Deep Reinforcement Learning-Based Autonomous Station-keeping for Geostationary Satellites

Liam Mays*, Jun Chen†
San Diego State University, San Diego, CA 92182

Maintaining geostationary satellites within tight orbital parameters is essential for continuous, high-quality services and reduced operational costs. Traditional station-keeping methods often rely on scheduled maneuvers and extensive ground intervention, limiting flexibility and adaptability. To address these limitations, this thesis proposes an autonomous Deep Reinforcement Learning (DRL) framework designed to discover optimal control policies specifically for geostationary Earth orbit (GEO) maintenance. By accurately modeling the orbital environment at GEO in MATLAB—including high-fidelity spherical-harmonic gravity, solar radiation pressure, and solar/lunar third-body effects—the DRL agent is trained in a realistically perturbed environment and learns to autonomously apply minimal corrective maneuvers without relying on explicit scheduling. This rigorous perturbation modeling contributes to more reliable policy performance under real-world conditions, improving station-keeping precision and fuel efficiency over low-fidelity baselines. This DRL-based station maintenance approach aims to establish a robust, fuel-efficient, and fully autonomous method for managing geostationary orbits, minimizing operational burdens, and extending satellite mission lifetimes.

I. Introduction

Geostationary orbits have played a critical role in modern satellite communication since their conceptualization. The idea was popularized by science fiction writer Arthur C. Clarke in the 1940s, who envisioned these orbits as revolutionary platforms for global telecommunications. Clarke recognized the unique property of geostationary orbits: satellites placed approximately 35,786 kilometers above Earth's equator match the planet's rotational period, allowing them to maintain a constant position relative to observers on the ground. This concept became a reality when the first geostationary satellite, Syncom 3, was successfully launched in 1963. Since then, satellites in geostationary orbits have become integral to communication networks, weather monitoring, broadcasting, navigation, and various other essential services.

Maintaining satellites in precise geostationary positions is increasingly critical due to the densely populated nature of this limited orbital region. Any orbital drift or perturbation-induced displacement can lead to communication interference, degraded service quality, operational inefficiencies, or even mission failure [1]. Perturbations such as gravitational anomalies arising from Earth's uneven mass distribution [2], solar radiation pressure [3], and gravitational influences from the Moon and Sun [4] continuously challenge satellite orbit stability. Consequently, periodic station-keeping maneuvers are necessary to correct these deviations and ensure operational integrity [5].

Traditional station maintenance, which involves periodic maneuver scheduling and extensive ground-based intervention, faces limitations in flexibility, responsiveness, and cost-effectiveness. As space activities escalate and orbital slots become more crowded, conventional methods struggle with these dynamic and increasingly complex environments [6]. The manual planning and execution of these corrective maneuvers often result in significant operational overhead and suboptimal resource use, highlighting the need for more autonomous, adaptable strategies [7, 8].

Autonomous station-keeping methods, particularly advanced techniques utilizing Deep Reinforcement Learning (DRL), present promising solutions to these limitations. DRL enables satellites to dynamically and autonomously adapt to perturbations in real-time without pre-scheduled interventions. Recent research has shown DRL's potential [9] in trajectory planning [10], low-thrust orbit insertion maneuvers [10], and general spacecraft guidance [12], illustrating the method's versatility in aerospace applications. Through reinforcement learning, a satellite can continuously improve its orbit maintenance strategy based on environmental feedback, optimizing fuel usage and potentially extending mission lifetimes significantly.

The research presented here introduces a novel autonomous DRL-based framework. By combining advanced DRL methods with high-fidelity orbital modeling in MATLAB, this research addresses both practical and economic

* M.S. Student, Department of Aerospace Engineering, AIAA Student Member, lmays4253@sdsu.edu
† Associate Professor, Department of Aerospace Engineering, AIAA Senior Member, Jun.Chen@sdsu.edu

challenges inherent in traditional station-keeping methods. The proposed autonomous framework aims not only to reduce operational burdens but also to offer robust and fuel-efficient solutions suitable for the demanding environment of geostationary orbit management.

II. Environment Modeling

This section outlines the methodological approach employed to create the environment in which the DRL agent can learn, detailing the satellite dynamics modeling, which forms the foundation for the development of a robust station-keeping framework.

In reality, geostationary satellites are subject to a variety of perturbative forces that drive their orbital elements away from ideal values. At GEO altitude, the dominant perturbations are Earth's non-spherical gravity field, solar radiation pressure, and third-body gravitational pulls from the Sun and Moon [3]. Figure 1 provides a conceptual illustration of the station-keeping challenge: perturbations push the satellite outside its nominal ±0.05° geodetic confinement box, requiring periodic corrective impulses.

