Low Cost ASV for High-Resolution Spatio-Temporal Aquatic Field Reconstruction via Dynamic Kernels

Rodney Staggers Jr *, Bharath Vedantha Desikan*, Jnaneshwar Das
Arizona State University, Tempe, AZ, USA
*contributed equally
Email: rdstagge@asu.edu, bvedant1@asu.edu, jdas5@asu.edu

***Abstract*—We present a low-cost (under US$2 500) autonomous surface vessel (ASV) capable of efficient 3D reconstruction of aquatic environmental fields. The design and autonomy stack for the vehicle is described. To reconstruct continuous scalar fields from spatially distributed sensor data, Gaussian process (GP) regression is employed, assessing kernel functions such as squared-exponential and Matérn, together with a non-stationary Matérn variant whose spatially varying length-scale captures local covariance and anisotropic structure. By optimizing kernel hyperparameters via log-marginal-likelihood maximization, the resulting predictive models produce interactive 2D and 3D water-quality visualizations through Folium. This scalar-field-reconstruction (SFR) workflow naturally integrates with the DeepGIS decision-support system, and the non-stationary formulation sharpens local estimates while concentrating uncertainty in poorly sampled regions, thereby guiding adaptive replanning and to the best of our knowledge, this is among the earliest field validations of a non-stationary Matérn kernel for lake water-quality mapping. For this initial effort in characterizing the dynamics of Tempe Town Lake in Arizona, we present reconstruction of the upper water column as a 2D field. Prediction mean and uncertainty maps are evaluated, and we close this paper with a discussion of our future directions for multi-robot mapping.**

I. Introduction

Autonomous environmental monitoring platforms can play a vital role in understanding and managing aquatic ecosystems (1). Current aquatic monitoring methods are still labor-intensive, spatially limited, and offer low spatio-temporal resolution. In contrast, in recent years robotic platforms equipped with integrated sensor suites have demonstrated collection of high-resolution spatiotemporal data across large water bodies, enabling more accurate assessments of environment as well as infrastructure health (2).

The R/V *Karin Valentine* (Figure 1) is a low size, weight, and power (SWaP) autonomous surface vessel (ASV) developed for this purpose, incorporating a robust suite of hardware optimized for autonomous *in-situ* data collection and optimal sample return for *ex-situ* analysis. Its avionics package includes a Pixhawk flight controller running v1.15 (and v1.16 *beta*) PX4 autopilot stack for low-level control, and an ODROID companion computer that runs ROS2 for high-level decision making and data handling. This system architecture supports real-time sensor integration and autonomy, while its payload—comprising a 1-D sonar for depth mapping and a multi-probe water-quality sonde—enables real-time bathymetric and biogeochemical analysis (Table I).

To transform raw spatial data into interpretable environmental insights, this work employs Gaussian Process (GP) regression ((3)) for reconstructing continuous scalar fields (4) such as temperature distributions. GP models offer a probabilistic, non-parametric approach that naturally captures spatial dependencies and provides quantification of uncertainty. By evaluating multiple kernel functions and optimizing their hyperparameters, this approach supports reliable field reconstruction for decision-support applications in aquatic research and resource management.

Novelty and Contributions.: This study stands at the forefront of aquatic remote sensing by uniting a sub-US$2 500, low-SWaP ASV with a *leading yet highly under-explored* non-stationary Matérn kernel. To the best of our knowledge, this is the one of the earliest field-validated application of that kernel to *in-situ* lake water-quality mapping, offering a very fresh outlook on hotspot detection and adaptive uncertainty quantification. Our approach sharpens scalar-field reconstructions, localises predictive uncertainty for informative re-sampling, and streams posterior statistics into an open DeepGIS/Folium visual-analytics stack, yielding an interactive 2-D/3-D workflow that closes the loop between data collection and mission planning. .

