A Data-Driven Census of Cislunar Orbital Stability Enabled by Volunteer Computing

Lezhe Gao [*] , David P. Anderson , Vitalii Koshura

Posted Date: 7 April 2026

doi: 10.20944/preprints202604.0498.v1

Keywords: cislunar dynamics; orbital stability; circular restricted three-body problem; volunteer computing; BOINC; Jacobi constant

doi:10.20944/preprints202604.0498.v1

Article

A Data-Driven Census of Cislunar Orbital Stability Enabled by Volunteer Computing

Lezhe Gao [1,*], David P. Anderson [2] and Vitalii Koshura [3]

[1] Independent Researcher, Room 2001, Building 5, Lane 1555, Yinhang Rd., Shanghai, China
[2] University of California, Berkeley, California, USA
[3] Independent Researcher, Friedrichshafen, Germany
* Correspondence: gaolezhe@outlook.com

Abstract

The cislunar space, governed by the circular restricted three-body problem (CR3BP), presents significant challenges for mission design due to its complex stability structure. Traditional high-fidelity numerical integration is computationally prohibitive for a systematic stability census of millions of orbits. Here, we present a novel approach based on global volunteer computing via the BOINC platform to overcome this barrier. Using the public "Million Orbit" dataset from Lawrence Livermore National Laboratory, we distributed the computation of Jacobi constant time series across thousands of volunteer devices, producing over 16 billion individual values. The resulting dataset is freely available. Analysis reveals that 91.68% of orbits belong to the high-energy Region V, 8.07% to the stable Region I, and only 0.24% to Region III, with Region II completely absent. A single rare Region IV orbit (ID 754482) was identified and analyzed. This work demonstrates the transformative potential of volunteer computing for large-scale astrodynamics research, providing a detailed stability map and a benchmark for future machine-learning applications.

Keywords: cislunar dynamics; orbital stability; circular restricted three-body problem; volunteer computing; BOINC; Jacobi constant

1. Introduction

The cislunar realm, the gravitational corridor encompassing Earth and the Moon, has re-emerged as a critical domain for future space exploration, satellite deployment, and deep space gateway operations. Its dynamics are primarily governed by the Circular Restricted Three-Body Problem (CR3BP), which gives rise to a complex tapestry of orbital families, interwoven with chaotic manifolds and instability regions. Navigating this environment requires a precise, global understanding of orbital stability.[1,2]

Traditional approaches rely on numerical integration of equations of motion, which, while accurate, are computationally prohibitive when conducting the vast parameter studies necessary for a comprehensive stability census. This creates a critical bottleneck: due to the sheer volume of calculations required, high-fidelity mapping of the cislunar space at the scale needed for systematic mission design remains an elusive goal, even with centralized supercomputing resources.

Meanwhile, the volunteer computing paradigm has matured over the past two decades, proving its capability to tackle problems of unprecedented scale. Pioneered by projects such as SETI@home, which harnessed idle personal computers to analyze radio telescope data in the search for extraterrestrial intelligence, the model demonstrated the potential of aggregating global, heterogeneous computing power.[3] This was extended to astrophysics by projects such as Einstein@Home, which uses volunteer computing to search for gravitational waves and pulsars, leading to numerous peer-reviewed discoveries.[4]

These successes demonstrate the capability of volunteer computing for high-throughput computing across diverse domains. However, within the field of classical orbital mechanics—a discipline with inherently parallelizable, high-throughput computational needs perfectly suited to this distributed model— volunteer computing campaigns focused on cislunar orbital stability analysis have been notably absent.

This work addresses this gap by applying volunteer computing to a foundational problem in astrodynamics. We use the fully public "Million Orbit" data set—a product of a massive supercomputing campaign by Lawrence Livermore National Laboratory[5,6]—as the input for a novel, distributed stability analysis. While the data set provides the raw trajectories, our project repurposes it for global stability analysis. We developed a containerized application running on the BOINC Central platform to calculate the Jacobi constant time series for all sampled points along each trajectory in the data set, performing a stability census at a scale and cost-effectiveness impractical for traditional centralized resources.

Our study yields a detailed stability map of the cislunar space, quantifying regions of long-term stability and rapid escape. Beyond the specific findings, this project establishes a novel, scalable methodology for astrodynamics research. It demonstrates that volunteer computing is not only a viable but a transformative tool for exploring complex gravitational systems, opening a new pathway for large-scale numerical investigations that complement and extend traditional high-performance computing approaches.

