DOI: 10.13009/EUCASS2023-643
Aerospace Europe Conference 2023 – 10TH EUCASS – 9TH CEAS

Green Propulsion Demonstrator "The LÄNDer"

Florian Merz⋆†, Marius Wilhelm⋆, Till Hörger⋆, Lukas Bischof⋆, Fabio Addario⋆
Moritz Schütz⋆, Dr. Christoph Kirchberger⋆, Prof. Dr. Stefan Schlechtriem⋆
⋆German Aerospace Center (DLR), Institute of Space Propulsion,
Langer Grund, 74239 Hardthausen

Florian.Merz@dlr.de† · Marius.Wilhelm@dlr.de · Till.Hoerger@dlr.de · Lukas.Bischof@dlr.de ·
Fabio.Addario@dlr.de · Moritz.Schuetz@dlr.de · Christoph.Kirchberger@dlr.de
†Corresponding author

Abstract

For future manned and unmanned space missions, landing systems are required, which are able to initiate and carry out soft autonomous landings on extraterrestrial celestial bodies. Development of rapid and robust guidance and control (GNC) systems as well as efficient, controllable and safe thrusters meeting a high level of autonomy and safety is crucial for successful lander missions. To meet those requirements, both component level view such as physico-chemical behavior of the propellants and the transients of the propulsion system itself as well as the system level view of the behavior of the holistic system are of utmost interest. To enable extensive testing and demonstration of new GNC methods and rocket engines, DLR Lampoldshausen has extended its research fields to develop a modular lander platform and a test bench with variable degrees of freedom. The aim is to combine extensive testing of GNC algorithms with sustainable propellant technology.

1. Introduction

The successful landing of the first stage of SpaceX's Falcon 9 rocket in December 2015 was a significant milestone for the space industry [34]. It marked the beginning of a new era of more sustainable and cost-effective launch systems via feasibility demonstration of reusable launch vehicles and emphasized the innovative potential of private companies in a mostly government-financed branch. One of the key components for the success of SpaceX's landing was the Guidance, Navigation & Control (GNC) system [11]. The deployed GNC enabled pinpoint landings of rockets, which can be reduced to the challenging and well-known unstable inverted pendulum problem. Until recently, GNC algorithms onboard designated lander space vehicles supported the successful guidance from Earth into orbit and descent to some extraterrestrial body with large uncertainty ellipses regarding the touchdown area, often spanning multiple kilometers [19]. Landing spacecraft precisely in highly uncertain environments such as sparsely studied celestial bodies or under the challenging conditions of Earth's turbulent atmosphere demands for more elaborate control algorithms. Though admittedly some sensors such as Global Navigation Satellite System (GNSS) receivers have likely been deployed for more precise state estimations that are not available for comparable tasks in space [11], the specific example of SpaceX underlines the creation of new mission types by mastering GNC technology. New GNC methods have thus to be developed to enable pinpoint landing of spacecraft landers with high degrees of autonomy and to meet the increasingly stringent demands on landing accuracy requirements for enhanced scientific return of missions [25].

The application of GNC with the aim of landing on extraterrestrial celestial bodies dates back to the 1960s. Back then, progress was mainly driven by the Cold War fueled space race between the Union of Soviet Socialist Republics and the United States. The first attempts, like Luna-2, led to intentional hard landings on the surface of the moon, i.e., impacting the surface without controlled descent. In 1966, the Soviet Luna-9 mission successfully performed the first soft landing on an extraterrestrial body, which marked the beginning of unmanned and manned soft landings by realising the former [28]. The latter was achieved only three years later, in 1969, with the first crewed mission to the Moon, Apollo 11. The effort to plan and execute such lander missions is based on the yielded scientific findings. Therefore, numerous successful unmanned landings on a variety of different celestial bodies like planets, their moons, or even asteroids took place following mankind setting foot on the Moon for the first time. With this specific achievement remaining unmatched for more than 50 years, a lot of work has gone and goes into re-enabling crewed missions to the Moon and even beyond. Consequently, there has been a notable increase in interest surrounding both manned and unmanned spacecraft landers.

DOI: 10.13009/EUCASS2023-643

Nowadays, space missions remain no longer limited to government-financed national space agencies. The iSpace lander mission “Hakuto-R M1” launched in December 2022 was expected be the first privately-led lunar lander mission [30], but failed due to an unforeseen error in the state estimation module [3]. A new space race arose with multiple commercial enterprises like OHB (LSAS), SpaceIL (Beresheet 2), SpaceX (Starship) and Blue Origin (ILV) developing lander technology and competing against each other for unmanned and manned missions, such as the National Aeronautics and Space Administration (NASA)’s Artemis program of sending back humans to the Moon. Present-day missions benefit from extensive research and numerous findings from the past. Despite technological advances, landing missions remain particularly challenging, as underlined by the recent crash landing of Beresheet and iSpace, emphasizing that there is no room for error [3, 14]. A successful landing requires flawless performance of GNC algorithms with sophisticated and redundant sensor and actuator concepts. As missions target increasingly distant celestial bodies, the importance of autonomous systems emerges due to transmission delays preventing the timely transmission of commands from Earth. Furthermore, modern missions face unprecedented economic, environmental, and safety constraints, demanding the development and deployment of new methods to address these challenges [12, 18]. Consequently, the Institute of Space Propulsion at the German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt e.V.) (DLR) facility in Lampoldshausen conducts research on sustainable alternatives to conventional rocket propellants and engines and thrust control via throttling [36]. In addition, Reinforcement Learning (RL)-based control of rocket engines powered by these green propellants is researched on site in the scope of applicability to lunar landing applications [35].

The necessity for extensive testing and demonstration of new methods on Earth-based and free-flying test benches is stated in [25], with Europe lacking experience on the field of GNC and dedicated test benches [16]. DLR Lampoldshausen thus decided to extend its research fields further from conventional rocket test benches towards lander test benches enabling x Degree of Freedom (DoF) testing. By developing both a modular lander platform and a test bench with variable DoF, extensive testing of GNC algorithms and rocket engines can be combined. With the aim for proving the applicability of green propellant technology studied by DLR, a lander demonstrator based on green propellant technology is developed in the context of a lighthouse project named “The LÄNDer”. The first steps have already been accomplished by implementation of an altitude controller [22] and an attitude controller [10]. The work covered different approaches, including conventional Proportional, Integral & Derivative (PID) control, optimal control and Model Predictive Control (MPC) of the altitude of the lander demonstrator. It was focused on cold gas propulsion using nitrogen and enabled 1-DoF control of the lander by design of a dedicated test bench. For the attitude controller, a 4-DoF test bench was developed. Promising techniques such as data-driven and Machine Learning (ML)-based methods have been investigated by [10].

2. Background

Guidance, Navigation & Control (GNC) is a crucial task for every mission to space. It is at the core of rocket launches, where predefined trajectories have to be tracked to ensure safe and efficient ascent into orbit and reaching of the desired orbital pose within the required precision boundaries. Every satellite in orbit is equipped with a GNC system to maintain or correct its orbit. Furthermore, GNC algorithms are also a central key for success of (autonomous) spacecraft landings. GNC consists of solving the following three main tasks:

- **Guidance** is the process of determining a trajectory from the current state to the desired state and deriving acting commands for the control loop to follow such path. Guidance in the context of landers involves directing the spacecraft from its current position to a target position. An example is the computation of a trajectory from the initial state when leaving orbit, towards a location deemed suitable and safe for landing. Due to limited surface topology data, landing sites can only be roughly estimated beforehand. Nowadays, potential landing sites are often identified based on data collected while orbiting, hence trajectories for the guidance can only be calculated as the lander orbits around the celestial body it is supposed to land on.

