expenditure. This approach is particularly well suited to CubeSat missions, where propulsion capabilities are inherently limited.

A key aspect of the EXTREMA methodology is that the outputs of these three pillars are not validated in isolation. Instead, they are progressively integrated into increasingly complex experimental scenarios, ultimately converging in a unified system-level validation environment. This integration is essential, as the behavior of autonomous systems emerges from the interaction between estimation, decision-making, and actuation, rather than from the performance of individual components [4].

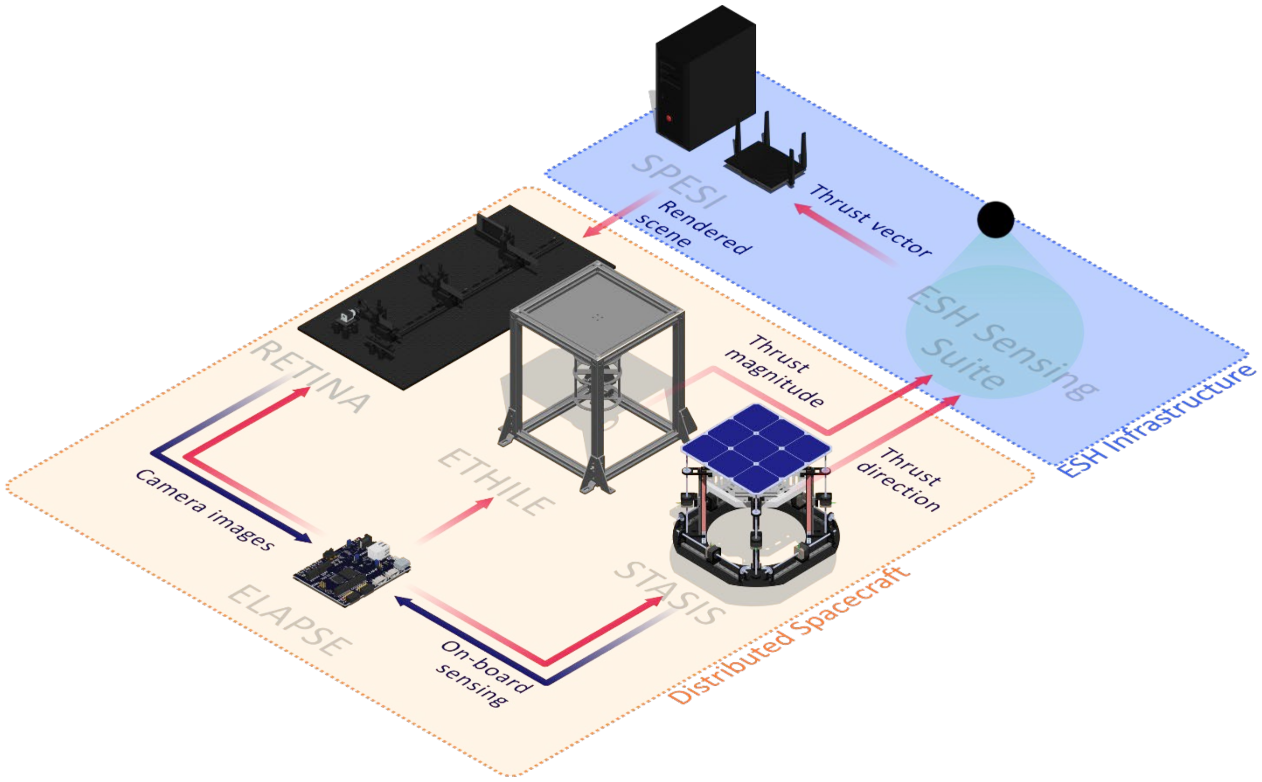


Figure 1. Schematics of the ESH simulation data flows and facilities.

The EXTREMA Simulation Hub (ESH) represents the experimental context of the project, where the algorithms developed within the three pillars are integrated and tested in a realistic, closed-loop environment. The ESH is conceived as a distributed hardware-in-the-loop (HIL) testing facility in which autonomous GNC systems interact with physical and virtual components in a manner that closely reproduces the conditions of an interplanetary mission.

At its core, the ESH combines multiple specialized experimental setups, each reproducing a specific subsystem of a spacecraft, as reported in Figure 1. The optical domain is addressed by RETINA [5], [6]

The actuation domain is represented by ETHILE, a thrust-in-the-loop test bench that emulates the behavior of spacecraft electric propulsion systems. Rather than relying on purely numerical models, ETHILE allows the guidance and control algorithms to interact with a physical representation of the propulsion subsystem, capturing effects such as actuation delays, control discretization, and hardware constraints. This is particularly relevant for low-thrust electric propulsion, where continuous actuation and tight coupling with guidance algorithms play a critical role.

Attitude dynamics are reproduced by STASIS [7], an air-bearing platform capable of simulating near-frictionless rotational motion. By hosting sensors and control hardware on a freely rotating structure, STASIS enables realistic testing of attitude determination and control algorithms. The platform is complemented by a high-accuracy ground truth system, which provides precise measurements of the spacecraft orientation and allows for rigorous validation of estimation and control performance.

