

International Journal of Statistics and Applied Mathematics



ISSN: 2456-1452
Maths 2025; 10(6): 312-315
© 2025 Stats & Maths
<https://www.mathsjournal.com>
Received: 18-04-2025
Accepted: 20-05-2025

Masashi Yoshida
Department of Mathematics,
Sapporo Research Institute of
Data Sciences, Sapporo,
Hokkaido, Japan

Rina Kato
Department of Mathematics,
Sapporo Research Institute of
Data Sciences, Sapporo,
Hokkaido, Japan

Examining floer homology computations for lagrangian intersections with applications to Kaziranga rhino census methodology

Masashi Yoshida and Rina Kato

DOI: <https://www.doi.org/10.22271/math.2025.v10.i6d.2488>

Abstract

Floer homology detects intersection points of Lagrangian submanifolds in symplectic geometry—a topological invariant that, surprisingly, can be re-interpreted as a counting framework for wildlife census data when observation trajectories are embedded in a phase space. This research examines computational methods for Floer homology on compact symplectic manifolds and applies the intersection-counting machinery to estimate the Indian one-horned rhinoceros population in Kaziranga National Park. Elephant-back and drone survey transects from five census zones, collected across five seasonal rounds between March 2021 and April 2023, were encoded as Lagrangian curves in a position-detection phase space. The Floer differential was computed using a finite-dimensional Morse-theoretic approximation with perturbation terms calibrated to detection heterogeneity. The resulting population estimate was 2,613 rhinos (95% CI: 2,418-2,831), compared with 2,587 from a standard distance-sampling model (95% CI: 2,371-2,824) and 2,649 from a mark-resight analysis. Computation time for the Floer approach averaged 3.7 seconds per zone, comparable to maximum-likelihood fitting of conventional models. Cross-validation with held-out transects yielded a mean absolute percentage error of 4.8%, versus 6.3% for distance sampling. These results suggest that symplectic-geometric counting provides a mathematically coherent alternative to classical abundance estimators, especially in settings where transect overlaps create natural Lagrangian intersection structure.

Keywords: Floer homology, Lagrangian intersection, symplectic geometry, rhino census, Kaziranga, population estimation, detection probability, Morse theory, phase space embedding, wildlife monitoring

Introduction

Specialists in wildlife census rarely look to symplectic geometry for tools, and symplectic geometers rarely think about rhinoceroses. Yet the mathematical structure underlying both fields shares a common thread: counting intersections of curves in a constrained space ^[1]. In Floer homology, one counts intersection points of Lagrangian submanifolds by examining pseudo-holomorphic strips that connect them—a count that turns out to be a topological invariant, unchanged by smooth deformations of the manifolds ^[2]. In wildlife census, one counts animals observed along survey transects and corrects for imperfect detection—a count that must account for the geometric overlap of observer paths ^[3].

Kaziranga National Park in Assam, India, shelters roughly two-thirds of the world's Indian one-horned rhinoceros (*Rhinoceros unicornis*) population ^[4]. Census methods have evolved from ground-based total counts on elephant back to hybrid designs incorporating drones, camera traps, and distance sampling ^[5]. A persistent challenge is that transects often cross, re-cross, and share boundary segments, creating a complex web of overlapping observation corridors whose detection probabilities interact in ways that distance-sampling models handle only approximately ^[3].

This research embeds survey transects as curves in a two-dimensional symplectic phase space (position $\times$ detection probability) and computes their Floer homology to extract a topologically robust intersection count that corresponds to the number of distinct individuals detected ^[1].

Corresponding Author:
Masashi Yoshida
Department of Mathematics,
Sapporo Research Institute of
Data Sciences, Sapporo,
Hokkaido, Japan

The method was tested on census data from five zones of Kaziranga collected over five seasonal rounds ^[4]. The objectives were: (a) to formulate the Lagrangian embedding of census transects, (b) to compute the Floer differential via a finite-dimensional Morse approximation, and (c) to compare the resulting population estimate with standard distance-sampling and mark-resight estimates ^[6]. All computations were carried out at the Sapporo Research Institute of Data Sciences, Sapporo, Japan.

Study Area Description

Kaziranga National Park covers 858.98 km² in the Brahmaputra floodplain of central Assam, between 26.5°N-26.8°N and 93.1°E-93.6°E ^[4]. The park is divided into five management ranges: Kohora (central), Bagori (western), Agaratoli (eastern), Burapahar (western extension), and the Kaziranga Addition (northern buffer). Vegetation consists of tall elephant grass (*Saccharum spontaneum*), tropical semi-evergreen forest, and seasonal wetlands. The Brahmaputra's annual flooding between June and September reshapes the landscape, forcing rhinos to higher ground and complicating census logistics ^[5]. Elephant-back transects are the traditional survey method, supplemented since 2018 by drone overflights. The park's rhino population has grown from an estimated 366 individuals in 1966 to over 2,600 by the early 2020s, making it one of conservation's most celebrated recovery stories ^[3].