- **Navigation** is the generic term for all tasks related to determination and estimation of the current state. Navigation filters are usually deployed to merge information from multi-sensor systems with the aim of minimizing uncertainties and hence providing the best possible state estimation together with assigned current uncertainties. The navigation task in the context of landers consists of fusing the information from different sensors to estimate the current pose. Applied to spacecraft landers, this involves estimation of the initial state when leaving orbit for the Powered Descent Initiation (PDI) and providing continuous updates until touchdown on the surface.

- **Control** is the low-level method generating actuator outputs for tracking of targets dictated by the high-level guidance. Control for planetary landers consists of activation of actuators for tracking of the trajectory (pose & time) generated by the guidance. This also includes generation of steering commands to prevent unwanted

DOI: 10.13009/EUCASS2023-643

oscillatory movements, i.e., stabilizing of the system. Closed loop feedback systems rely on state estimations provided by the navigation module for computation of actuator outputs. Control algorithms are thus the link between the guidance and navigation modules.

GNC is commonly implemented in a cascade as depicted in Figure 1. The control module receives the desired state as an input from the guidance module, while the estimated state is outputted from the navigation module. The control logic commonly consists of cascaded controllers, i.e., an inner loop for stabilization of states and an outer loop for tracking of the trajectory [16, 33]. With regard to spacecrafts, GNC is subject to some differences in direct comparison to Earth-based systems. Like any other constraint-free body, spacecraft dynamics are six DoF motion problems, i.e., three translatory and three rotatory DoFs have to be controlled. Translatory motion in vertical direction can be manipulated through variation of the main thrust. The other two translatory motions in horizontal direction are usually enforced either by tilting of the thrust vector in the direction of commanded translatory position change. The vertical component of the main thruster vector decreases when tilted, hence the outer loop usually compensates this effect by increasing the main thrust. The Environment for Autonomous GNC Landing Experiments (EAGLE) lander demonstrator developed by DLR made use of a sliding-mode control [16]. Chang’E3, a Chinese lunar lander project, deployed conventional PID controllers [38]. Another privately run project striving to send a lander to Moon, Autonomous Landing and Navigation Module (ALINA), deploys Linear–Quadratic Regulator (LQR) type controllers [37]. In space applications, the analytical derivation of optimal open loop control for altitude is referred to as the “suicide burn” technique. As the name suggests, margin for error is very low: starting to decelerate too early would result in reaching zero velocity above ground, with the rest of the descent being a free-fall. Initiating the burn too late ends with an impact with non-zero velocity on ground.

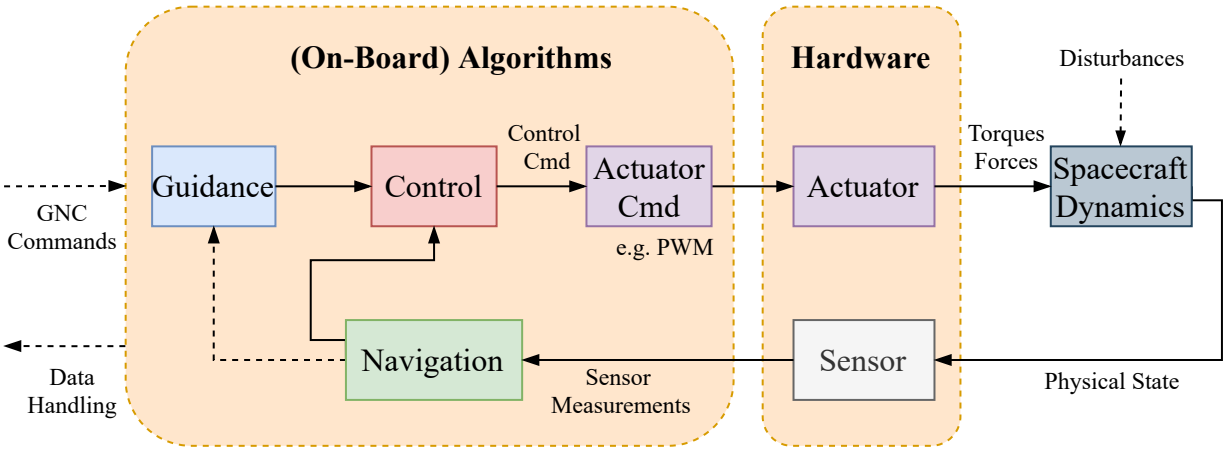

Figure 1: Schematic overview of GNC, inspired by [7]

Although the fundamental functions of GNC have not undergone any significant alterations over the past few decades, more sophisticated models have been devised and accuracy requirements are increasingly stringent [17]. According to a roadmap paper published by the Jet Propulsion Laboratory (JPL), additional sensors, improved filter capabilities and overall robustness would enable a wider range of mission scenarios such as pinpoint landings [25]. Six DoF GNC has been identified as a key research topic for enabling further scientific insights from such missions. Focus on new fuel-optimal and real-time capable techniques has been advised. Furthermore, it has been stated that very-high-frequency GNC systems capable of withstanding highly dynamic environments with high gravitational forces and atmospheric disturbances are required to meet the intensifying accuracy demands of Mars landings. Future lander missions demand for increasing accuracy and autonomy of GNC systems [17]. A typical landing trajectory can be seen in Figure 2. Benchmark methods reveal large potential for further increases on achievable accuracies and hence scientific insights gained from lander missions. Data-driven approaches play a significantly increasing role in conventional control engineering. Approximation of computationally expensive optimization-based controllers such as MPC using Neural Networks (NNs) is already common and recently developed methods like RL may further enhance robustness and enable computationally less demanding control.

Tipaldi et al. state, that “traditional optimization-based control approaches are dependent on the accuracy of the underlying dynamic models, and can fail if such models do not include correct gravitational models, disturbances, and other nonlinear terms” [31]. The authors underline the consequent “need of using robust and adaptive approaches, which can cope with the typical issues of spaces missions, i.e., partially known environment and the limited on-board

DOI: 10.13009/EUCASS2023-643

memory and computational resources” and identify RL-based approaches as promising candidates for space GNC applications due to their generalization capabilities. This statement could also be verified in literature: if the model perfectly represents the reality then optimization-based control approaches such as MPC slightly outperform RL in terms of cost, especially when evaluating on data outside the training data range. With increasing modeling errors, RL was found to perform better in direct comparison to MPC [21]. Due to uncertainties such as atmospheric disturbances, spacecraft aerodynamic coefficients, uncertain initial state estimate or deviations from the expected gravitation, spacecraft landers are usually exposed to such modeling error scenarios [25]. This year’s European Space Agency (ESA) International Conference on GNC dedicates a topic to Artificial Intelligence (AI) in GNC, which further emphasizes expected potentials of ML. The conference call for papers lists potential applications of ML in GNC: modeling and system identification, filtering, control, failure detection and increased autonomy [2]. ML methods are thus prominent candidates to meet the aforementioned requirements of increased filter capabilities and overall robustness [25]. In fact, the recent Mars 2020 lander mission already made use of AI to improve estimation accuracies by deploying a Terrain Relative Navigation (TRN) system and has achieved an unprecedented accuracy [4].

With RL being a recent topic of research, previous work has mainly focused on simulation environments. A common method for validation of such methods involves performing Hardware-in-the-Loop (HIL) tests or using High-Fidelity Models (HFMs) [21]. But to the authors knowledge there exists to this day no RL-based Attitude Control System (ACS) on a deployed spacecraft. A six DoF RL-based control approach has been presented in [17]. The authors used Domain Randomization (DR) for initial lander mass, acceleration and inertial tensor to increase the robustness of the Proximal Policy Optimization (PPO)-based algorithm. They covered the powered descent to the surface for a three DoF and a six DoF problem, thus replacing both guidance and control modules by a RL agent. Pinpoint landings with significantly higher accuracy than current benchmarks could be achieved in simulations with the trained controller being robust to noise and parameter uncertainty. RL-based ACS has already shown promising performance [33]: through DR, an attitude controller outperforming a Quaternion Rate Feedback (QRF) controller on a large range of mass could be trained. The RL agent was trained for any potential spacecraft mass, ranging from the mass/inertia of a small CubeSat up to the International Space Station (ISS). The RL controller demonstrated its capability to generalize by outperforming the QRF while being trained on a large range of possible inertia and the QRF being tuned to the specific inertia both controllers have been evaluated on. Unfortunately, very little information on the NN architecture is given.