2. Methods

2.1. The "Million Orbit" Reference Data Set

The computational foundation of this census is the fully public "One Million Open-source Cislunar Orbits" data set, generated through a massive supercomputing campaign by Lawrence Livermore National Laboratory (LLNL).[5,6] This data set provides a bases for the study of cislunar dynamics within the Circular Restricted Three-Body Problem (CR3BP) framework.

The data set comprises one million unique spacecraft trajectories, each numerically integrated over a simulated period of six years. The orbits are initialized from geosynchronous Earth orbit to regions beyond the Moon, encompassing the dynamically complex regions between the Earth and the Moon, including Lagrange points and their associated manifolds. Each orbit is stored in a standardized Hierarchical Data Format version 5 (HDF5) file[7], containing time-series arrays of the spacecraft's three-dimensional position and velocity vectors in the Earth-Moon rotating frame. [5,6]

For our distributed computing paradigm, this data set presents ideal characteristics:

1) Web access: All data files are hosted on a public LLNL repository with unrestricted access.

2) Inherent modularity: The million orbits are stored in independently accessible files and logical groups, allowing for trivial decomposition into discrete computational units.

3) High scientific value: As a product of a verified high-performance computing (HPC) integration, it provides a benchmark-quality source for secondary analysis, distinguishing our stability metrics from those derived from ad-hoc simulations.

In this study, we treat this data set not as the subject of new numerical integration, but as the primary input for a subsequent, large-scale analytical computation—the derivation of Jacobi constant time series and stability indices for every data point across all trajectories.

2.2. Task Definition: Aligning BOINC Work Units with Existing Data Structure

The architecture of the source "Million Orbit" data set is modular, with trajectories divided into 20,000 HDF5 files, each containing 50 independent orbital groups. This structure enabled a direct and efficient mapping to the BOINC model[8]: each job (or 'work unit') handled one complete HDF5 file, thereby eliminating the need for further data segmentation or pre-processing.

However, because the total size of all HDF5 files amounts to terabytes, we did not embed them in work units. Instead, each work unit contains only a JavaScript Object Notation (JSON) data interchange format file[9] (input.json), which contains two key parameters:

doi:10.20944/preprints202604.0498.v1

1. Data Source Parameters, including the URL of the target .h5 file on the LLNL server[10] and the specific file identifier.
2. Task ID, which is used to identify tasks and to enable debugging.

The task of each work unit was to: (i) fetch the assigned HDF5 file, (ii) calculate the time series of the Jacobi constant for all 50 trajectories within it, and (iii) return the derived data set (also in JSON) along with integrity-validation metadata in a result file.

This “one-file-per-task” granularity—corresponding to 50 orbits per computation—was chosen to balance multiple factors. It ensured that individual tasks required several hours of computation on typical volunteer hardware, optimizing the use of donated resources while remaining within practical completion time windows. This granularity also made the management overhead of 20,000 total tasks tractable and provided fault tolerance. The failure of any single task affected only 0.005% of the total data set, allowing for swift re-computation with negligible impact on the overall campaign. To ensure result accuracy, we used redundant computing: we randomly run some of the jobs twice and check if the results agree.

2.3. Design and Implementation of the BOINC Application

To execute the stability census on volunteer resources, we developed a custom application using BOINC's Universal Docker Application (BUDA) framework. The application was designed with three primary goals: (1) integration with the existing BOINC infrastructure, (2) fault tolerance to handle the heterogeneous and intermittent nature of volunteer nodes, and (3) reproducibility of the computational environment.

The code for fetching the HDF5 files, parsing orbit data, and computing the Jacobi constant time series (as described in the jacobi_calculator.py module), was written in Python. To ensure a consistent runtime environment across the diverse range of volunteer operating systems (Windows, macOS, Linux), we encapsulated the application, along with all its dependencies (NumPy, SciPy, h5py, etc.), in a Docker container. The Dockerfile defines an image based on python:3.9-bullseye, which includes all the required libraries. This containerization strategy guarantees that each work unit executes in an identical software environment, eliminating the "it works on my machine" problem inherent in distributed computing.[11]

Each BOINC work unit corresponds to a single JSON task file (input.json), as detailed in the 'Task Definition' section. The application workflow, orchestrated by buda_main.py, is as follows:

Initialization: The application sets up the configurations and reads input.json to obtain the target HDF5 file's URL and other parameters.