These physical facilities are coordinated by a central computing unit implemented as a FlatSat system, which hosts the EXTREMA fLAtsat Processing SystEm (ELAPSE) [8]. ELAPSE acts as an onboard computer emulator, executing navigation, guidance, and control algorithms in an environment that closely replicates the constraints of flight hardware. The software architecture is designed to be modular and extensible, enabling rapid integration of new algorithms and facilitating iterative development within a continuous testing framework.

The dynamic evolution of the spacecraft is handled by SPESI [9], a real-time stochastic orbital propagator that acts as the glue between all components of the simulation. SPESI collects measurements from the physical facilities, such as thrust vectors from ETHILE and attitude information from STASIS [10], and integrates them to propagate the spacecraft state. At each simulation step, it also generates the corresponding deep-space scene, which is fed back to the optical facility, thereby ensuring consistency between the simulated environment and the physical sensing chain.

A distinguishing feature of the ESH is the tight coupling between all subsystems within a real-time simulation loop. The onboard computer processes sensor data, estimates the spacecraft state, computes control actions, and commands the hardware components, whose outputs are then fed back into the simulation. This closed-loop architecture enables the emergence of realistic system-level behaviors, including delays, nonlinearities, and cross-coupling effects, which are typically difficult to capture in purely numerical simulations without building complex and computationally heavy models.

Achieving meaningful hardware-in-the-loop validation requires strict synchronization between physical components and numerical nodes. In the ESH, this is ensured through a real-time execution framework in which all operations within a simulation step must be completed within a fixed time budget. This includes sensor data acquisition, state estimation, control computation, dynamics propagation, and scene rendering. The requirement for hard real-time execution imposes significant constraints on both software design and hardware infrastructure, but it is essential to guarantee consistency between the simulated and physical domains.

To address the lengthy time scales associated with interplanetary missions, the ESH incorporates an accelerated simulation framework that preserves the dynamical properties of the system while reducing the duration of experiments [11]. This allows the validation of complete mission scenarios, including long-duration transfers and capture phases, within practical timeframes. The ability to perform repeated simulations under varying conditions is crucial for assessing the robustness of autonomous algorithms and for identifying rare or critical behaviors.

Finally, the entire simulation environment is monitored and controlled through MoniCA, a unified software interface that aggregates telemetry from all devices and provides real-time visualization and command capabilities. MoniCA enables users to observe the internal state of the system, interact with the simulation, and collect data for post-processing and analysis, thereby completing the experimental infrastructure of the ESH. The overall facility can be appreciated in Figure 2.

2 TESTING-AS-A-SERVICE

The EXTREMA Simulation Hub has demonstrated the technical feasibility and advantages of early, integrated hardware-in-the-loop validation. Specifically, it has been employed to simulate complete deep-space interplanetary cruises performed by a self-driving spacecraft.

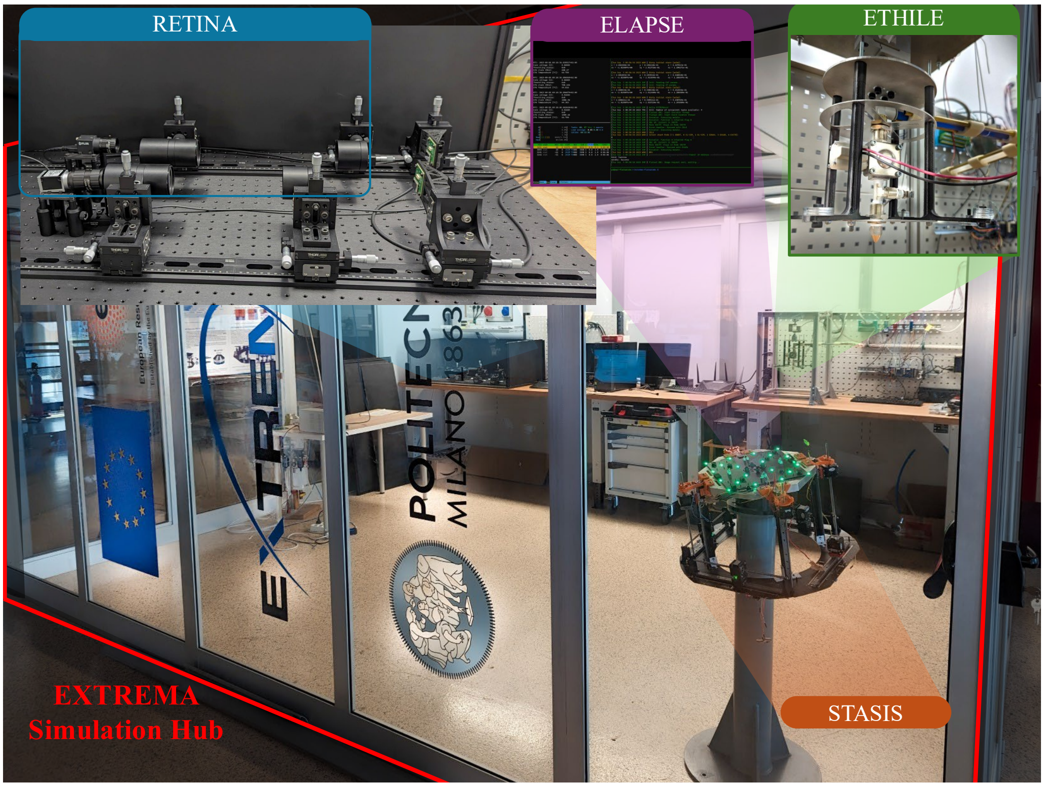


Figure 2. Facilities within the EXTREMA Simulation Hub.