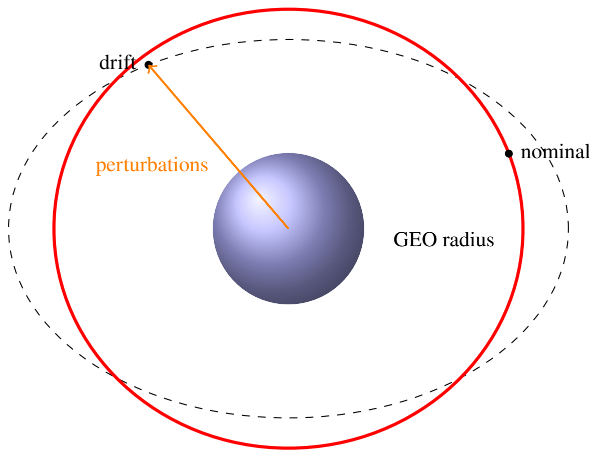

Fig. 1 Conceptual illustration of perturbed satellite at GEO

To capture these real-world effects quantitatively, we introduce higher-fidelity orbital models in the subsequent subsections, beginning with spherical-harmonic gravity.

A. Spherical Harmonics

The Earth is not a perfect sphere, but rather an oblate spheroid with variations in its mass distribution [1]. To account for these irregularities in Earth's gravitational potential, we use spherical harmonics [2]. This model expands the gravitational potential V as an infinite series of harmonics (Equation (1)), allowing for greater accuracy. The gravitational potential in spherical harmonics is expressed as:

$$V = \frac{\mu}{r} \sum_{l=0}^{\infty} \left(\frac{a_c}{r}\right)^l \sum_{m=0}^{l} P_{l,m}(\sin\phi) \left(C_{l,m}\cos(m\lambda) + S_{l,m}\sin(m\lambda)\right) \tag{1}$$

where:

- r is the radial distance from the Earth's center,
- a_c is the Earth's equatorial radius,
- ϕ is the geocentric latitude,
- λ is the geocentric longitude,
- $P_{l,m}$ are the associated Legendre polynomials,
- $C_{l,m}$ and $S_{l,m}$ are the spherical harmonic coefficients for degree l and order m.

The associated acceleration due to the spherical-harmonic gravity field is simply the gradient of V. Equation (2) forms the basis for our high-fidelity gravity model.

$$\mathbf{a}_{\mathrm{SH}} = \nabla V. \tag{2}$$

The primary term in this expansion corresponds to the simple two-body potential $V = \frac{\mu}{r}$, which assumes a spherically symmetric Earth. However, the $\mathbf{a}_{\mathrm{SH}}$ encapsulates all the $J_2, J_3, \ldots$ perturbations through the harmonic coefficients $C_{l,m}, S_{l,m}$.

B. Solar Radiation Pressure

Solar Radiation Pressure (SRP) is a perturbative force caused by the momentum transfer from sunlight as it strikes a satellite. The calculation of SRP begins by determining the position of the Sun relative to the satellite. Using the ephemeris data for the Sun, the satellite's position vector $\mathbf{r}$ relative to the Sun is obtained [4]. The force due to solar radiation is computed based on the distance between the Sun and the satellite, as well as the solar flux at that distance.

To account for variations in solar intensity due to Earth's elliptical orbit, the solar flux SF is calculated using Vallado's formulation [3]:

$$SF = \frac{1358}{1.004 + 0.0334\cos(\text{days from aphelion})} \tag{3}$$

where 1358 W/m^2 is the average solar flux at 1 AU, and the denominator accounts for variations in solar distance.

The solar radiation pressure p_{SRP} is then given by:

$$p_{SRP} = \frac{SF}{c} \tag{4}$$

where c is the speed of light in a vacuum.

The resulting acceleration on the satellite due to SRP is calculated as [3]:

$$\mathbf{a}_{SRP} = -\frac{c_R \cdot p_{SRP} \cdot A_{\mathrm{sat}}}{m_{\mathrm{sat}}} \frac{\mathbf{r}_{\mathrm{rel,Sun}}}{|\mathbf{r}_{\mathrm{rel,Sun}}|} \tag{5}$$

where:

- c_R is the satellite's reflectivity coefficient,
- A_{sat} is the cross-sectional area exposed to sunlight,
- m_{sat} is the mass of the satellite,
- $\mathbf{r}_{\mathrm{rel,Sun}}$ is the vector from the satellite to the Sun.

