Markov Spectral Analysis of Arctic Buoys

Karun Ram, Dimitrios Giannakis

Department of Mathematics, Dartmouth College

Abstract

We construct a **Markov-operator representation** of the dynamical system of **Arctic Ocean currents**, using a collection of $N \approx 1.62 \times 10^6$ data points collected from ice-tethered buoys over ≈ 20 years [IAB23]. By **analyzing the eigenspectrum** of a finite-rank approximation to the Markov operator, **we identify (near)-invariant sets under the forward evolution of the system**, and their corresponding **basins of attraction**. With this, we decompose the Arctic Ocean into a set of **"weakly interacting" regions**, and develop a **probabilistic model** for the **evolution of surface material distribution** over time. Planning for future Arctic ice studies, oil spill mitigation, and climate research are among the many applications of our findings.

Motivation and Previous Work

The dynamical system governing surface currents in the Arctic Ocean are of immense interest across disciplines. One of the most comprehensive reviews of the dynamics of surface currents is given in [TM20], and we hope to expand upon these results by studying the evolution of the system over time, and to build a robust probabilistic model of the dynamics.

Previous studies apply probabilistic modelling techniques (e.g. Ulam's method) to obtain similar classifications of surface current dynamics at different scales.

- [FSS14]: simulated dataset of ocean currents, partitions global ocean into minimally interactive regions in order to identify formation of "garbage patches."
- [Mir+17]: collection of Lagrangian buoys in the Gulf of Mexico, identifies regions over which surface material transport is limited.

Modelling the dynamics of the Arctic Ocean present two unique challenges:

- **Data noise**: Tracers often disappear and reappear (for example, they are often trapped under other ice sheets, or are covered in snow), which makes the data inconducive to direct analysis via Ulam's method
- **Sparsity**: Due to the inaccessibility of the Arctic, there are geographical regions and time periods over which the frequency of buoys is sparse, which results in an inaccurate model.

To resolve these, we employ techniques in [GSZ15], constructing an orthonormal smooth eigenbasis of Gaussian functions (as opposed to the indicator basis that Ulam's method gives).

Method

Data processing: We first denoise the data, removing obviously erroneous readings and decoupling trajectories with large periods of inactivity. We use conservative lower and upper bounds on sea ice velocity to remove beached and "collected" buoys. We then perform linear interpolation on each trajectory.

Once filtered, we are left with a collection of trajectories $\mathcal{C} = \{C_i : i \in \{1, 2, ..., k\}\}$, where each trajectory is a sequence of 3-tuples (x, y, t) representing time-ordered coordinates in $\mathbb{R}^2$ (representing the stereographic projection of a buoy's position onto the plane). Our goal is to model the dynamical system $\Phi^q : X \to X$ (where $X \subset \mathbb{R}^2$ is our region of interest), representing the forward evolution of buoy position driven by surface currents, parametrized by the timestep q (we set $q = 90$ days).

Constructing smooth orthonormal basis: Let $H = L^2(X, \mu)$ be a Hilbert space where μ is a well-chosen measure (e.g. area). Let $k : X \times X \to \mathbb{R}$ be a map s.t.

$$k(x, y) = \exp\left(\frac{||x - y||^2}{\sigma(x)\sigma(y)\epsilon^2}\right)$$

where based on a heuristic optimization procedure, we set $\epsilon \approx 0.4812$, and $\sigma(x)$ is a measure of data density around x. k induces a natural integral linear operator $K : H \to H$, which is self-adjoint and compact. The eigenfunctions of K form a smooth orthonormal basis for H.

Finite-rank approximation: We approximate K by a finite-rank matrix K_N, by considering an (approximately) equidistributed subset $X_N \subset X$ which we sample from our data. The eigenvectors of K_N correspond to the evaluation of the eigenfunctions of K on X_N. So we construct an approximate basis

$$\beta_{N,\ell} = \{\phi_i : 1 \leq i \leq \ell\}$$

where ϕ_i is the ith eigenvector of K_N, sorted in decreasing order by eigenvalue magnitude. By a result in [GSZ15], for a well-chosen ℓ, $B := \text{span}(\beta_{N,\ell})$ approximates $L^2(X_N, \mu)$ well. Empirically, we found that $\ell = 500$ gave a sufficient approximation without capturing too much noise.