However, its modularity and the Agile philosophy adopted in the integration of autonomous guidance and navigation algorithms in EXTREMA pave the way for a wider range of applications and testing capabilities.

Recent market analyses highlight a structural imbalance between the rapid growth in spacecraft development and the availability of adequate testing and validation infrastructure. The proliferation of CubeSats and small satellite missions, driven by reduced launch costs and increased private investment, has significantly shortened development cycles. However, V&V practices have not evolved at the same pace. Many organizations, particularly startups and small teams, lack access to high-fidelity test environments, relying instead on software-only simulations or late-stage system testing [12]. This mismatch contributes to elevated mission failure rates and integration risks, especially for systems incorporating advanced autonomy [13].

Furthermore, the increasing complexity of onboard software, coupled with the adoption of agile development methodologies, creates a need for continuous testing frameworks analogous to those used in terrestrial software engineering. In other domains, continuous integration and deployment (CI/CD) pipelines enable rapid iteration and systematic validation of code changes [14]. In the space sector, however, the cyber-physical nature of spacecraft systems, where software interacts tightly with hardware and dynamics, makes it difficult to directly transfer these practices without appropriate infrastructure [15].

AXESS (Accelerated X-in-the-loop Environment for Spacecraft Systems Testing) is conceived as a response to this gap. Its objective is to transform the capabilities demonstrated within ESH into a scalable, remotely accessible platform that enables continuous, high-fidelity V&V for a wide range of users.

2.1 AXESS Architecture

AXESS extends the ESH paradigm by introducing a service-oriented framework, depicted in Figure 3. Thanks to the new paradigm, users can remotely access and configure hardware-in-the-loop experiments through standardized APIs. Monitoring of the experiments and collection of the results is enabled by the service interface.

Therefore, direct interaction with the ESH facility is generalized to the execution of test scenarios on a subset of hardware test benches. Upon users request and configuration, the AXESS framework orchestrates the underlying systems with the aim of running the desired tests and reporting back to the user.

The overall system architecture is composed of different modules which can be appreciated in the scheme in Figure 4.

At the core of AXESS TaaS is the orchestration layer, which coordinates the execution of experiments across distributed assets. This layer manages resource allocation, scheduling, and synchronization, ensuring that multiple users can access the system without compromising real-time performance or polluting results. The orchestration framework interfaces with the existing ESH components (i.e., RETINA, ETHILE, STASIS, ELAPSE, and SPESI), exposing their functionalities through well-defined APIs.

Users interact with the platform via a web-based interface, where they can define test campaigns by specifying mission scenarios, initial conditions, and performance metrics. These configurations are translated into executable pipelines that deploy user-provided software onto the ELAPSE FlatSat, configure the relevant hardware components, and initiate the simulation loop. During execution, telemetry data is streamed in real time through the MoniCA interface, allowing users to monitor system behavior and intervene if necessary.

A key aspect of the architecture developed is the abstraction of hardware complexity. Users are not required to manage low-level details of the experimental setup. Instead, they can focus on the behavior of their algorithms within a controlled environment. In this way, also reproducibility is easily guaranteed because execution leverages the same conditions. The abstraction introduced is essential, especially among users that do not have direct and low-level experience hardware systems.

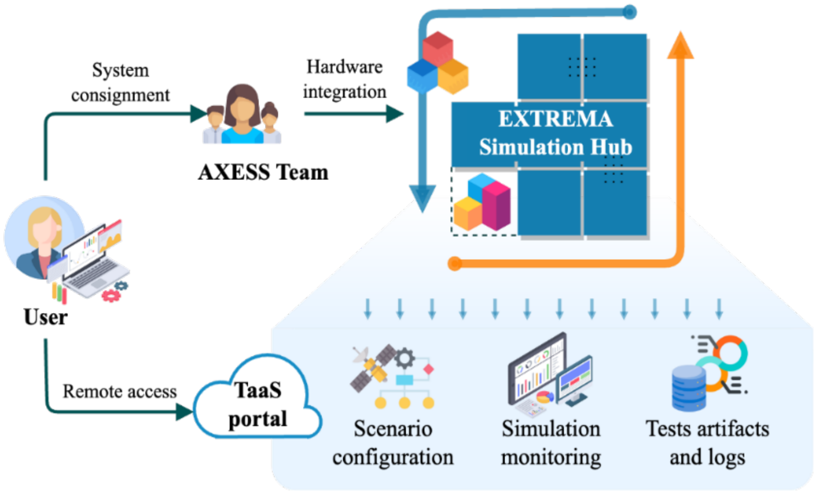


Figure 3. Scheme of the AXESS Test-as-a-Service paradigm.