This acceleration acts in the direction opposite to the vector from the satellite to the Sun, gradually altering the satellite's orbit over time.

C. Third-Body Perturbations from the Moon and Sun

Additionally, third-body gravitational effects from the Sun and Moon introduce further perturbations. The gravitational acceleration on the satellite due to a third body is described by [3]:

$$\mathbf{a}_{3B} = -\frac{\mu_{\oplus}\mathbf{r}_{\oplus sat}}{r_{\oplus sat}^3} + \mu_3\left(\frac{\mathbf{r}_{sat3}}{r_{sat3}^3} - \frac{\mathbf{r}_{\oplus 3}}{r_{\oplus 3}^3}\right) \tag{6}$$

where:

- $\oplus$ Denotes Earth,
- **3** Denotes the third-body (Sun or Moon),
- **sat** Denotes the satellite,

Having detailed each of the principal perturbative forces—Earth's non-spherical gravity field, solar radiation pressure, and third-body influences from the Sun and Moon—shown below:

$$\mathbf{a} \;=\; \mathbf{a}_{\mathrm{SH}} + \mathbf{a}_{\mathrm{SRP}} + \mathbf{a}_{\mathrm{3B}}. \tag{7}$$

we now assess their combined impact on a geostationary satellite's orbital elements. Figure 2 presents the time evolution of semi-major axis, eccentricity, and inclination when all perturbations act simultaneously. The gradual divergence from their initial values underscores why continuous station-keeping maneuvers are essential, and motivates the autonomous DRL strategy developed in this work.

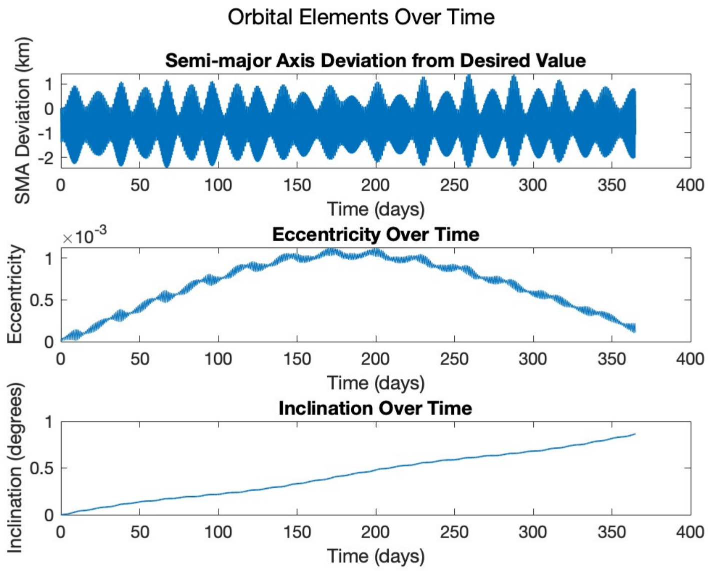

Fig. 2 Propagation of orbital elements with the inclusion of perturbations.

III. DRL Framework and Markov Decision Process (MDP)

Given the demonstrated divergence from desired orbital parameters, a unique DRL framework was developed within MATLAB to autonomously explore and learn the complex orbital dynamics environment. By interacting with a carefully constructed high-fidelity simulation environment, the agent learns to maintain the desired geostationary orbit through the efficient application of impulsive maneuvers.

A. MATLAB Simulation Environment

All satellite dynamics are propagated using MATLAB's built-in `ode45` solver, which employs an adaptive Runge–Kutta scheme to accurately capture the motion. At each decision epoch, the DRL agent selects an impulsive $\mathbf{\Delta}$V vector that is applied directly to the satellite's velocity state; following this instantaneous "kick," `ode45` resumes integration from the updated state for a duration determined by the agent's adaptive time-stepping policy. This loop—perturbed dynamics → agent action → velocity update → `ode45` propagation—enables seamless modeling of impulsive maneuvers within a high-fidelity environment. The use of `ode45` is particularly appropriate here because it is a standard astrodynamics tool, offering robust error control, and wide community validation.