As powered descent to celestial bodies surfaces requires burning fuel, decreasing masses and inertia are highly relevant for spacecraft landers. The proved adaptability of RL to a wide range of mechanical properties makes it a tempting choice for such applications. It is also worth mentioning that other methods from the ML domain such as as evolutionary algorithms [13] or the aforementioned approximation of costly online optimization through NN-based function approximators are promising approaches and an ongoing topic of research. As this work focuses on RL, these methods are not further covered.

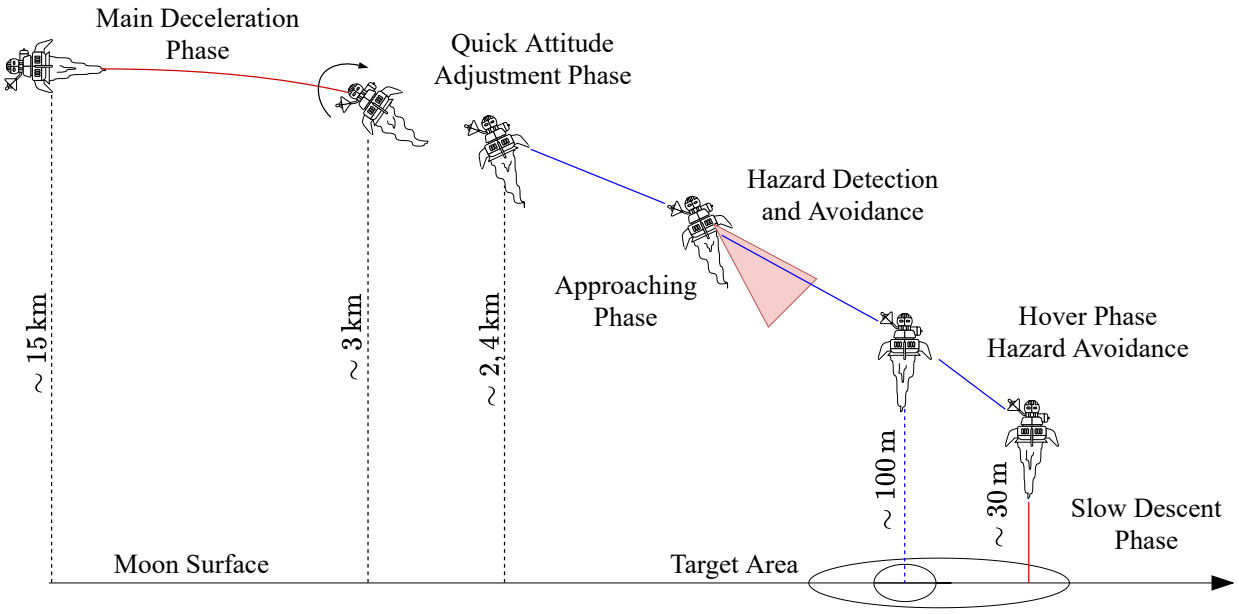

Figure 2: Typical trajectory for landing missions, inspired by [38]

DOI: 10.13009/EUCASS2023-643

3. "The Länder" test bed configurations and model dynamics

Within the scope of the work at the DLR, two experimental set-ups were developed to investigate different control strategies. This chapter describes the selection of the components and the associated measurement technology and programming for the test bed. The experimental setup should initially provide only one controllable degree of freedom, in this case the translational motion perpendicular to the earth. In the second iteration, a test setup was developed to provide four controllable DoF. Both setups are currently using gaseous nitrous as medium, but will be further developed to use monopropellant or bipropellant thrusters.

3.1 1-DoF test setup

The 1-DoF test setup provides a translational movement along a linear rail with a total length of 6 m. A flexible nitrogen supply hose was chosen and mounted to the side to not interfere with the travel path. To prevent the lander from touching down without braking, a spring damper is fitted at the foot end of the linear rail. The schematic representation of the structure can bee seen on the left and the lander itself can be seen on the right in Figure 3.

The landing demonstrator consists of a 3D-printed (PETG) base-plate with two valves, a pressure sensor as well as a laser distance sensor and an accelerometer. The valves can handle a pressure up to 60 bar, allowing a thrust of $F_T = 30\,\mathrm{N}$ per valve. With a weight of $m_L = 2.9\,\mathrm{kg}$ it results in a thrust-to-weight ratio of roughly $\frac{F_T}{m_L \cdot g} = 2$.

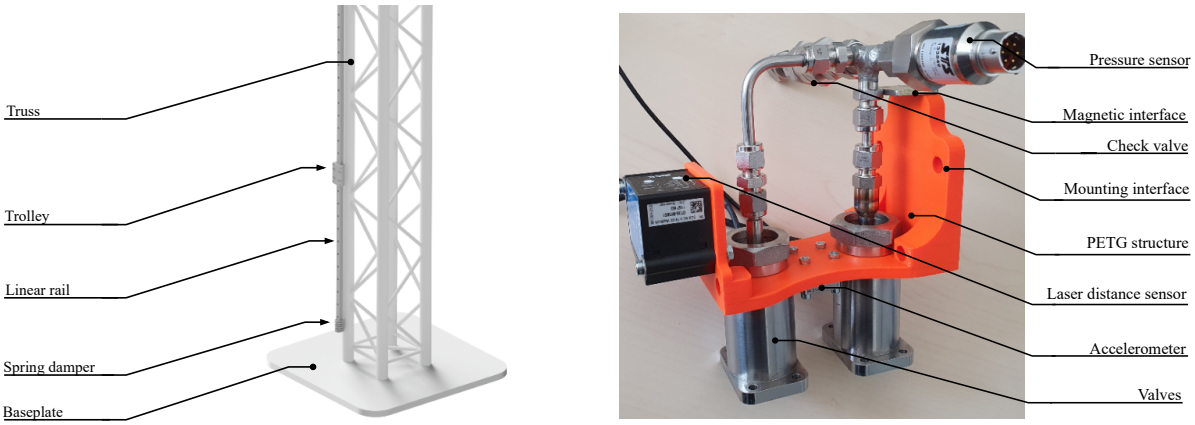

Figure 3: "The Länder" 1-DoF test setup at DLR Lampoldshausen

An electromagnet is required in addition to the mechanical components of the lander, which has a magnetizing effect in its de-energized state. The electromagnet can then be used to launch the lander from a defined height. The measurement data acquisition and control of the test bench is performed by means of an external ADwin-Pro-II- real-time measurement system with a 300 MHz T11 processor from Jäger Messtechnik. The controller can be directly connected to the real-time measurement system via Python or Matlab. This has the advantage that new controller strategies can be implemented quickly. Especially in the field of AI control, Python offers large libraries. Since the Python programming language is not considered to be real-time capable, tests were carried out to determine the time delay and its effect on robustness.

3.2 4-DoF test setup

The 4-DoF demonstrator consists of a Carbon Fiber Reinforced Plastic (CFRP) sandwich base plate fitted with tubes, valves, sensors and electronics. The base plate is a dodecagon with a diameter of 900 mm and thickness of 13.5 mm with 10 mm foam core. Its acs) utilizes nitrogen cold gas fed through flexible hoses from an external feed. Currently, attitude is calculated from measurements of an accelerometer mounted as close at the Center of Mass (CoM) as possible. By switching solenoid valves gas is led trough a 3D-printed nozzle to generate torque. An assembly of the test bench with the lander attached is shown in Figure 4.