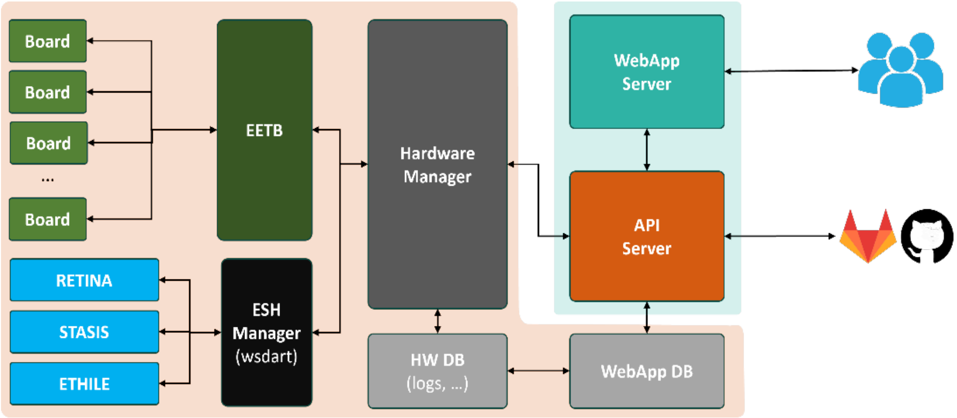


Figure 4: Scheme of the AXESS configurable system architecture.

2.2 Data Management, Metrics, and Benchmarking

A central component of the AXESS vision is the systematic collection and exploitation of experimental data. Each test campaign generates a rich set of telemetry, including sensor measurements, state estimates, control actions, and ground truth data from SPESI. This data is stored and organized to support both immediate analysis and long-term knowledge accumulation.

The availability of standardized datasets enables the definition of common performance metrics across different users and applications [16], [17]. These metrics may include navigation accuracy, control effort, convergence time, robustness to disturbances, and computational load. By evaluating algorithms against shared benchmarks, AXESS facilitates objective comparison and accelerates technological maturation.

Moreover, the accumulation of experimental data across multiple users and scenarios opens the possibility of identifying trends and best practices. For example, recurring failure modes or performance bottlenecks can be systematically analyzed, informing both algorithm design and system architecture. This collective knowledge represents a key added value of the platform, transforming individual test campaigns into contributions to a broader ecosystem.

2.3 Accessibility, Scalability, and Security

To achieve its goal of democratizing access to high-fidelity testing, AXESS must address challenges related to scalability and security. From a scalability perspective, the platform must support multiple concurrent users while maintaining real-time performance. This requires careful scheduling of hardware resources, as well as the ability to partition experiments in time and space.

Security is equally critical, particularly in a multi-user environment where proprietary algorithms and sensitive data are involved. AXESS incorporates mechanisms for secure access control, data isolation, and encrypted communication, ensuring that users can operate with confidence in the integrity and confidentiality of their experiments.

In addition, the platform is designed to be extensible, allowing new hardware components and simulation capabilities to be integrated over time. This ensures that AXESS can evolve alongside technological advancements and continue to meet the needs of the space community.

3 PRELIMINARY DEMONSTRATIONS

To demonstrate the platform, two application examples, or demo, are here presented.

3.1 Optical testbench

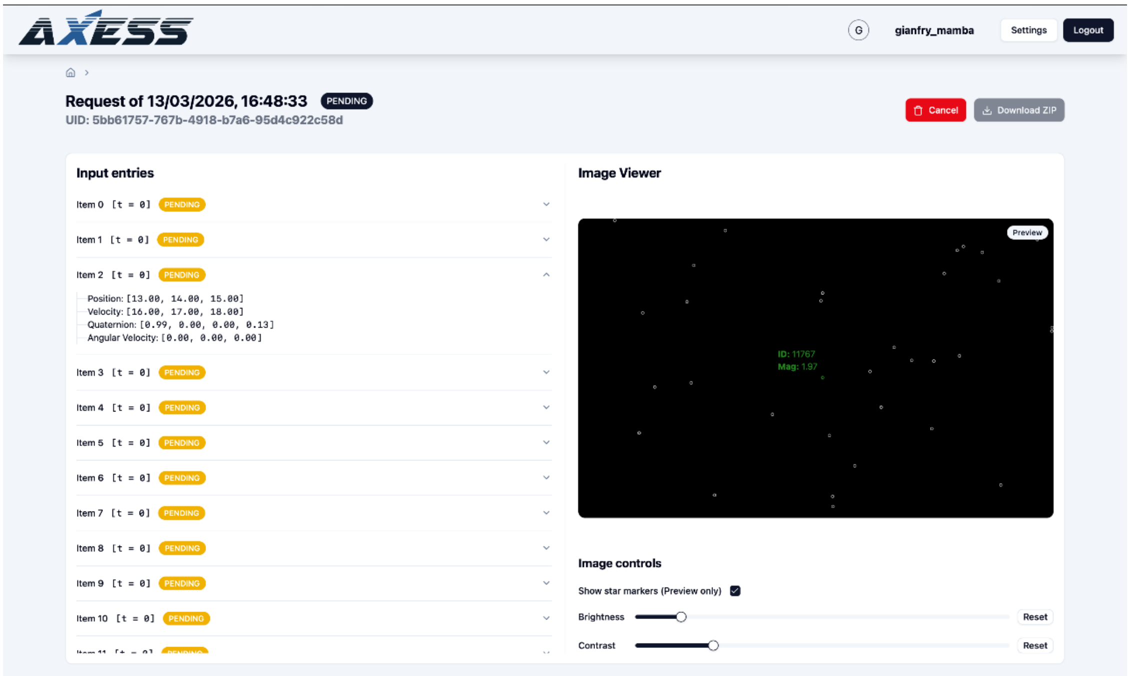
The first demonstrative platform developed within the AXESS project is the one concerning the access to the RETINA facility within the ESH. RETINA is a variable-magnification optical stimulator specifically designed to emulate deep-space navigation scenarios, such as the far-range scenarios encountered during the operational scenario envisioned by the EXTREMA project. In this navigation regime, the planets and stars are unresolved, appearing as small blobs of pixels in which it is not possible to distinguish any kind of morphological or geometrical feature. RETINA is composed of a double lens system that allows the modification of the magnification and therefore enables the deployment of cameras with different angular Field of View (FOV) without requiring hardware modifications [6].