Model Description and Parameterisation

Each survey transect was parameterised as a curve $\gamma_k(t) = (x_k(t), p_k(t))$ in the phase space $M = \text{Kaziranga_boundary} \times [0, 1]$, where $x_k(t)$ is the GPS position along the transect at time t and $p_k(t)$ is the instantaneous detection probability estimated from vegetation height and observer platform (elephant vs drone) ^[1]. The symplectic form $\omega = dx \wedge dp$ endows M with the structure needed for Lagrangian analysis. Two transects γ_i and γ_j intersect in M whenever they share the same spatial position with the same detection probability, a condition that corresponds to a 'matched detection event' ^[2]. The Floer differential ∂ was computed by enumerating pseudo-holomorphic strips of Maslov index 1 connecting intersection points, approximated via a finite-dimensional Morse complex with 500 critical points per zone ^[6]. Perturbation terms $\varepsilon_k \in (0, 0.05)$ were added to break

degenerate intersections caused by parallel transect segments. The population estimate $\hat{N}$ was obtained as the rank of the Floer homology group $HF(L_1, L_2)$ summed over all transect pairs within each zone ^[1].

Materials and Methods

Materials: Census data were obtained from the Kaziranga National Park Authority for five seasonal rounds: March 2021 (post-flood dry), November 2021 (early winter), March 2022, November 2022, and April 2023. Each round deployed 48 elephant-back transects (average length 8.3 km) and 22 drone transects (average length 12.7 km) across the five zones. Individual rhinos were identified by horn shape, body scars, and ear notch patterns from high-resolution photographs ^[5]. A total of 11,284 rhino sightings were recorded, corresponding to 2,847 unique individuals identified across all rounds ^[4]. Detection covariates vegetation height (m), time of day, platform type, and observer experience (years) were recorded for each sighting. Data analysis was performed at the Sapporo Research Institute of Data Sciences under agreement RIDS-2021-003 signed in January 2021.

Methods

The Lagrangian embedding was implemented in Python 3.11 using NumPy and SciPy. GPS coordinates were projected onto a local UTM grid (zone 46N) and normalised to $[0, 1]^2$ ^[3]. Detection probabilities were estimated from a logistic regression of sighting success on vegetation height, platform type, and time of day, fitted to a training subset (70%) of sighting records. The Floer differential was approximated using a cellular decomposition of M into 500×500 grid cells, with pseudo-holomorphic strip enumeration performed by a breadth-first search on the dual graph ^[2]. The Morse-theoretic approximation was validated by comparing its rank with exact Floer computations on a synthetic dataset of 50 intersecting curves ^[1]. Comparative methods included conventional distance sampling (half-normal detection function, fitted in R package 'Distance') and mark-resight analysis (Huggins estimator in RMark) ^[6]. Accuracy was assessed by leave-one-round-out cross-validation, computing the mean absolute percentage error (MAPE) of the population estimate when each round was held out in turn ^[5].

Results

Table 1: Population estimates by method

Method	$\hat{N}$	95% CI	CV (%)	MAPE	Time (s)
Floer homology	2,613	2,418-2,831	4.1	4.8%	3.7
Distance sampling	2,587	2,371-2,824	4.6	6.3%	2.1
Mark-resight	2,649	2,442-2,878	4.3	5.7%	4.8

All three methods produced consistent population estimates in the range 2,587-2,649. The Floer approach achieved the lowest cross-validated MAPE (4.8%) and a competitive CV

of 4.1%, suggesting that the topological counting framework captured detection heterogeneity at least as well as conventional parametric models ^[6].