DOI: 10.13009/EUCASS2023-643

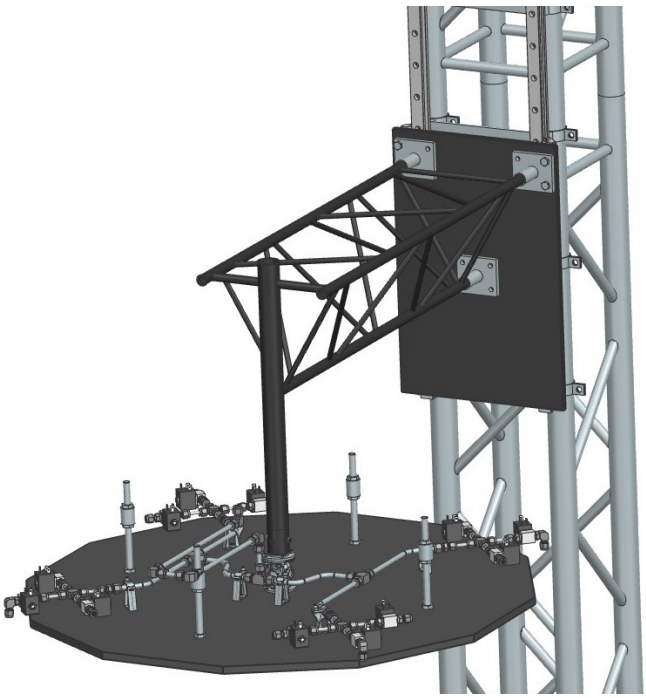

Figure 4: "The Länder" 4-DoF test setup at DLR Lampoldshausen

To allow rotational movement along all three rotational axes the demonstrator is fastened on a ball bearing (to allow yaw movement) and a universal joint (to allow pitch and roll movement). All three of these axes meet in a single point at the center of the universal joint’s cross. Thus, the center of mass (COM) can be adjusted using balance weights to be inside this pivot point, cancelling out all weight induced torque. This principle was chosen over a pin and cone bearing as described in [24], so that forces in all directions can be transferred, enabling controlled movement along translational axes.

Translational movement along a single axis enabled by a linear guide rail and a trolley. The universal joint and therefore the demonstrator is held by a lightweight truss consisting of carbon fiber reinforced plastic (CFRP) tubes. By using lightweight trolley structures, only minimal interference in the spacecraft’s dynamics is achieved, without using complex pulley systems as described in [20].

3.3 Embedded Flight Controller

For the 4-DoF setup the ADwin-Pro-II computer was replaced with a modular on-board avionics suite. Modularity is achieved by splitting the tasks between separate Microcontroller Units (MCUs), which communicate over a shared Controller Area Network (CAN)-bus. Concretely the avionics suite is split into 3 seperate STM32L4 MCUs that have the following tasks:

1. **Sensor-Interface and State-Estimation**: The first MCU interfaces with sensors and runs a state-estimation algorithm. In our example a complementary filter was used to filter incoming accelerometer and gyroscope data. The estimated state and sensor data (if needed) are then written onto the CAN-bus.

2. **Attitude-Controller and Actuator-Interface**: The second MCU takes the estimated states from the CAN-bus and feeds them into the control algorithm described in section 4. The MCU then executes the output of the control algorithm by controlling the solenoid-valves leading to the cold-gas-thrusters. The output is also written to the CAN-bus.

3. **Data-Logging and Human-Interface**: The third MCU was used to log the contents of the CAN-bus onto an SD-Card. It is also responsible to relay user commands, which it receives over a serial connection to the other MCUs. This is used to set the attitude set-point, which the second MCU is trying to reach.

This architecture allows for easier modifications to subsystems, and easier HIL-testing. In our example the second MCU may simply be swapped for a more powerful device once the RL-algorithms become to computationally expensive for the current MCU.

DOI: 10.13009/EUCASS2023-643

3.4 Modeling of attitude dynamics

The lander is a non-linear and depending upon the position of the CoM also unstable system, that consequentially has to be stabilized by means of a controller. The total angular momentum of a rigid body can be expressed as

$$\mathbf{h} = \int_{\mathcal{B}} (\mathbf{r} \times \mathbf{v}) dm, \tag{1}$$

where $\mathbf{r}$ is the distance from the CoM to the mass element dm and $\mathbf{v}$ represents its velocity [9], refer to Figure 5. If the

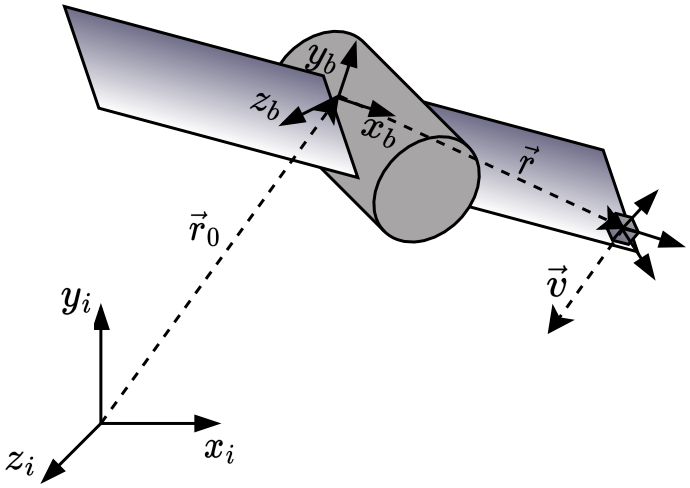

Figure 5: Motion of rigid body motion in different frames

body rotates around its CoM, the velocity of the mass point dm can be expressed as the cross product of angular rate and distance from the CoM as

$$\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}. \tag{2}$$

Injecting this formulation into Equation 1 yields

$$\mathbf{h} = \int_{\mathcal{B}} (\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}))\, dm. \tag{3}$$

By multiplying out Equation 3,

$$\begin{aligned} \mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}) = &[(y^2 + z^2)\,\omega_1 - (xy) \quad \omega_2 - (xz) \quad \omega_3]\,\mathbf{e_1} \\ &+ [-(xy) \quad \omega_1 + (x^2 + z^2)\,\omega_2 - (yz) \quad \omega_3]\,\mathbf{e_2} \\ &+ [-(xz) \quad \omega_1 - (yz) \quad \omega_2 + (x^2 + y^2)\,\omega_3]\,\mathbf{e_3} \end{aligned}$$

one can write the expression of angular momentum in the body frame $\mathbf{h}_B$ as

$$\mathbf{h}_B = \mathbf{I}\boldsymbol{\omega}, \tag{4}$$

with the inertia matrix $\mathbf{I}$. Applying the second fundamental law of rigid body dynamics

$$\frac{d\mathbf{h}}{dt} = \sum_i \mathbf{M_i} \tag{5}$$

to Equation 1 resolves in the formulation of angular moment in body frame $\mathbf{h}_B$ as a function of external torques in body frame $\mathbf{M}_{B,i}$, i.e.,

$$\dot{\mathbf{h}}_B + \omega_\mathbf{b} \times \mathbf{h}_B = \sum_i \mathbf{M}_{B,i}. \tag{6}$$

As the lander is propelled with gas from external sources, mass and inertia of the lander are constant. Equation 6 can thus further be simplified using Equation 4 to

$$\mathbf{I}\dot{\boldsymbol{\omega}}_B + \boldsymbol{\omega}_B \times (\mathbf{I}\boldsymbol{\omega}_B) = \sum_i \mathbf{M}_{B,i}. \tag{7}$$

DOI: 10.13009/EUCASS2023-643

4. Control Strategies

In the following, three control strategies that have already been applied on the test bed are briefly described in more detail.