To demonstrate the versatility of the RETINA facility and its optical emulation system, the AXESS project proposes to enable direct access to the testbed via its platform. In this context, this demo serves as demonstrator of streamlined access to hardware-based facilities through modern digital interfaces, enabling new testing workflows.

The Demo is accessible via a web-based UI which allows submitting remote requests for render-cast-acquire optical cycles for deep-space night sky conditions. The output of this process is constituted by images with excellent emulation fidelity that leverage both the state-of-the-art geometrical and radiometric calibration of the facility [6], [18]. The overall workflow is the following:

1. The user submits a list of camera position, velocity and inertial attitude orientation a in deep space using the UI.
2. The images are pre-processed, to retrieve a preview of each scene. The preview is generated via the rendering pipeline and does not include real camera effects but allows the user to see the ID of the stellar and planetary object that will be observable in the scene.
3. Once the user confirms the acquisition, the scenes are processed in a queue. For each image:
    a. The scene is rendered according to the radiometric and geometric calibration of the facility.
    b. The scene is displayed on the RETINA screen.
    c. The camera acquires the associated image.
    d. The image is retrieved by the coordinator.
4. Once the acquisition is completed, the images are packed in a .zip file and are automatically downloaded via the web interface.

The demo has been deployed online, allowing complete remote access to the facility and allowing the execution of tests at different locations without the need of on-site presence. Access is currently granted on a per-request basis. Figure 4 shows an example of the demo UI, with the image preview and request queue.



Request of 13/03/2026, 16:48:33 PENDING
UID: 5bb61757-767b-4918-b7a6-95d4c922c58d

Input entries

Item 0 [t = 0] PENDING
Item 1 [t = 0] PENDING
Item 2 [t = 0] PENDING
- Position: [13.00, 14.00, 15.00]
- Velocity: [16.00, 17.00, 18.00]
- Quaternion: [0.99, 0.00, 0.00, 0.13]
- Angular Velocity: [0.00, 0.00, 0.00]

Item 3 [t = 0] PENDING
Item 4 [t = 0] PENDING
Item 5 [t = 0] PENDING
Item 6 [t = 0] PENDING
Item 7 [t = 0] PENDING
Item 8 [t = 0] PENDING
Item 9 [t = 0] PENDING
Item 10 [t = 0] PENDING

Image Viewer

Image controls

Show star markers (Preview only) ☑
Brightness
Contrast

Figure 5. Screenshot of the AXESS Demo 1 web-based user interface.

3.2 Embedded software testbench

A pivotal step in the mission design testing phases lies in evaluating algorithms to certify their maturity for onboard deployment through a comprehensive verification and validation process that involves embedded boards deployment [19], [20]. Specifically, an incremental approach envisions the utilization of processor-in-the-loop (PIL) testing [21], [22]. However, as previously discussed, the internal availability of processing boards can be limited by budget and timing constraints. In fact, setting up and operating embedded systems is traditionally a poorly automated process, with a steep learning curve.

In this context, one of the most transformative aspects of the system is the adoption of DevOps as a set of approaches to foster the continuity of the validation phase of spacecraft subsystems. Specifically, Continuous Integration Continuous Deployment (CI/CD) is applied to shape the experiments execution and retrieval of results. As Figure 1 illustrates, the pipelines are configured and triggered by the user, since the requirements are imposed by the providers. In this way, even a change to the software under testing can trigger the automated testing pipeline. Currently, System-on-chip (SoC) boards mounting Zynq-7000 and Ultrascale+ chips have been employed to demonstrate the pipeline functionality, whereas the CI/CD framework is governed by Gitlab Runners. The current boards portfolio represents high-performance options which are particularly relevant as enabling technologies for onboard systems that aim to support higher autonomy functions, such as onboard trajectory guidance [23].

Once configured, these pipelines typically include the following steps:

1. Deployment of the piece of software into the ELAPSE environment, in bare-metal, or under a Unix OS;
2. Execution of predefined HIL scenarios;
3. Collection of simulation artifacts such as logs, reports and performance metrics;
4. Optional comparison against baseline results.