II. Related Work

Previous efforts at robotic aquatic sampling have explored both mobile swarms and single-vessel approaches. A cyber-physical system has been used to reconstruct spatiotemporal aquatic fields using a swarm of low-cost robotic fish (5). Their method emphasizes adaptive sampling through a rendezvous-based mobility scheme, where the sensors periodically regroup to maintain wireless connectivity and plan their next movement. In contrast, our approach employs a single autonomous surface vehicle following a systematic lawnmower path to collect data, with scalar field reconstruction performed offline

using kernel-based methods. This strategy simplifies platform coordination and avoids the complexity of distributed swarm control, while still enabling high-resolution coverage of the field.

The deployment of autonomous surface vessels for water sampling is not a novel notion. In recent years, groups such as have taken advantage of advancements in flight controllers to perform lithological and bathymetric measurements on aquatic environments (6; 7). On the autonomy front, tasks such as station-keeping and obstacle avoidance have been demonstrated in aquatic environments (8; 9). The main contribution of our works are a) low SWaP and low cost rapidly-prototyped ASV design optimized for multi-robot operations, and b) demonstrating a science laboratory catering to needs of local ecosystems (Arizona in our case). On this ASV, we demonstrate optimal estimation of the environmental fields measured, providing a baseline for the waterbody. To do so, we build up on prior research where such tasks have been executed in a principled manner using Gaussian process regression (10).

In order to reconstruct these continuous fields with adaptive uncertainty quantification, Paciorek and Schervish introduced a non-stationary extension of the Matérn covariance by letting the local anisotropic length-scale be a positive-definite matrix $\Sigma(\mathbf{x})$ that varies with location, thereby sharpening both mean predictions and uncertainty contours in heterogeneous domains (). However, naïve inference with this kernel requires $\mathcal{O}(n^3)$ Cholesky factorizations and $\mathcal{O}(n^2)$ storage, making it infeasible for more than a few thousand points without specialized approximations (). Moreover, treating $\Sigma(\mathbf{x})$ itself as a Gaussian process leads to challenges in identifiability and slow convergence of MCMC when fitting the model, limiting practical use to $n \lesssim 10^3$ (11). An alternative was proposed by Sampson and Guttorp (1992), who warp the original geographic coordinates into a latent "dispersion" space to recover stationarity. Although elegant, their multidimensional-scaling plus spline-interpolation pipeline involves ad-hoc tuning of stress functions, can yield non-unique mappings, and delivers only modest prediction improvements in practice (12).

III. System Description

The R/V Karin Valentine (Figures 1 and 2) was designed to incorporate real-time data collection using PX4 firmware, an open-source flight stack built on the real-time embedded operating system (RTOS), NuttX. This system can integrate the peripheral inputs of a particular airframe configuration (UART/Serial, PWM-driven motors, cameras, etc.) with its internal odometry to execute a mission plan. In essence, this infrastructure supports high-level control and autonomy.

This idea of crafting autonomous flight plans is only possible by leveraging several packages and software libraries. In particular, the MAVROS package enables communication between our onboard computer and the flight controller. With MAVLink serving as its baseline protocol, the system is capable of directing commands, transmitting messages between the flight controller, its supporting GUI (which in our case is QGroundControl), peripheral devices, telemetry apparati, and companion computers. Along with that, via the MAVLink protocol, QGroundControl serves as an essential hub to coordinate the waypoints for our boustrophedon surveys.