4.1 Conventional control using PID-type controller

The probably best known controller type is the PID controller. The aim of each feedback controller is to achieve some desired set point via manipulation of variables. The output y of the continuous PID controller and therefore the manipulated variable consists of a weighted sum of proportional (P), integral (I) and derivative (D) responses to the input u:

$$y = K_P u + K_I \int_0^t u d\tau + K_D \dot{u}. \tag{8}$$

PID control is widely used because of the advantageous combination of its components capabilities:

- The proportional part computes a control variable proportional to the control deviation, which is beneficial for a fast response and elimination of deviations. However, elimination of steady-state errors solely based on P-controllers is only ensured if the plant consists of at least one integrator, which is not necessarily the case.
- Therefore, an integral controller is deployed in combination with the P-controller. By operation of such a controller, the elimination of stationary control deviations can be guaranteed, however if the controlled system already has integrating behaviour, the addition of an integral controller may lead to instabilities. The downside of I-control, namely slow settling of control outputs, is counteracted by the combination with the P-controller.
- The output of the derivative part is proportional to the rate of change of the control input. This property can be favorable when demanding fast responses to varying control inputs, but large factors K_D can also lead to large oscillation amplitudes or even instability. Especially stochastic deviations resulting from measurement noise are of a major concern and derivative control is thus recommended to be deployed solely with filtered measurements.

Usually, outputs of the PID controller have to be limited as demanded actuator values are constrained or bound to operating conditions [6]. This can be illustrated on the basis of a rocket engine with a minimal and maximum achievable thrust: a controller should consequently only demand thrusts that can be provided by the actuators. One major concern when using integral controllers is the so-called integral windup. It can typically be observed when the controller output or plant saturates, making the controller unable to decrease the error faster and therefore the integral part growing. Integral windup leads to performance degradation in form of large overshoot and settling time, oscillations and instability. These can be counteracted by either a back-calculation approach or by stopping integration when reaching saturation, which is also known as clamping [1, 8].

A well-tuned PID controller often satisfies performance requirements at minimal computation costs. Furthermore, well-tuned PID controllers show nearly optimal performance in undisturbed environments [26]. PID controllers with anti-windup satisfy the Bounded-Input Bounded-Output (BIBO) stability criterion. In addition, PID control is a well studied strategy with successful demonstrations in numerous applications.

4.2 Optimal Control

In optimal control, a cost functional is maximized or minimized in the search for the optimal control as a function of time. The calculation requires the dynamic model, boundary conditions (initial and final conditions), and path constraints as constraints.

Trajectory calculations in the past and even today have mostly been made under the assumption of a point mass model [17]. This is partly because the complexity of optimization should be kept within limits, as the numerical effort increases rapidly even with the simplified assumption of a point mass model.

A simple example that can be solved analytically is the *suicide burn*. This approach relies on the special case of linear control. The name already indicates the drawback of this control strategy. If the correct timing of the engine ignition during the descent of a lander or rocket stage is missed, the thrust phase will either end several meters above the ground, or the descent rate will not be completely reduced, resulting in a crash. The advantage of a suicide burn is its time-optimality and, with maximum I_{sp} (specific impulse), also fuel-optimality. The following sections present the solution approach for a time-optimal landing of a cold-gas lander.

DOI: 10.13009/EUCASS2023-643

First, the cost functional is defined as:

$$J[u] = t_f \tag{9}$$

which minimizes the time required for the point mass to travel from position 0 to position 1. Consequently, the dynamic system is described by the motion equations:

$$\begin{aligned} \dot{s} &= v, \\ \dot{v} &= u \end{aligned} \tag{10}$$

where the control variable u represents acceleration. Additionally, initial and final conditions are taken into account:

$$\begin{aligned} s(0) = h_0, v(0) = 0 \\ s(t_f) = 0, v(t_f) = 0 \end{aligned} \tag{11}$$

The starting point is the altitude h_0 and the end point is the ground $h = 0\,\mathrm{m}$, where the trajectory starts from rest and the velocity should be mitigated at landing. The control variable u is constant for two subintervals each. This type of control is also called „bang-bang“ control. In the first partial interval, until the onset of thrust, only the acceleration due to gravity g acts. In the second part of the flight phase, the thrust F_T and therefore the acceleration $a = \frac{F_T}{m_L} - g$ is acting on the lander.

Finally, the distance and the velocity of the partial intervals can be calculated for the equations of motion. For the first partial interval $0 \leq t \leq t_s$ follows:

$$\begin{aligned} \dot{v} = u = -g, v(0) = 0 \Rightarrow v(t) = -g \cdot t, \\ \dot{s} = v, s(0) = h_0 \Rightarrow s(t) = h_0 - \frac{1}{2} \cdot g \cdot t^2 \end{aligned} \tag{12}$$

For the second subinterval, the firing time τ is defined as $\tau = t - t_s$. Analogous to the first subinterval follows:

$$\begin{aligned} v\Big|_{\tau=0} &= -g \cdot t_s, \dot{v} = u = a, \\ &\Rightarrow v(t) = -g \cdot t_s + a \cdot \tau, \\ s\Big|_{\tau=0} &= h_0 - \frac{1}{2} \cdot g \cdot t_s^2, \dot{s} = v, \\ &\Rightarrow s(t) = h_0 - \frac{1}{2} \cdot g \cdot t_s^2 + \frac{1}{2} \cdot a \cdot \tau^2 \end{aligned} \tag{13}$$

By inserting the end conditions, a system of equations with two equations and two unknowns is created:

$$\begin{aligned} v(t_f) = -g \cdot t_s + a \cdot (t_f - t_s) = 0, \\ s(t_f) = h_0 - \frac{1}{2} \cdot g \cdot t_s^2 + \frac{1}{2} \cdot a \cdot (t_f - t_s)^2 = 0 \end{aligned} \tag{14}$$

Resolved to t_s and t_f follows:

$$\begin{aligned} t_s &= \sqrt{\frac{2 \cdot a \cdot h_0}{(a+g) \cdot g}} \quad \text{with} \quad a > -g, \\ t_f &= \sqrt{\frac{2 \cdot (a+g) \cdot h_0}{a \cdot g}} \quad \text{with} \quad a > 0 \end{aligned} \tag{15}$$

where the constraints mean that the suicide burn only applies to landers with a thrust-to-weight ratio $\frac{F_T}{m_L \cdot g} > 1$. The results for the simulation and the real experiment are compared in subsection 5.2.

4.3 Modern Control - Reinforcement learning

Within the last decade, Reinforcement Learning (RL) successfully proved its ability to solve complex tasks such as playing Atari games [23] and outperforming human experts [27]. With maturing algorithms and a highly active research community, RL applied to classic control problems gained in popularity and revealed promising results, sometimes identifying surprising and innovative solutions with unforeseen reward maxima that have not been considered by experts [5]. RL is part of the Artificial Intelligence (AI) domain, which can be mainly divided into the three following categories:

DOI: 10.13009/EUCASS2023-643

- **Supervised learning**, which usually requires human annotation of data. Input data, often referred to as features, is fitted via function approximators such as Gaussian Processes (GPs) or NNs to outputs, also known as targets or labels. The aim of supervised learning is to learn the best possible approximation of the underlying function, which may either be used for classification or regression tasks.

- **Unsupervised learning** can be used when trying to extract information from unlabeled data. Through algorithms such as K-Means, information clusters can be identified and separated from each other. This corresponds to the definition of data science as a way of extracting “knowledge and insights from noisy, structured and unstructured data” [15].

- **Reinforcement Learning** does not need annotated data as the so-called agent learns solely through interaction with its environment. By applying similar methods as in supervised learning (NNs, GPs), good actions taken by the agent are rewarded and thus reinforced, hence the name of this ML domain.