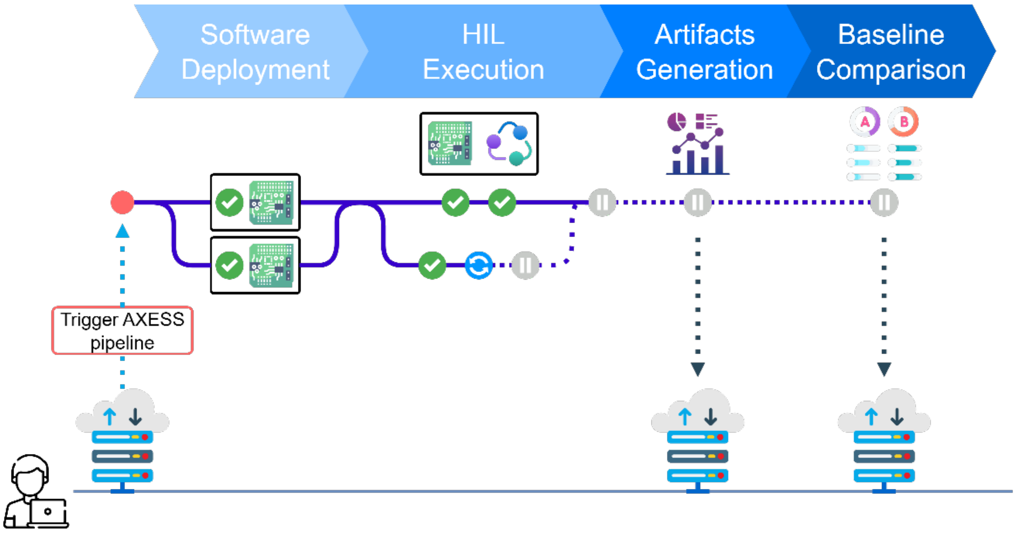


Figure 6: Schematic illustrating the jobs of the AXESS pipeline for embedded deployment.

By automating this process, AXESS enables continuous verification of system behavior under realistic conditions. This is the realization of a Runtime Verification (RV) environment which is particularly important for autonomous systems. where small changes in algorithmic parameters can lead to significant differences in emergent behavior. The system shall be tested not only to measure its performance in achieving certain goal conditions. Additionally, the RV framework aims to certify that the intent of the designer is satisfied, and the system is acting within the operational and safety bounds.

The integration of CI/CD workflows also supports reproducibility and traceability. Each test run is associated with a specific software version, configuration, and dataset, allowing results to be systematically compared and audited. This capability is essential not only for internal development, but also for external validation and certification processes.

Importantly, AXESS adapts these practices to the constraints of space systems. Unlike purely software environments, HIL testing involves limited physical resources and strict timing requirements. As a result, the platform must balance automation with resource management, ensuring efficient utilization of the infrastructure while maintaining deterministic execution.

4 MARKET IMPACT

The introduction of AXESS must be understood within the broader context of structural limitations highlighted in the market analysis. The rapid expansion of the small satellite sector has not been matched by a proportional increase in testing and validation capabilities. As a result, access to high-fidelity facilities remains limited, often characterized by long scheduling queues, high operational costs, and a lack of flexibility for iterative development cycles.

This mismatch is particularly critical for missions relying on autonomous GNC. Unlike traditional subsystems, autonomy cannot be fully validated through analytical methods or isolated simulations, as its behavior emerges from the interaction between perception, decision-making, and actuation under dynamic conditions. The absence of accessible, system-level testing environments therefore creates a bottleneck in the maturation of these technologies.

AXESS directly addresses this gap by introducing a distributed and service-oriented model for hardware-in-the-loop testing. By decoupling user access from physical presence and enabling remote interaction with experimental facilities, it significantly increases the effective availability of high-fidelity validation infrastructure. This shift is analogous to the transition observed in cloud computing, where shared resources enable scalable and on-demand access to computational capabilities.

4.1 Enabling New Development Workflows

One of the most significant impacts of AXESS lies in its ability to reshape development workflows for space systems. Traditional approaches are largely sequential, with design, implementation, and testing occurring in distinct phases. This structure is increasingly misaligned with the needs of modern missions, where rapid iteration and continuous refinement are essential.

The Testing-as-a-Service paradigm enables a transition toward iterative and data-driven development. By integrating hardware-in-the-loop validation into continuous integration pipelines, AXESS allows teams to test early and often, reducing the risk of late-stage integration issues. This capability is particularly valuable for startups and small teams, which often operate under tight resource constraints and cannot afford extensive redesign cycles.

Moreover, the ability to run automated test campaigns under varying conditions supports a more systematic exploration of the design space. Instead of validating a single nominal scenario, developers can assess performance across a wide range of uncertainties, leading to more robust and resilient systems. This approach aligns with emerging best practices in safety-critical software development, where extensive testing under diverse conditions is essential to ensure reliability.

4.2 Lowering Barriers to Entry and Supporting Innovation

The market analysis emphasizes that a significant portion of innovation in the space sector is driven by small and medium-sized enterprises, as well as research institutions. However, these actors often face substantial barriers when it comes to accessing advanced testing infrastructure. The cost and complexity of building and operating HIL facilities are prohibitive, leading to reliance on simplified simulations that may not capture critical system-level effects.

By providing access to a shared, high-fidelity testing environment, AXESS lowers these barriers and enables a broader range of actors to develop and validate advanced technologies. This democratization of access is expected to have a direct impact on innovation, as it allows more teams to experiment with novel approaches and rapidly iterate on their designs.

In addition, AXESS supports collaboration between academia and industry by providing a common platform where algorithms and systems can be tested under comparable conditions. This facilitates technology transfer and reduces the gap between research and operational deployment.

4.3 Standardization and Benchmarking Across the Community

A recurring challenge in the current ecosystem is the lack of standardized validation procedures and performance metrics. Different organizations often rely on proprietary testing setups and evaluation criteria, making it difficult to compare results or assess the maturity of competing solutions. This fragmentation slows down technology adoption and increases uncertainty for decision-makers.