B. MDP Definition

To formalize the station-keeping problem for the DRL agent, we define the underlying Markov Decision Process (MDP) as follows:

States At each decision instant the agent observes a 9-dimensional state vector:

$$\mathbf{s}_t = \left[\, \widetilde{\Delta a},\ \widetilde{\Delta e},\ \widetilde{\Delta i},\ \lambda,\ F_{\mathrm{NS}},\ F_{\mathrm{EW}},\ \theta_{\mathrm{asc}},\ \theta_{\mathrm{desc}},\ \nu_{\mathrm{EW}} \right]^\top, \tag{8}$$

where

- $\widetilde{\Delta a}$, $\widetilde{\Delta e}$, $\widetilde{\Delta i}$ are the semi-major axis, eccentricity, and inclination deviations normalized by their station-keeping bounds,
- λ is the satellite's geocentric longitude,
- F_{NS}, F_{EW} are the remaining fuel fractions (normalized) allocated for north-south and east-west maneuvers,
- θ_{asc}, θ_{desc} are the normalized angular distances from the ascending and descending nodes,
- ν_{EW} is the normalized true anomaly relevant for east-west burns.

Actions A unique method was employed to allow the agent to find and exploit fuel-efficient maneuvers. The agent issues a 6-dimensional action:

$$\mathbf{a}_t = \left[u_{\mathrm{NS}},\ u_{\mathrm{EW}},\ \Delta V_x,\ \Delta V_y,\ \Delta V_z,\ \tau\right]^\top, \tag{9}$$

where

- $u_{\mathrm{NS}}, u_{\mathrm{EW}} \in [0, 1]$ are decision flags for north-south and east-west burns,
- $\Delta V_x, \Delta V_y \in [-\Delta V_{\mathrm{max,EW}}, \Delta V_{\mathrm{max,EW}}]$ and $\Delta V_z \in [-\Delta V_{\mathrm{max,NS}}, \Delta V_{\mathrm{max,NS}}]$ specify the Cartesian components of the impulsive maneuver,
- $\tau \in [0, 1]$ is a "wait" scaler that determines the next propagation interval from 0.5 to 16 days.

Reward At each simulation time step t, the agent receives

$$r_t = \begin{cases} r^{\mathrm{track}}(t), & \text{sub-steps of burn times } t, \\ r^{\mathrm{track}}(t) + r^{\mathrm{action}}(t), & \text{at burn times } t, \end{cases} \tag{10}$$

where

$$r^{\mathrm{track}}(t) = -w_{\mathrm{orbit}}\, \|\Delta\mathbf{EO}(t)\| \quad \text{and} \quad r^{\mathrm{action}}(t) = -\big(w_{\mathrm{fuel}}\|\Delta\mathbf{V}(t)\| + w_{\mathrm{phase}}\,\Phi(t)\big).$$

Here, $\Delta\mathbf{EO}(t)$ is the orbital-element error vector, $\|\Delta\mathbf{V}(t)\|$ is the impulsive burn magnitude when applied, $\Phi(t)$ is an orbital phase penalty at the burn location, and $w_{\mathrm{orbit}}, w_{\mathrm{fuel}}, w_{\mathrm{phase}} > 0$ are weighting factors.

The weights $w_{\mathrm{orbit}}, w_{\mathrm{fuel}}, w_{\mathrm{phase}}$ and the precise forms of the states, actions, and rewards were designed based on preliminary studies and will be systematically optimized in the full study.

DRL Agent Employed an off-policy, actor–critic Deep Deterministic Policy Gradient (DDPG) agent to handle the continuous action and observation spaces inherent to station-keeping. A high-level workflow diagram (Figure 3) illustrates how the agent's policy and value networks interact with the MATLAB simulation environment to select and evaluate impulsive maneuvers.