Fig. 1: R/V Karin Valentine docked at the Tempe Town Lake in Arizona, USA

TABLE I: Sonde Parameters

Parameter	**Units**
Temperature	°C
pH	–
Depth	m
Conductivity	μS/cm
Dissolved Oxygen Saturation	%
Dissolved Oxygen Concentration	mg/L
Chlorophyll	μg/L
CDOM	ppb
Turbidity	NTU

To complete this system, a motorized winch is fixed to the vessel to facilitate vertical profiling along the water column. When the boat has reached a point of entry, the boat will drift and begin to descend the sonde, which will be attached to the winch, by an incremental delta (1-meter) for 3-5 seconds. After that interim has expired, the winch will continue the descent of the sonde until the device is within one meter of the sonar's ascertained bathymetric value. Once the system has reached this distance, the winch will reel the housing unit up to the water surface, and then proceed to the next waypoints.(Figure 3)

For the operation of this research, this venture initially functioned off of a Holybro Pixhawk 4 (FMUv5), but now operates off of the Hex Cube Black Pixhawk Flight Controller (reliant on the FMUv3 hardware). The boat measures $130\,\text{cm} \times 88.6\,\text{cm} \times 42.3\,\text{cm}$, with a total mass of $21.3\,\text{kg}$. It is powered by two $12\,\text{V}$, $20\,\text{Ah}$ $LiFePO_4$ batteries connected in parallel, and operates at cruise speeds ranging from 0.7 to $2\,\text{m/s}$. The waypoint navigation system uses a $0.5\,\text{m}$ threshold radius for target proximity.

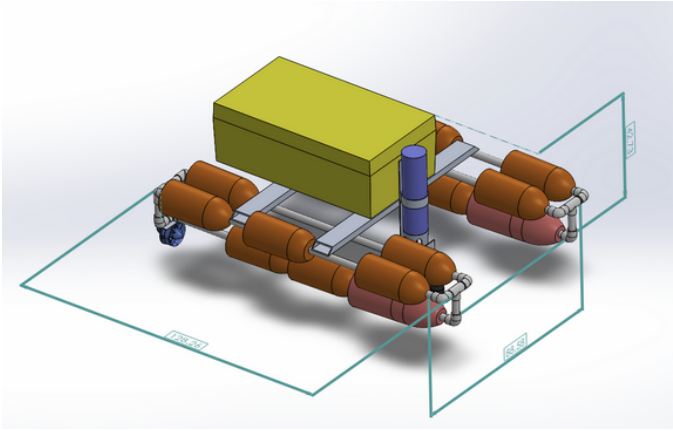


Fig. 2: CAD model of R/V Karin Valentine dimensionalized in centimeters

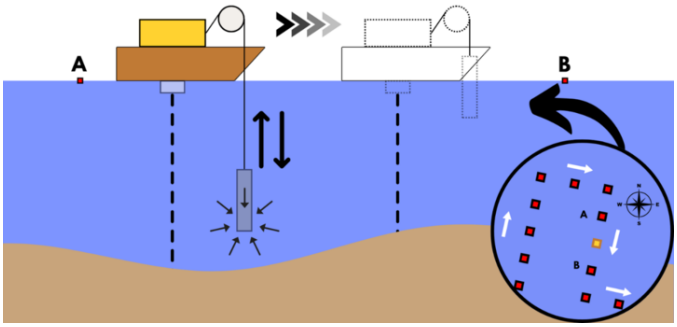


Fig. 3: GPS, bathymetry, and water quality data are collected as the R/V travels to subsequent waypoints.

A. Data Collection

The surveys were modelled to roughly resemble a boustrephedon pattern with five-meter turns in a 100-meter x 100-meter section of the lake. Given that the sonde, sonar, and GPS data were asynchronously collected, the first step in collating the dataset involved aligning the GPS and sonar data to the timing of the sonde data. Since the sonde data (which was collected at a rate of 1Hz) had the minimum number of data points, the other two parameters were truncated by matching the GPS and sonar points that were collected closest to the time at which an individual sonde measurement was taken. Subsequently, a 3x3 plot and a CSV file are constructed using sonde readings and their assigned geospatial coordinates (Figure 4). Additionally, the process is replicated to align the sonar readings with its prescribed GPS coordinates (Figure 5).