AXESS addresses this issue by enabling the definition and dissemination of standardized test scenarios and benchmarks. By providing a common reference framework, it allows different approaches to be evaluated under identical conditions, improving transparency and comparability. This is particularly important for autonomy algorithms, where performance can be highly sensitive to environmental assumptions and implementation details.

The establishment of shared benchmarks also supports the development of certification pathways for autonomous systems. As regulatory bodies increasingly focus on the validation of onboard decision-making capabilities, the availability of standardized testing environments and metrics will be essential for demonstrating compliance and building trust.

4.4 Data-Driven Ecosystem and Knowledge Accumulation

Beyond individual test campaigns, AXESS enables the creation of a data-driven ecosystem in which experimental results are systematically collected and analyzed. Each experiment contributes to a growing repository of data describing system behavior under a wide range of conditions. This accumulation of knowledge has several important implications.

First, it enables longitudinal analysis of algorithm performance, allowing developers to track improvements over time and identify persistent challenges. Second, it supports the identification of common failure modes and the development of mitigation strategies. Third, it creates opportunities for advanced data analytics, including the application of machine learning techniques to optimize system design and operation.

The availability of shared datasets also fosters collaboration and reproducibility. Researchers can build upon previous work, validate their results against established benchmarks, and contribute new scenarios to the community. This collective approach accelerates progress and reduces duplication of effort.

4.5 Strategic Positioning and Long-Term Impact

From a strategic perspective, AXESS aligns with broader trends in the space sector, including the increasing importance of autonomy and its verification, the shift toward smaller and more agile missions, and the growing role of commercial actors. By addressing a critical bottleneck in the development process, it has the potential to significantly enhance the competitiveness of the European space ecosystem.

The market analysis highlights the risk of technological dependence in the absence of sufficient domestic infrastructure for testing and validation. AXESS contributes to mitigating this risk by providing a European-based capability that supports the full lifecycle of autonomous system development. This is particularly relevant in the context of strategic autonomy, where access to critical technologies and infrastructure is a key concern.

In the long term, the adoption of Testing-as-a-Service models for spacecraft systems may lead to a fundamental shift in how missions are developed and validated. As continuous, high-fidelity testing becomes the norm, the distinction between design and operations may become increasingly smaller, with systems being continuously updated and revalidated throughout their lifecycle.

4.6 Limitations and Adoption Challenges

Despite its potential, the adoption of AXESS is not without challenges. The reliance on shared physical infrastructure introduces constraints in terms of resource availability and scheduling, which must be carefully managed to ensure fair and efficient access. Additionally, the integration of CI/CD practices into space system development requires a cultural shift, as well as the development of new standards and best practices.

There are also technical challenges related to interoperability, as different users may employ diverse hardware platforms, software architectures, and communication protocols. Ensuring seamless integration within a common testing environment requires the definition of standardized interfaces and robust abstraction layers.

Finally, issues related to data security and intellectual property must be addressed to build trust among users. The success of AXESS as a community platform will depend on its ability to provide strong guarantees in these areas while maintaining ease of use and flexibility.

5 CONCLUSIONS

The increasing adoption of autonomous spacecraft operations, particularly for CubeSats and deep-space missions, requires a fundamental evolution in V&V methodologies. Traditional late-stage testing approaches are no longer sufficient for systems in which Guidance, Navigation, and Control autonomy must operate reliably under uncertainty, limited communication, and strict onboard constraints.

This work presented the EXTREMA Simulation Hub (ESH) as a distributed hardware-in-the-loop environment designed to move V&V upstream in the development process. By integrating facilities such as RETINA, ETHILE, STASIS, ELAPSE, and SPESI within a synchronized real-time loop, ESH enables realistic testing of autonomous navigation, guidance, and control algorithms under representative mission conditions. Early experimental campaigns demonstrated the value of this approach in exposing cross-subsystem interactions, validating closed-loop behaviors, and improving design robustness before flight qualification.

Building on this foundation, AXESS extends ESH into a Testing-as-a-Service platform capable of providing remote, scalable, and continuous access to high-fidelity validation infrastructure. Through web-based interfaces, standardized APIs, CI/CD pipelines, and automated data collection, AXESS introduces a new paradigm in spacecraft testing, where validation becomes continuous, reproducible, and accessible to a broader community of users.

Beyond technical capabilities, AXESS contributes to the creation of a shared ecosystem based on common benchmarks, reusable scenarios, and comparable performance metrics. This supports technology transfer, reduces barriers to entry for startups and SMEs, and strengthens Europe's strategic autonomy in advanced spacecraft validation.

Future work will focus on expanding the range of accessible facilities, refining standardization and certification pathways, and increasing the level of automation and multi-user scalability. The long-term objective is to establish AXESS as a reference infrastructure for autonomous spacecraft development, enabling safer, faster, and more reliable missions through continuous X-in-the-loop validation.

ACKNOWLEDGMENTS

Project EXTREMA has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No. 864697). Project AXESS has received funding from the ERC under grant agreement No. 101189358.

REFERENCES

[1] I. A. D. Nesnas *et al.*, “Autonomous Exploration of Small Bodies Toward Greater Autonomy for Deep Space Missions,” *Front. Robot. AI*, vol. Volume 8-2021, 2021, doi: 10.3389/frobt.2021.650885.

[2] S. Bhaskaran, “Autonomous Navigation for Deep Space Missions,” in *SpaceOps 2012 Conference*. doi: 10.2514/6.2012-1267135.