Remote Data Fetching: It downloads the required HDF5 file from the public LLNL repository.

Checkpointed Computation: It iterates through each of the 50 orbit groups within the downloaded file. A useful feature for volunteer computing is a checkpointing mechanism. Every 300 seconds, the application saves its intermediate state—including the computed results for that orbit and the index of the last completed group—to a checkpoint file. If a volunteer's computer is interrupted and later resumes the same work unit, the application detects the existing checkpoint file and resumes processing from where it left off, rather than starting over. This prevents the loss of computational progress and ensures efficient use of donated cycles.

Output and Cleanup: Upon successful processing of all 50 orbit groups, the application generates the final output.json file containing all results and metadata. If the process encounters a non-recoverable error, it logs the error to the output file and exits gracefully.

3. Results

3.1. Computation Overview

The volunteer computing campaign successfully processed all 20,000 work units, covering the complete set of one million orbits. A total of approximately 16 billion individual Jacobi constant values were computed, with an average of ~16,000 points per orbit.

The software and result dataset has been published to Zenodo. [11,12]

3.2. Orbital Stability Classification

Based on the classical CR3BP energy thresholds (C_1, C_2, C_3, C_4) for $\mu = 0.01215$, each orbit was assigned to one of five stability regions according to the proportion of time steps falling into each region. The distribution is as follows (Table 1):

Table 1. Distribution of one million cislunar orbits across the five classical stability regions.

Region	Condition	Number of Orbits	Percentage
Region I	$C > C_1$	80,748	8.07%
Region II	$C_1 \geq C > C_2$	0	0%
Region III	$C_2 \geq C > C_3$	2,430	0.24%
Region IV	$C_3 \geq C > C_4$	1	0.0001%
Region V	$C \leq C_4$	916,821	91.68%

Notably, Region II is completely absent, and Region IV is represented by a single orbit (ID 754482). The Jacobi constant time series for this unique orbit is shown in Fig. 1. Its behavior exhibits a distinctive pattern that may indicate a transitional state between stability and escape.

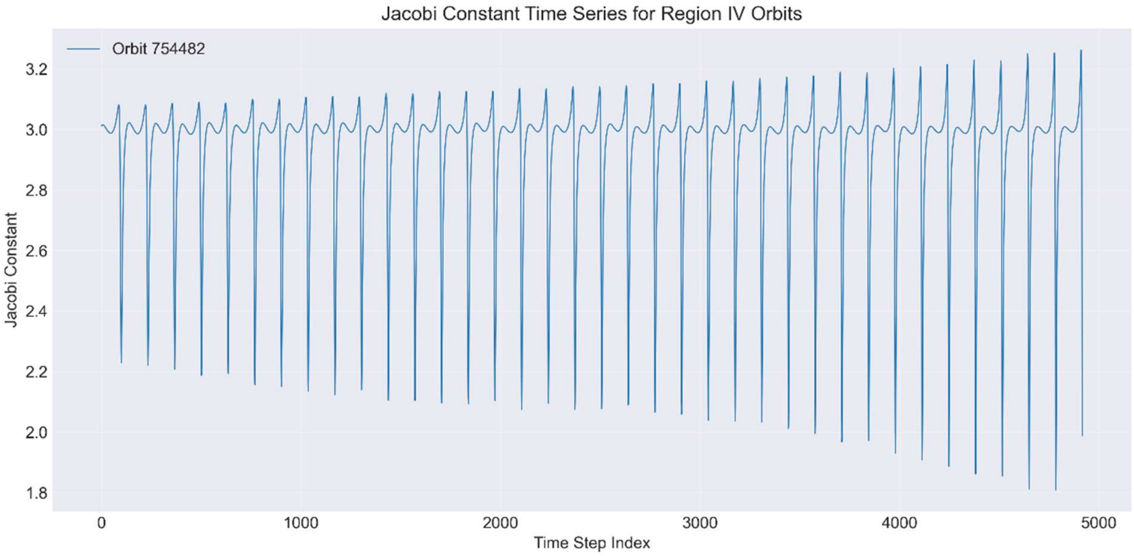

Figure 1. Jacobi constant time series for the sole Region IV orbit (orbit ID 754482).