IV. SCALAR FIELD RECONSTRUCTION

A. Gaussian Process Overview

We model the continuous scalar field $f : \mathbb{R}^2 \to \mathbb{R}$ (e.g., water temperature) via Gaussian Process regression (3; 13). Given n noisy observations

$$y_i = f(\mathbf{x}_i) + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}(0, \sigma_n^2),$$

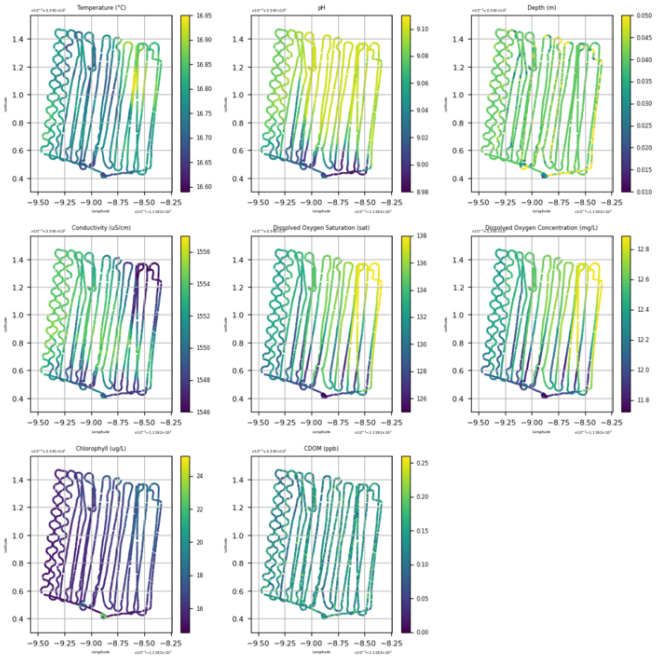


Fig. 4: Various Sonde readings plotted against GPS

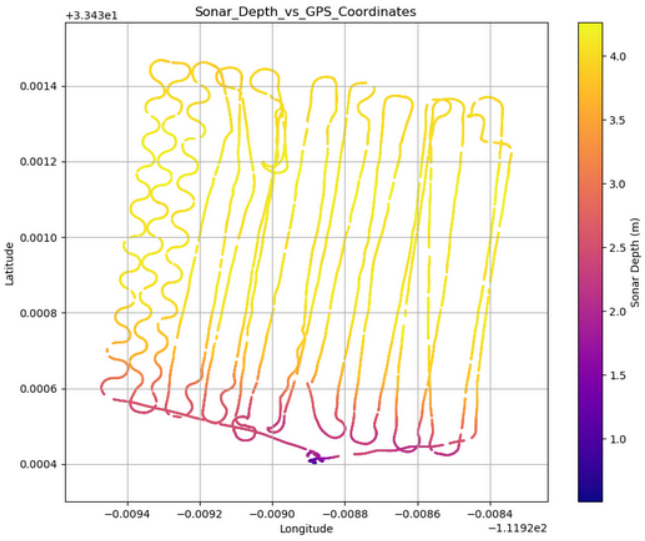


Fig. 5: Bathymetric depth plotted against GPS

at inputs $X = \{\mathbf{x}_i\}_{i=1}^n \subset \mathbb{R}^2$, we place a GP prior