[3] C. Steiger *et al.*, “GNC Operations for the BepiColombo Mission to Mercury: First In-flight Experience,” in *8th European Conference for Aeronautics and Space Sciences (EUCASS)*, 2019. doi: 10.13009/EUCASS2019-218.

[4] R. Amini *et al.*, “FRESCO: A Framework for Spacecraft Systems Autonomy,” in *2021 IEEE Aerospace Conference* , Big Sky, Montana, USA, 2021.

[5] F. Ornati, P. Panicucci, E. Andreis, and F. Topputo, “RETINA: a highly-versatile optical facility for camera-in-the-loop testing of spaceborne vision-based sensors,” in *46th AAS Guidance, Navigation and Control Conference*, 2024, pp. 1–19.

[6] P. Panicucci, F. Ornati, and F. Topputo, “Design of a Low-Aberration Variable-Magnification Optical Stimulator for Vision System Hardware-in-the-Loop Testing,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 61, no. 6, pp. 15224–15243, Dec. 2025, doi: 10.1109/TAES.2025.3580393.

[7] G. Di Domenico and F. Topputo, “STASIS: an Attitude Testbed for Hardware-in-the-Loop Simulations of Autonomous Guidance, Navigation, and Control Systems,” in *73rd International Astronautical Congress (IAC 2022)*, 2022, pp. 1–20. [Online]. Available: https://hdl.handle.net/11311/1221313

[8] D. Perico, G. Di Domenico, G. Merisio, and F. Topputo, “ELAPSE: a FlatSat software and processing unit for deep-space autonomous GNC systems testing,” in *46th AAS Guidance, Navigation and Control Conference*, 2024, pp. 1–17.

[9] C. Giordano and F. Topputo, “SPESI: a Real-Time Space Environment Simulator for the EXTREMA Project,” in *33rd AAS/AIAA Space Flight Mechanics Meeting*, 2023, pp. 1–13. [Online]. Available: https://hdl.handle.net/11311/1228274

[10] F. Ornati, G. Di Domenico, P. Panicucci, and F. Topputo, “High-accuracy, vision-based attitude estimation system for air-bearing spacecraft simulators,” *J. Spacecr. Rockets*, vol. 62, no. 2, pp. 579–591, 2025, doi: 10.2514/1.A36073.

[11] G. Di Domenico, “Development of a hardware-in-the-loop simulation framework for interplanetary transfers on smaller timescales,” 2019.

[12] P. W. Kenneally, S. Piggott, and H. Schaub, “Basilisk: A Flexible, Scalable and Modular Astrodynamics Simulation Framework,” *Journal of Aerospace Information Systems*, vol. 17, no. 9, pp. 496–507, 2020.

[13] M. S. Feather *et al.*, “Demonstrations of System-Level Autonomy for Spacecraft,” in *2021 IEEE Aerospace Conference (50100)*, 2021, pp. 1–18. doi: 10.1109/AERO50100.2021.9438427.

[14] S. Teixeira, R. Arrais, and G. Veiga, “Cloud Simulation for Continuous Integration and Deployment in Robotics,” in *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, 2021, pp. 1–8. doi: 10.1109/INDIN45523.2021.9557476.

[15] T. Gutierrez, A. Bergel, C. E. Gonzalez, C. J. Rojas, and M. A. Diaz, “Systematic Fuzz Testing Techniques on a Nanosatellite Flight Software for Agile Mission Development,” *IEEE Access*, vol. 9, pp. 114008–114021, 2021, doi: 10.1109/ACCESS.2021.3104283.

[16] A. Taitler *et al.*, “The 2023 International Planning Competition,” *AI Mag.*, vol. 45, no. 2, pp. 280–296, Jun. 2024, doi: 10.1002/aaai.12169.

[17] J. Lebreton *et al.*, “Training Datasets Generation for Machine Learning: Application to Vision Based Navigation,” 2024. [Online]. Available: https://arxiv.org/abs/2409.11383

[18] F. Ornati, P. Panicucci, A. Pizzetti, F. Capolupo, and F. Topputo, “On the Radiometric Calibration of Optical Hardware-in-the-Loop Stimulators,” *J. Spacecr. Rockets*, pp. 1–14, Apr. 2026, doi: 10.2514/1.A36543.

[19] B. Broekman and N. Edwin, *Testing Embedded Software*. Addison-Wesley, 2003.

[20] G. Lentaris *et al.*, “High-Performance Embedded Computing in Space: Evaluation of Platforms for Vision-Based Navigation,” *Journal of Aerospace Information Systems*, vol. 15, no. 4, pp. 178–192, 2018, doi: 10.2514/1.I010555.

[21] M. Cols Margenet, H. Schaub, and S. Piggott, “Flight Software Development, Migration, and Testing in Desktop and Embedded Environments,” *Journal of Aerospace Information Systems*, vol. 18, no. 4, pp. 157–174, 2021.

[22] J. Schierman *et al.*, “Run-time verification and validation for safety-critical flight control systems,” in *AIAA Guidance, Navigation and Control Conference and Exhibit*, 2008, p. 6338.

[23] C. Hofmann, A. C. Morelli, and F. Topputo, “Performance Assessment of Convex Low-Thrust Trajectory Optimization Methods,” *J. Spacecr. Rockets*, vol. 60, no. 1, pp. 299–314, 2023.