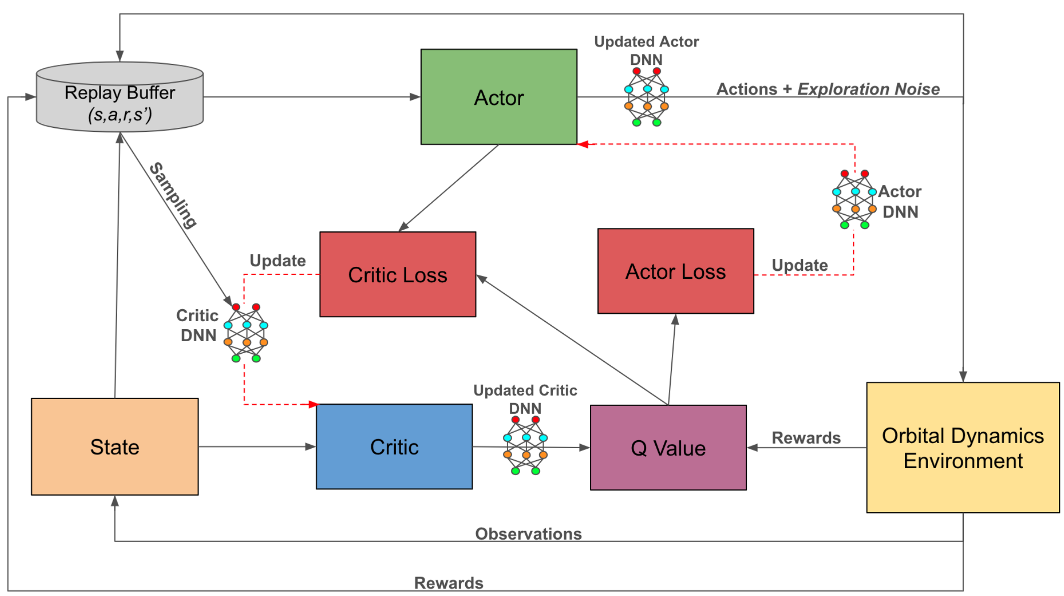

Fig. 3 Block diagram of workflow

In the full paper, the trained DRL agent will be evaluated over multiple one-year simulation runs to quantify its station-keeping performance. Key metrics include the ability of the satellite to remain within a $\pm 0.05°$ geodetic latitude and longitude confinement box, and the total ΔV expended per year (m/s/year). Successful performance is defined by maintaining orbital deviations within the specified bounds while keeping annual fuel usage at or below the allocated budget, thereby demonstrating both high positional accuracy and fuel efficiency.

Additionally, this study will evaluate the DRL agent's robustness across multiple longitudinal slots in the geostationary belt. Because Earth's gravitational equilibrium points cause perturbation magnitudes to vary with longitude [3], we will compare station-keeping performance and fuel requirements at several representative positions. These extensions will demonstrate the generality of the autonomous framework under diverse orbital conditions and inform optimal deployment strategies for real-world GEO missions.

References

[1] Machuca, P., "Orbital Dynamic Earth Environment Modeling," , 2024. Lecture notes, San Diego State University, Spring 2024.

[2] National Geospatial-Intelligence Agency, "WGS84 Earth Info," , ???? `https://earth-info.nga.mil/index.php?dir=wgs84&action=wgs84`.

[3] Vallado, D. A., *Fundamentals of Astrodynamics and Applications*, 4th ed., Microcosm Press, Hawthorne, CA, 2013.

[4] NASA Jet Propulsion Laboratory, "NAIF SPICE Toolkit Documentation," , ???? `https://naif.jpl.nasa.gov/pub/naif/toolkit_docs/C/info/intrdctn.html`.

[5] [Skipper, J. K., Li, S., Pitre, G., Sim, F. A. S., and Provencher, R., "Autonomous Reconstruction and Calibration of GEO Stationkeeping Maneuvers," *SpaceOps 2006 Conference Proceedings*, American Institute of Aeronautics and Astronautics, 2006. Paper AIAA 2006-5590.

[6] Herrera, I., Armando, "Reinforcement Learning Environment for Orbital Station-Keeping," Master's thesis, The University of Texas Rio Grande Valley, Edinburg, TX, December 2020.

[7] LaFarge, N. B., Howell, K. C., and Folta, D. C., "An Autonomous Stationkeeping Strategy for Multi-Body Orbits Leveraging Reinforcement Learning," *Proceedings of the AIAA/AAS Astrodynamics Specialist Conference*, American Institute of Aeronautics and Astronautics, 2020.

[8] Zhang, J., Shen, A., and Li, L., "Minimum-Fuel Geostationary East-West Station-Keeping Using a Three-Phase Deep Neural Network," *Acta Astronautica*, Vol. 204, 2023, pp. 500–509.

[9] Tipaldi, M., Iervolino, R., and Massenio, P. R., "Reinforcement Learning in Spacecraft Control Applications: Advances, Prospects, and Challenges," *Annual Reviews in Control*, Vol. 47, 2022, pp. 1–23.

[10] Kapilavai, D., Roychowdhury, S., Fu, Y., Yu, L., and van Bloemen Waanders, B., "Trajectory Planning for Aerospace Vehicles using Deep Reinforcement Learning," *AIAA SciTech 2021 Forum*, American Institute of Aeronautics and Astronautics, 2021.

[11] Sullivan, C. J., and Bosanac, N., "Using Reinforcement Learning to Design a Low-Thrust Approach into a Periodic Orbit in a Multi-Body System," *AAS/AIAA Space Flight Mechanics Meeting*, American Astronautical Society & AIAA, Boulder, CO, 2021.

[12] Hovell, K., and Ulrich, S., "On Deep Reinforcement Learning for Spacecraft Guidance," *AIAA Scitech 2020 Forum*, Carleton University, Ottawa, Ontario, Canada, 2020.