

Gaussian processes at the Helm(holtz): A better way to model ocean currents

Understanding the behavior of ocean currents has the potential to improve ecosystem management, forecasting of oil spill dispersion, and the comprehension of ocean transportation. To this end, ocean researchers release GPS-tagged buoys to track ocean currents. Since we expect current dynamics to be smooth but highly non-linear, Gaussian processes (GPs) offer an attractive model. In particular, one existing approach is to consider the velocities of the buoys as sparse observations of a vector field in two spatial dimensions and one time dimension. But we show that applying a GP, e.g. with a standard square exponential kernel, directly to this data fails to capture real-life current structure, such as continuity of currents and the shape of vortices. By contrast, these physical properties are captured by divergence and curl-free components of a vector field obtained through a Helmholtz decomposition. So we propose instead to model these components with a GP directly. We show that, because this decomposition relates to the original vector field just via mixed partial derivatives, we can still perform inference given the original data with only a small constant multiple of additional computational expense. We illustrate our method on simulated and real oceans data.

Preliminary Version 1: Inferring (vector fields over) ocean currents using Gaussian Processes

Inference of time-varying vector fields from sparse observations is an important problem in oceanography and atmospheric science. For example, ocean currents are modeled as smooth 2D vector fields, the locations and velocities of drifting buoys provide sparse observations, and estimates for the underlying vector field could help predict the dispersion of oil spills.

Gaussian processes (GP) have some features that are useful for tackling this problem, for example the possibility of quantifying uncertainty, dealing with noisy observations, and including prior knowledge in the model. Nonetheless, with a naive use of GPs ([cite Lodise paper?](#)) the inferred currents do not reflect the physical properties of currents, e.g. the presence of vortices and (convergent) fronts.

In this work, we show that if we transform our smooth vector field of interest into two scalar fields (with particular properties) via the Helmholtz decomposition, these physical features can be inferred more accurately. In particular, we prove that putting two independent GP priors on the scalar fields still allows us to perform Bayesian inference on the original vector field, and the posterior samples obtained in this way seem to better capture the oceans' properties of interest.

[We also believe that this method can be extended to other relevant applications, including vector fields for atmospheric or climate modelling, a future work direction which will be further investigated.](#)

Preliminary Version 2: Bridging Gaussian processes & vector fields using Helmholtz decomposition

Vector fields play a main role in modelling important natural phenomena, such as the speed and direction of a moving fluid throughout space or the strength and verse of a magnetic force as it changes from one point to another. Exploiting well-known physical equations and simulations, climate scientists are able to retrieve the vector field representation for many problems of interest, e.g. concerning the oceans or the atmosphere. Unfortunately, simulations have some inevitable limitations, leading to trade-offs between accuracy and tractability. Machine learning could in theory provide some useful tools to overcome this challenge, but how to integrate it with physical simulations is not straightforward.

In this work, we propose a method meant to help towards this direction, based on the Helmholtz decomposition. This fundamental vector calculus theorem states that any smooth 2D vector field can be decomposed into the sum of two scalar vector fields, one curl-free and the other divergence free, related to the original field via (mixed) partial derivatives. We exploit this result modelling our data through a Gaussian Process (GP) regression where we put two independent GP priors on the scalar fields, and we show that this still leads to a GP on the original 2D vector field.

Our method is then validated on an oceanographic application of interest, i.e. reconstructing ocean currents starting from drifter data obtained through gps measurement ([Cite the project where they got real data?](#)). This problem has recently received added attention for various real-world issues pertaining this, e.g. predicting the transport and dispersion of oil spills and plastics across marine ecosystems. In our setting, we model both simulated and real world observations as points on a vector field, for which we have the position (a latitude-longitude pair) and the velocity across each of these two directions. Furthermore, we consider how this vector field evolves across time, by properly including a timewise component in the GP kernel.

Remarkably, we discover that ...