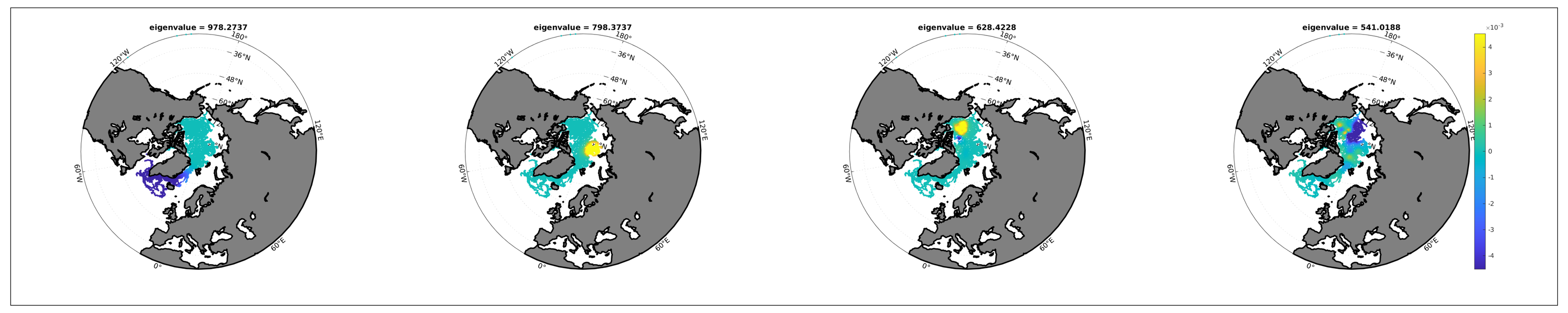

Figure 1: Selected smooth eigenbasis elements

Constructing Markov operator: We next compute an operator $P : H \to H$ such that $(P^* \circ f)(x) = f(\Phi^q(x))$. We use the finite-rank approximation again and we compute $P_N : L^2(X_N, \mu) \to L^2(X_N, \mu)$ instead. Under the suitably restricted operator $\Phi^q_N = \Phi^q \Big|_{X_n}$, we have that

$$(P^*_N \circ f)(x) = f(\Phi^q_N(x)).$$

and in particular, if $X_N = \{x_i : 1 \leq i \leq n\}$ is a trajectory separated by timesteps of 1 (in units of q),

$$(P^*_N \circ f)(x_i) = f(x_{i+q}).$$

If we further project f onto B, then we can construct a reduced Markov operator $P_{N,\ell}$ that well approximates P_N. Further, it can be shown that

$$P_{N,\ell} = \frac{1}{N}\phi^T \psi$$

where ϕ is a $N \times \ell$ matrix whose ith column is ϕ_i: that is, the ith element in $\beta_{N,\ell}$, and the ith column of ψ is the "translation" of ϕ_i by q timesteps. In particular, to model the forward evolution of a probability distribution p over X_n, we first project p onto B, giving some vector $\overline{p}$, then compute $B\overline{p}$.

Eigenvector analysis: We then analyze the eigenspectrum of $P_{N,\ell}$, which we then interpret over X_N using the methods described in [FSS14] and [Mir+17]. To summarize this analysis method, we consider level sets of eigenvectors. Left eigenvectors correspond to near-invariant sets, the corresponding right eigenvectors correspond to basins of attraction, and eigenvalues correspond to decay rates of the invariant sets.