3.3. High-Energy Orbits

A total of 173 orbits were found to have maximum Jacobi constant exceeding 100. All of these belong to Region V and correspond to high-energy (possibly escape) trajectories. These values have been verified independently.

doi:10.20944/preprints202604.0498.v1

4. Discussion

The overwhelming majority of orbits (91.68%) fall into Region V, the highest energy region, indicating that most trajectories in the LLNL dataset are either distant retrograde orbits or escape orbits. This is consistent with the dataset's design, which spans from GEO to beyond the Moon.

Region I, corresponding to the most stable low-energy orbits (e.g., halo, Lyapunov families), comprises 8.07% of the dataset. These orbits are of particular interest for long-term missions such as lunar gateways.

The complete absence of Region II orbits suggests a dynamic gap: no trajectories occupy the energy range between the stable low-energy region and the intermediate Region III. This may reflect a natural separation in the phase space of the Earth-Moon system.

The single Region IV orbit (ID 754482) is a rare find. Its Jacobi constant hovers just above C_4 for most of the six-year integration, occasionally dipping below. This behavior could indicate a near-escape trajectory that remains bound for a long duration, possibly due to weak chaos or resonance effects. Further analysis of its Lyapunov exponent would be valuable.

The 173 high-energy orbits ($C > 100$) likely represent numerical artifacts or extremely fast escapes. Manual verification confirmed they are not the result of software errors.

The volunteer computing approach proved highly effective: 20,000 tasks were completed in approximately two months, at zero financial cost. This demonstrates the potential of BOINC for large-scale astrodynamics studies.

Acknowledgments: We are deeply grateful to the thousands of volunteers who contributed their idle computing power through the BOINC Central platform, making this large-scale stability analysis possible. Their generosity and enthusiasm are the foundation of this work. We also extend our sincere thanks to Travis Yeager, Ph.D., and the Lawrence Livermore National Laboratory (LLNL) for generating and publicly releasing the *One Million Open-source Cislunar Orbits* dataset, which served as the essential input for our computations.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process: During the preparation of this work the author(s) used DeepSeek in order to polish the article, to improve readability, to refine logic flow and to optimize the codes which is needed for analyzing the data. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

References

1. Leiva, A.M., Briozzo, C.B., 2006. Control of chaos and fast periodic transfer orbits in the Earth–Moon CR3BP. Acta Astronaut. 58(8). https://doi.org/10.1016/j.actaastro.2005.12.006
2. Szebehely, V., 1967. Theory of Orbits: The Restricted Problem of Three Bodies. Academic Press, New York, pp. 13–16.
3. Anderson, D.P., Cobb, J., Korpela, E., et al., 2002. SETI@home: an experiment in public-resource computing. Commun. ACM 45(11). https://doi.org/10.1145/581571.581573
4. Knispel, B., Eatough, R.P., Kim, H., et al., 2013. EINSTEIN@HOME discovery of 24 pulsars in the Parkes Multi-beam Pulsar Survey. Astrophys. J. 774(2), 93. https://doi.org/10.1088/0004-637x/774/2/93
5. Yeager, T., Higgins, D., McGill, P., et al., 2025. One Million Open-source Cislunar Orbits [dataset]. Res. Notes AAS 9(8), 215. https://doi.org/10.3847/2515-5172/adf8ec
6. Yeager, T., Higgins, D., McGill, P., et al., 2025. An open benchmark of one million high-fidelity cislunar trajectories. arXiv:2512.11064.
7. Open Geospatial Consortium, 2018. OGC Hierarchical Data Format Version 5 (HDF5) Core Standard. https://portal.ogc.org/doi/18-043r3 (accessed 15 February 2026).
8. Anderson, D.P., 2020. BOINC: A Platform for Volunteer Computing. Journal of Grid Computing 18(1). https://doi.org/10.1007/s10723-019-09497-9

9. RFC Editor, 2017. The JavaScript Object Notation (JSON) Data Interchange Format. https://www.rfc-editor.org/info/rfc8259 (accessed 15 February 2026).
10. Index of cislunar data. https://gdo-cislunar.llnl.gov/cislunar_data_six_year_1.0GEO_to_18.2GEO/ (accessed 15 February 2026).
11. Gao, L., 2026. BOINC BUDA Application for Cislunar Orbit Stability Analysis v1 (Version 1)[software]. Zenodo. https://doi.org/10.5281/zenodo.18835416
12. Gao, L., Anderson, D. P., & Koshura, V., 2026. Jacobi Constant Time Series for One Million Cislunar Orbits Derived via Volunteer Computing [dataset]. Zenodo. https://doi.org/10.5281/zenodo.19212057