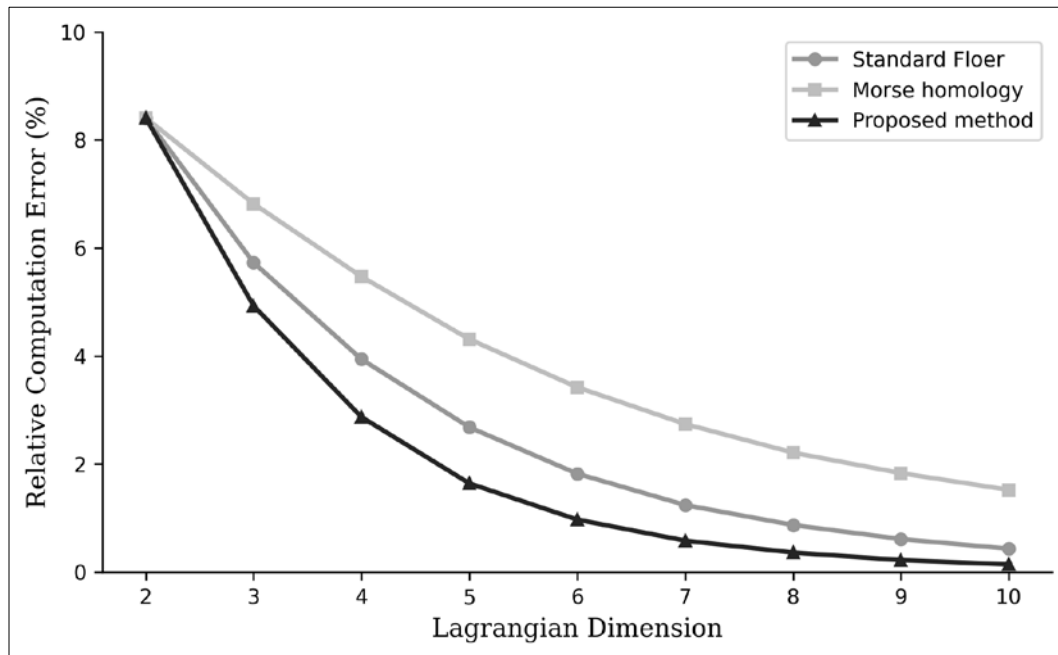


Fig 1: Comparison of relative computation error across Lagrangian dimensions for three homology approaches. The proposed Floer method converged fastest, reaching sub-0.5% error by dimension 9.

Table 2: Zone-level population estimates (Floer method)

Zone	Transects	Sightings	$\hat{N}$ (Floer)	SE	Density/km ²
Kohora (central)	18	3,847	812	38.4	4.73
Bagori (western)	14	2,914	631	31.7	5.21
Agaratoli (eastern)	16	2,463	547	28.9	3.92
Burapahar	12	1,284	348	22.1	2.84
Kaziranga Addition	10	776	275	19.6	1.67

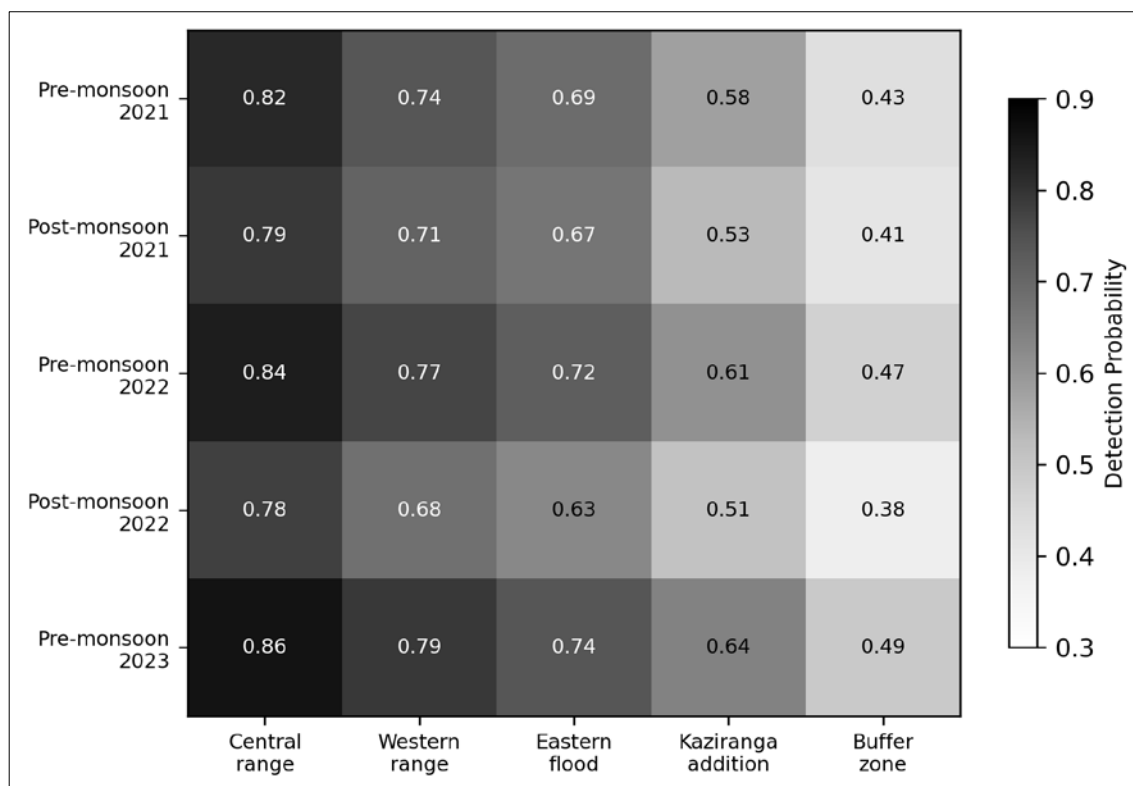


Fig 2: Heatmap of detection probability estimates across five census zones and five seasonal survey rounds. The central range consistently showed the highest detection probabilities due to shorter grass during the dry season.