At the core of RL are states, actions and rewards. States $\mathbf{S}_t$ contain information regarding the current observation and are thus comparable to the states $\mathbf{x}(t)$ in conventional control engineering. They are usually manipulated in a way to achieve some target state. This can be done by actions $\mathbf{A}_t$, where set values $\mathbf{u}(t)$ are the equivalent in control engineering. Rewards R_t are scalars evaluating the value of the previous state and action. In case of classic control engineering, the environment is known as the model plant and the transition dynamics are usually derived from models of the system. If transition dynamics are known, RL methods such as Dynamic Programming (DP) are guaranteed to converge towards optimal control [29]. If transition dynamics are unknown or intractable, RL is a powerful method for derivation of close-to-optimal control algorithms without derivation of a model and solely based on interaction with the environment [32] even states “RL is a form of optimal control without a model”. A comparison of terminology and symbols for RL and control engineering is given in Table 1.

Table 1: Comparison of RL and conventional control terminology

	Reinforcement learning	**Control engineering**
actor	agent	controller
states	$\mathbf{S}_t$	$\mathbf{x}(t)$
manipulations	actions $\mathbf{A}_t$	inputs $\mathbf{u}(t)$
plant model	transition probabilities	state-space representation
	$p(\mathbf{S}_{t+1}, R_t \vert \mathbf{S_t}, \mathbf{A}_t)$	$\dot{\mathbf{x}}(\mathbf{t}) = f(t, \mathbf{x}(t), \mathbf{u}(t))$
optimization	maximizing the expected return G_t	minimizing some cost function J

RL founds on the idea of the Markov Decision Process (MDP) as depicted in Figure 6. The algorithm, also referred to as agent, interacts with its environment through selection of actions $\mathbf{A}_t$. Based on the current observation of the state $\mathbf{S}_t$, the agent will try to pick actions $\mathbf{A}_t$ maximizing the expected reward R_{t+1} for transitioning to the next state $\mathbf{S}_{t+1}$. When applying $\mathbf{A}_t$ to the environment, the current state $\mathbf{S}_t$ is updated by the transition dynamics inherent to the environment and the reward received for transitioning into a new state is given by R_{t+1}. The agent will thus learn to interact with the environment and try to maximize the cumulative reward, also known as return G_t [29].

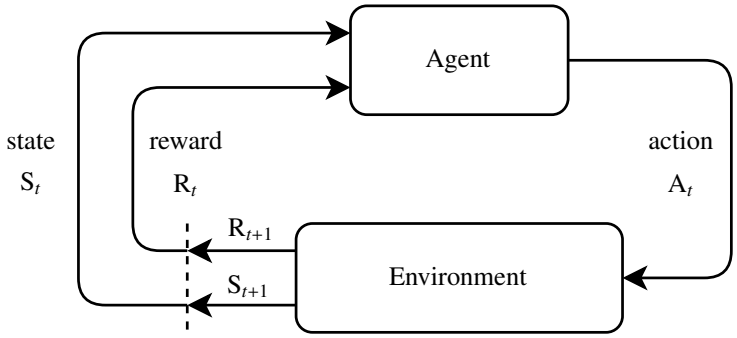

Figure 6: Markov decision process [29]

DOI: 10.13009/EUCASS2023-643

5. Results and Discussion

5.1 PID control results

The PID controller was investigated for a setpoint specification that included hovering, rapid rising, fast falling and following a given trajectory in the form of a sine function. The result can be seen in Figure 7. The setting variables are $K_P = 60, K_I = 65.45, K_D = 13.75$. An input pressure of $p_0 = 35\,\text{bar}$ was chosen, resulting in a theoretical thrust of $F_T = 40\,\text{N}$. For the simulation, the same parameters were chosen, with a mass of $m_L = 2.9\,\text{kg}$. The orange line shows the commanded trajectory, the red line describes the simulated PID controller and the blue line is the measured height during a real test.

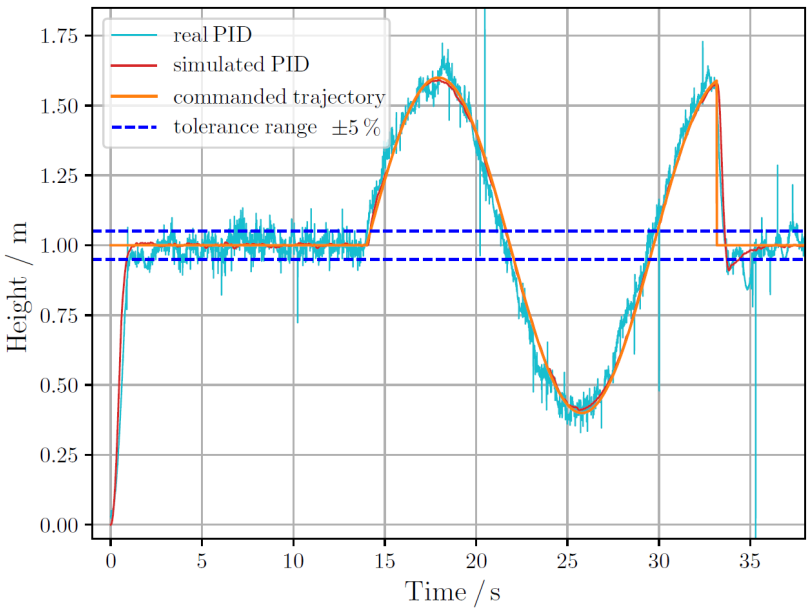

Figure 7: Comparison between simulated and real PID control

It could be demonstrated that a PID controller is a suitable control strategy because it exhibits approximately perfect tracking behavior. This is mainly due to the integrating part, which can compensate well for the uncertainties in the model and the mass. Nevertheless a PID controller has limited capabilities because new gain variables are needed for each task (hovering, landing, climbing). At this point it should be mentioned that so far no correction elements and no filters have been used, and only the distance sensor was used for height measurement.

5.2 Optimal Control Results

The time- and fuel-optimal fall (suicide burn) was executed from a height of 2.07 m. In Figure 8 the analytical solution described in subsection 4.2 is shown as dotted line. For the simulation, the total mass of the lander $m_L = 2.7\,\text{kg}$ and a thrust of $F_T = 35\,\text{N}$ were chosen. This results in $a = \frac{F_T}{m_L} - g = 3.16\,\text{m}\,\text{s}^{-2}$. After $t_s = 0.32\,\text{s}$ of the free fall, the thrust begins and ends at $t_f = 1.32\,\text{s}$ with the touchdown on the ground, after $t_{burn} = 0.996\,\text{s}$ burning time. The time-optimal case is contrasted with the free fall (dash-dotted), which already impacts the ground unbraked after 0.65 s. The real test is shown as a solid line. To achieve a thrust of $F_T = 35\,\text{N}$, the pressure must be preset in the real test. This was taken from previous tests and is approximately $p_0 = 32\,\text{bar}$. The influence of the supply hose on the total weight cannot be determined concretely,thrust of $F_T = 35\,\text{N}$, the pressure must be preset in the real test. This was taken from previous tests and is approximately $p_0 = 32\,\text{bar}$. The influence of the supply hose on the total weight cannot be determined concretely, but increases the higher the lander takes off. The evaluation showed that the actual pressure was $p_0 = 35\,\text{bar}$. The oscillation at the end is caused by the spring damping, which prevents damage to the lander due to a too fast touchdown. The evaluation showed that the actual pressure was $p_0 = 35\,\text{bar}$. The oscillation at the end is caused by the spring damping, which prevents damage to the lander due to a too fast touchdown.