Preliminary Results

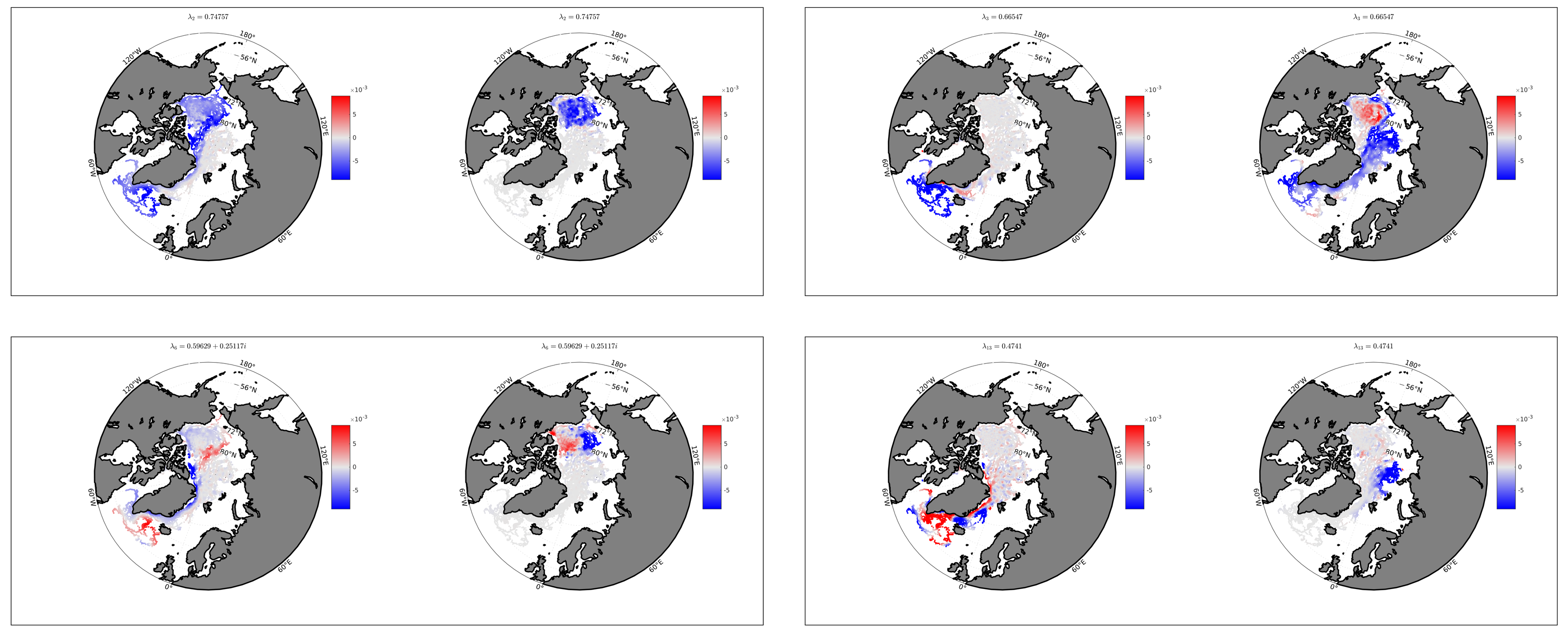

Figure 2: Selected left-right eigenvector pairs of the reduced Markov operator visualized on X_N

The result were compared to results to a baseline set of results using Ulam's method to ensure robustness of our parameter selections. Qualitatively, the interpretation of the level sets concurs broadly with existing empirical understandings of the Arctic Ocean (specifically the Beaufort Gyre and the Transpolar Drift), and with the baseline results obtained using Ulam's Method (not shown here). The lack of coverage of the dataset, as well as the large timespan over which data was accumulated, mean that we need to be wary of drawing any stronger conclusions. With that said, these results are promising, and lend credence to the methodology.

Future Directions

First, we intend on expanding our dataset considerably. We have acquired an updated dataset that is ≈ 100 times larger than the original, and upon analysis it will allow us to alleviate many of the issues arising from data sparsity. We also intend on subdividing the dataset into shorter time periods, which will likely prevent the "averaging out" of shorter-term trends First, we intend on expanding our dataset considerably. We have acquired an updated dataset that is ≈ 100 times larger than the original, and upon analysis it will allow us to alleviate many of the issues arising from data sparsity. We also intend on subdividing the dataset into shorter time periods, which will allow our assumption of the Markov property to be satisfied more strongly, as well as providing us with insights into how the underlying dynamical system has evolved over time.

References

[IAB23] IABP. *International Arctic Buoy Project*. 2023.

[TM20] Mary-Louise Timmermans and John Marshall. "Understanding Arctic Ocean Circulation: A Review of Ocean Dynamics in a Changing Climate". In: *Journal of Geophysical Research: Oceans* 125.4 (2020).

[FSS14] Gary Froyland, Robyn M. Stuart, and Erik van Sebille. "How well-connected is the surface of the global ocean?" In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 24.3 (Sept. 2014). ISSN: 1054-1500.

[Mir+17] P. Miron et al. "Lagrangian dynamical geography of the Gulf of Mexico". In: (2017). arXiv: 1703.10684 [physics.ao-ph].

[GSZ15] Dimitrios Giannakis, Joanna Slawinska, and Zhizhen Zhao. "Spatiotemporal Feature Extraction with Data-Driven Koopman Operators". In: Proceedings of Machine Learning Research 44 (2015), pp. 103–115.