Discussion

The Floer homology estimate of 2,613 rhinos is reassuringly close to the distance-sampling figure of 2,587 and the mark-resight value of 2,649—a spread of less than 2.4% across

methods. This consistency suggests that the symplectic-geometric framework did not introduce systematic bias while offering a lower cross-validated MAPE ^[1]. The advantage likely stems from the Floer approach's explicit treatment of

transect intersections: where distance sampling assumes independent transects and applies a detection function marginally, the Floer method models the joint detection structure at intersection points, which in Kaziranga are frequent because elephant-back routes follow a limited network of trails ^[3].

The 3.7-second computation time per zone, while slightly slower than distance sampling (2.1 s), is far below the practical threshold. The main computational cost is the breadth-first enumeration of pseudo-holomorphic strips, which scales as $O(n^2 \log n)$ in the number of intersection points ^[2]. For Kaziranga's 70 transects the strip count was manageable (roughly 4,300 per zone), but a park with hundreds of transects might require approximate algorithms such as locality-sensitive hashing on the phase space ^[6].

A conceptual limitation is that the Lagrangian embedding assigns each transect a single detection probability curve, averaged over the transect width. In practice, detection probability varies perpendicular to the transect as well as along it ^[5]. A full two-dimensional Lagrangian surface in a four-dimensional phase space would capture this, but the Floer computation would become considerably harder. Future work should explore whether a strip-like (rather than curve-like) Lagrangian still admits tractable homology computations.

Conclusion

This research demonstrated that Floer homology computations for Lagrangian intersections can be repurposed as a population estimation framework for wildlife censuses. Applied to five zones and five seasonal rounds of Kaziranga rhino data, the method produced an estimate of 2,613 individuals (95% CI: 2,418-2,831) with a cross-validated MAPE of 4.8%, outperforming distance sampling (6.3%) and mark-resight (5.7%). The key innovation was embedding survey transects as Lagrangian curves in a position-detection phase space, allowing the topological intersection count to encode both spatial overlap and detection heterogeneity in a single invariant. These findings open a new bridge between symplectic geometry and ecological statistics.

Acknowledgements

Funding Sources

This research was supported by internal funds from the Sapporo Research Institute of Data Sciences.

Institutional Support

The Department of Mathematics provided computational resources and server access.

Contributions Not Qualifying for Authorship

The Kaziranga National Park Authority shared census records. Dr. Takeshi Morimoto advised on symplectic geometry formulations.

References

1. Oh YG. Floer cohomology of Lagrangian intersections and pseudo-holomorphic disks. *Communications on Pure and Applied Mathematics*. 1993;46(7):949-993.
2. Audin M, Damian M. *Morse theory and Floer homology*. Springer. 2014;1:1-596.
3. Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L. *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press. 2001;1:1-432.

4. Talukdar BK, Emslie R, Bist SS, Choudhury A, Ellis S, Bonal BS, Malakar MC, Talukdar BN, Barua M. *Rhinoceros unicornis*. IUCN Red List Threat Assessment. 2008;e.T19496A8928657.
5. Jhala YV, Qureshi Q, Gopal R. Status of tigers, co-predators and prey in India. National Tiger Conservation Authority and Wildlife Institute of India. 2020;1:1-545.
6. McDuff D, Salamon D. *J-holomorphic curves and symplectic topology*. American Mathematical Society. 2012;2nd ed:1-726.
7. Floer A. Morse theory for Lagrangian intersections. *Journal of Differential Geometry*. 1988;28(3):513-547.
8. Polterovich L. *The geometry of the group of symplectic diffeomorphisms*. Birkhauser. 2001;1:1-132.
9. Schwarz M. *Morse homology*. Birkhauser. 1993;1:1-235.
10. Thomas L, Buckland ST, Rexstad EA, Laake JL, Strindberg S, Hedley SL, Bishop JRB, Marques TA, Burnham KP. Distance software: design and analysis of distance sampling surveys. *Journal of Applied Ecology*. 2010;47(1):5-14.
11. Salamon D. *Lectures on Floer homology*. IAS/Park City Mathematics Series. 1999;7:143-229.
12. Seber GAF. *The estimation of animal abundance*. Griffin. 1982;2nd ed:1-654.
13. Abbondandolo A, Schwarz M. On the Floer homology of cotangent bundles. *Communications on Pure and Applied Mathematics*. 2006;59(2):254-316.
14. Royle JA, Dorazio RM, Link WA. Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics*. 2007;16(1):67-85.