DOI: 10.13009/EUCASS2023-643

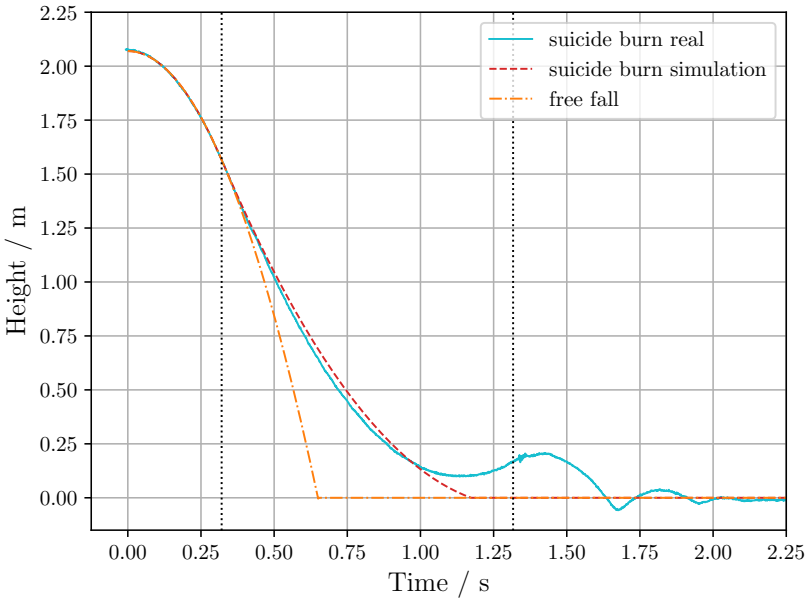

Figure 8: Comparison between simulated and real suicide burns

The suicide burn is time-optimal, but as the small-scale experiment already shows the system must be known exactly. Small deviations from the analytical model can have devastating effects. In this experiment, the pressure was set slightly too high, resulting in a suicide burn that starts a few milliseconds too early and causes a small overshoot. For planning an interplanetary trajectory through the simplified model of the mass point, where milliseconds are not necessarily critical, optimal control is used as part of the GNC in the form of guidance. A further development in this direction would be conceivable by means of a MPC controller.

5.3 RL Control Results

After training the RL agent in simulation, the controller is extracted from the simulation environment and deployed on the MCU for evaluation on the real system. It is worth noting that this type of evaluation corresponds to a “zero-shot transfer” as the controller is transferred to the test bench without further tuning or DR. For the evaluation of the controller, the lander demonstrator is deflected at a maximum possible angle on the test bench and has to stabilize the system to an Euler angle of 0 °. The RL algorithm is still under development. However, initial results have already been shown for 1-DoF and 2-DoF. Several attempts were made, both with domain randomization and without.

An evaluation of this controller is presented in Figure 9. Metrics of this evaluation are listed in Table 2. For this type of evaluation, it was found that the simulated results and the evaluation of the controller on the test bench results in similar performance. DR thus contributed to improve the performance of the controller on the real system. This is because deviations from the trained model are likely and result in severe performance degradation without DR. If the real properties of the system were perfectly known, a zero-shot transfer would probably work well. However, the DR captures the uncertainty with respect to the properties of the system and thus improves the performance on the real system while reducing the performance in simulation as it has to deviate from the optimal policy for the assumed mean model. The evaluation of the controller on the test bench shows no overshoot and approaches its target almost aperiodically, which has previously been identified as performance desiderata.

Table 2: Performance metrics of tuned RL controller for roll evaluated on test bench

	Rise time / s	Steady state / °	Variance / °
RL (continuous), eval 1	4.0	-0.5	0.1
RL (continuous), eval 2	5.0	-0.3	0.9

DOI: 10.13009/EUCASS2023-643

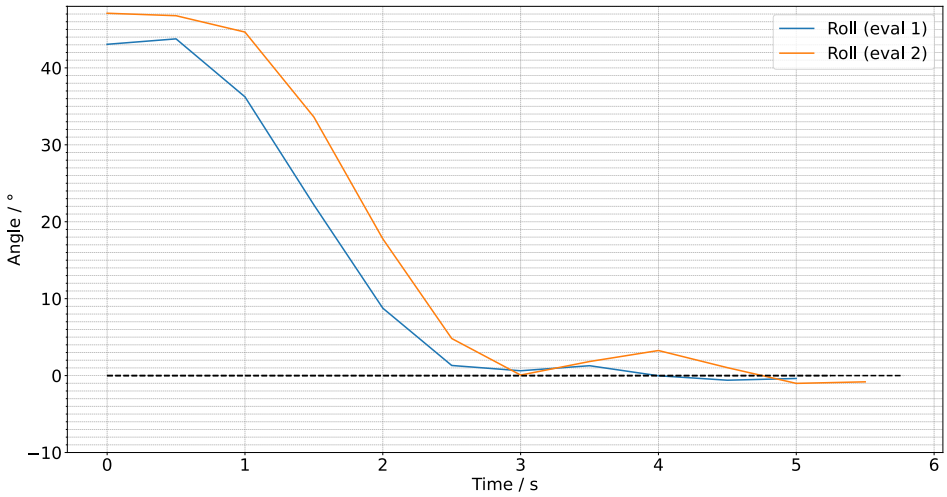

Figure 9: Experimental validation of tuned RL controller for roll

6. Conclusion and Outlook

The Green Propulsion Demonstrator project "The Länder" conducted by the German Aerospace Center (DLR) explored two experimental low cost setups: the 1-DoF and 4-DoF test setups, which provided platforms for investigating control strategies. The 1-DoF setup enabled translational motion perpendicular to the Earth. The 4-DoF setup allowed additional rotational movement along all three axes and featured the use of an embedded system for the ACS. These setups facilitated the evaluation and testing of control strategies in a controlled environment.

Three control strategies were examined: conventional control using PID-type controllers, optimal control, and ML-based control using RL. Conventional control with PID controllers offered a cost-effective approach, the performance was assessed on a test bench at DLR. The optimal control strategy focused on time and efficiency optimization and relied on accurate system knowledge. The ML control strategy, specifically RL, was implemented on an embedded system, allowing for real-time control and adaptability to uncertainties.

The ML controller implemented on the embedded system underwent evaluation on the test bench, which involved deflecting the lander demonstrator at a maximum angle and requiring it to stabilize the system to a desired Euler angle. Initial results showed promising performance with minimal overshoot and successful approach to the target. The RL controller’s performance metrics, such as rise time, steady state, and variance, indicated its effectiveness in controlling the system.

The combination of the test bench setup and the implementation of the RL controller on an embedded system showcased the DLR’s commitment to advance research in aerospace control. The test bench provided a controlled environment for evaluating and refining control strategies, while the integration of the RL controller on an embedded system demonstrated the feasibility of implementing ML-based control in real-time applications.

Looking ahead, "The Länder" has several avenues for future investigation and development. These include the exploration of Model Predictive Control (MPC), the utilization of green propulsion, and the extension of the test setup to a 6-DoF configuration with gravity control.

MPC is a control strategy that optimizes control inputs over a future time horizon while considering dynamic constraints. By incorporating predictions of system behavior, MPC can provide improved performance and robustness compared to traditional control approaches.

Furthermore, the project can explore the use of hot gas propulsion. By studying the control strategies specific to green propulsion, it contributes to the development of the first propulsive landing device with thrusters that use green propellants, developed at DLR Lampoldshausen.

Expanding the test setup to a 6-DoF configuration with gravity control would enable a more comprehensive evaluation of control strategies. In addition to translational and rotational motion, the 6-DoF setup would incorporate control over gravity, allowing for the simulation of different gravitational environments. This extension would provide valuable insights into the performance of control strategies under varying gravitational conditions, expanding the applicability of the findings to different space missions and environments.

DOI: 10.13009/EUCASS2023-643

References

[1] Encyclopedia of systems and control. vol. 2: N - z.

[2] Esa international conference on guidance, navigation control systems (gnc): Topics.

[3] ispace announces results of the "hakuto-r" mission 1 lunar landing.

[4] Nasa’s perseverance rover lands on mars: Here’s how it is using ai.

[5] The surprising creativity of digital evolution: A collection of anecdotes from the evolutionary computation and artificial life research communities. Publisher: arXiv Version Number: 4.

[6] Dirk Abel. *Umdruck zur Vorlesung Regelungstechnik und Ergänzungen (Höhere Regelungstechnik)*. Verlagshaus Mainz, 43. auflage edition.

[7] Gianluca Antonelli, Thor Fossen, and Dana Yoerger. *Underwater Robotics*, volume 1, pages 987–1008. 01 2008.

[8] Karl Johan Astrom and Lars Rundqwist. Integrator windup and how to avoid it. In *1989 American Control Conference*, pages 1693–1698, 1989.

[9] Giulio Avanzini. Spacecraft attitude dynamics and control. *Politecnico di Torino*, 9:13, 2008.

[10] Lukas Bischof. Robust attitude control system based on reinforcement learning for a spacecraft lander demonstrator. In *Institutsseminar*, 2023.

[11] Lars Blackmore. Autonomous precision landing of space rockets. In *Frontiers of Engineering: Reports on Leading-Edge Engineering from the 2016 Symposium*, volume 46, pages 15–20. The Bridge Washington, DC, 2016.

[12] Helmut K. CIEZKI, Victor ZHUKOV, Lukas WERLING, Christoph KIRCHBERGER, Clemens NAUMANN, Martin FRIESS, and Uwe RIEDEL. Advanced Propellants for Space Propulsion A Task within the DLR Interdisciplinary Project "Future Fuels". 2019.

[13] Matthew A. Cooper and Brendon Smeresky. An overview of evolutionary algorithms toward spacecraft attitude control. In Timothy Sands, editor, *Advances in Spacecraft Attitude Control*, chapter 4. IntechOpen, Rijeka, 2020.

[14] Jason Davis. Software Command to Fix Faulty Sensor Doomed Beresheet.

[15] Vasant Dhar. Data science and prediction. *Commun. ACM*, 56(12):64–73, dec 2013.

[16] Michael Dumke, Marco Sagliano, Piyapat Saranrittichai, Guilherme Fragoso Trigo, and Stephan Theil. Eagle - environment for autonomous gnc landing experiments. In *10th International ESA Conference on Guidance, Navigation and Control Systems*, Juni 2017.

[17] Brian Gaudet, Richard Linares, and Roberto Furfaro. Deep reinforcement learning for six degree-of-freedom planetary landing. 65(7):1723–1741.

[18] Amir S. Gohardani, Johann Stanojev, Alain Demairé, Kjell Anflo, Mathias Persson, Niklas Wingborg, and Christer Nilsson. Green space propulsion: Opportunities and prospects. *Progress in Aerospace Sciences*, 71:128–149, November 2014.

[19] M. Golombek, D. Kipp, N. Warner, I. J. Daubar, R. Fergason, R. L. Kirk, R. Beyer, A. Huertas, S. Piqueux, N. E. Putzig, B. A. Campbell, G. A. Morgan, C. Charalambous, W. T. Pike, K. Gwinner, F. Calef, D. Kass, M. Mischna, J. Ashley, C. Bloom, N. Wigton, T. Hare, C. Schwartz, H. Gengl, L. Redmond, M. Trautman, J. Sweeney, C. Grima, I. B. Smith, E. Sklyanskiy, M. Lisano, J. Benardini, S. Smrekar, P. Lognonné, and W. B. Banerdt. Selection of the InSight landing site. 211(1):5–95.

[20] DE Hewes and TC Obryan. Operational features of the langley lunar landing research facility. Technical report, 1967.

[21] Yuan Lin, John McPhee, and Nasser L. Azad. Comparison of deep reinforcement learning and model predictive control for adaptive cruise control. 6(2):221–231.

DOI: 10.13009/EUCASS2023-643

[22] Florian Merz. Grundlegender vergleich unterschiedlicher regelstrategien zur schubvariation und implementierung ausgewählter verfahren in einem 1-dof-kaltgas-demonstrator. In *Institutsseminar*, 2022.

[23] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. 518(7540):529–533.

[24] Justin Moidel. *Rotational motion platform for emulating spacecraft attitude control*. PhD thesis, Citeseer, 2010.

[25] Marco B. Quadrelli, Lincoln J. Wood, Joseph E. Riedel, Michael C. McHenry, MiMi Aung, Laureano A. Cangahuala, Richard A. Volpe, Patricia M. Beauchamp, and James A. Cutts. Guidance, navigation, and control technology assessment for future planetary science missions. 38(7):1165–1186.

[26] Alaa Sheta, Malik Braik, Dheeraj Reddy Maddi, Ahmed Mahdy, Sultan Aljahdali, and Hamza Turabieh. Optimization of PID controller to stabilize quadcopter movements using meta-heuristic search algorithms. 11(14):6492.

[27] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of go with deep neural networks and tree search. 529(7587):484–489.

[28] Evgeny Slyuta. The Luna program. In *Sample Return Missions*, pages 37–78. Elsevier, 2021.

[29] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: an introduction*. Adaptive computation and machine learning. MIT Press.

[30] Rocky Swift. Japan’s ispace launches commercial moon lander, in potential world first.

[31] Massimo Tipaldi, Raffaele Iervolino, and Paolo Roberto Massenio. Reinforcement learning in spacecraft control applications: Advances, prospects, and challenges. 54:1–23.

[32] Sebastian Trimpe and Friedrich Solowjow. Reinforcement learning and learning-based control, April 2022.

[33] Vedant, James T. Allison, Matthew West, and Alexander Robin Mercantini Ghosh. Reinforcement learning for spacecraft attitude control. *Proceedings of the International Astronautical Congress, IAC*, 2019-October, 2019. Funding Information: This material is based upon work partially supported by the National Science Foundation under Grant No. CMMI-1653118. Publisher Copyright: Copyright © 2019 by the International Astronautical Federation (IAF). All rights reserved.; 70th International Astronautical Congress, IAC 2019 ; Conference date: 21-10-2019 Through 25-10-2019.

[34] Mike Wall. Wow! SpaceX lands orbital rocket successfully in historic first.

[35] Günther Waxenegger-Wilfing, Kai Dresia, Jan Deeken, and Michael Oschwald. Machine learning methods for the design and operation of liquid rocket engines–research activities at the dlr institute of space propulsion. *arXiv preprint arXiv:2102.07109*.

[36] Lukas Werling, Dominic Freudenmann, Sophie Ricker, Marius Wilhelm, Felix Lauck, Friedolin Strauss, Konstantin Manassis, Maxim Kurilov, Anna Petrarolo, Till Hörger, Michele Negri, and Christoph Kirchberger. Research and test activities on advanced rocket propellants at dlr’s institute of space propulsion in lampoldshausen. 06 2022.

[37] Matthias Winter, Hans Krüger, Stefano Fari, Jose Luis Redondo Gutierrez, Bronislovas Razgus, Svenja Woicke, David Seelbinder, Calum Hervieu, Adriaen van Camp, Willem Magalhaes Oliveira, Marco Solari, Andreas Wenzel, Stephan Theil, Bolko Maass, Marco Sagliano, Markus Schlotterer, Guilherme Fragoso Trigo, Ansgar Heidecker, René Schwarz, and Stefanie Bremer. Alina moon lander gnc - architecture, design and test results. In *11th International ESA Conference on Guidance, Navigation & Control Systems*, Juni 2021.

[38] Deng-Yun Yu, Ze-Zhou Sun, and He Zhang. *Technology of Lunar Soft Lander*. Springer Singapore.