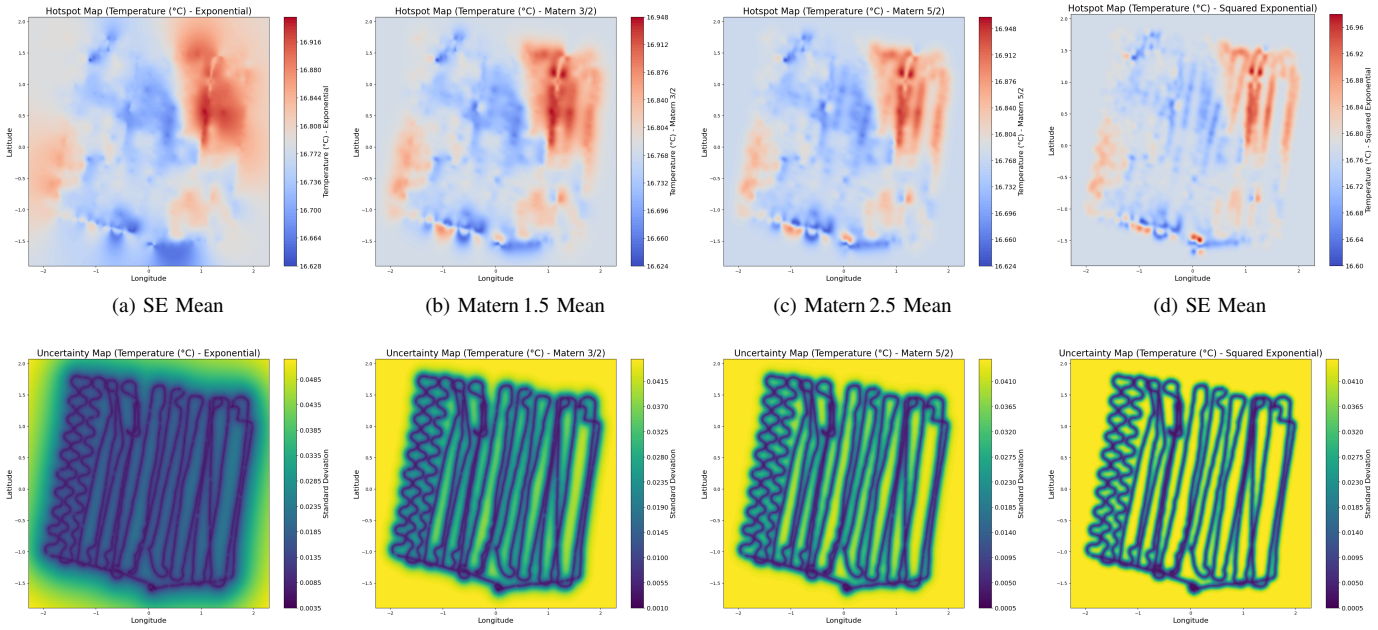
$$f(\mathbf{x}) \sim \mathcal{GP}\big(0,\, k(\mathbf{x}, \mathbf{x}')\big),$$

so that the joint distribution over $\mathbf{y}$ and predictions at test locations X_* is

$$\begin{bmatrix} \mathbf{y} \\ f_* \end{bmatrix} \sim \mathcal{N}\big(\mathbf{0},\, \begin{bmatrix} K(X,X) + \sigma_n^2 I & K(X,X_*) \\ K(X_*,X) & K(X_*,X_*) \end{bmatrix}\big).$$

B. Stationary Kernels

Under stationarity, kernels depend only on the Euclidean distance $r = \|\mathbf{x} - \mathbf{x}'\|$. We consider:



(a) SE Mean (b) Matern 1.5 Mean (c) Matern 2.5 Mean (d) SE Mean

(e) SE Uncertainty (f) Matern 1.5 Uncertainty (g) Matern 2.5 Uncertainty (h) SE Uncertainty

Fig. 6: Collected data analysis (Dec. 6): mean (top row) and uncertainty (bottom row) maps for stationary kernels.

- **Squared Exponential (SE):**

$$k_{\mathrm{SE}}(r) = \sigma_f^2 \exp\!\Big(-\tfrac{r^2}{2\ell^2}\Big),$$

- **Matérn:**

$$k_{\mathrm{Matern}}(r) = \sigma_f^2 \,\frac{2^{1-\nu}}{\Gamma(\nu)} \Big(\sqrt{2\nu}\,\tfrac{r}{\ell}\Big)^{\nu} K_\nu\Big(\sqrt{2\nu}\,\tfrac{r}{\ell}\Big).$$

C. Motivation for Non-Stationary Kernels

Standard Gaussian process models with stationary covariance kernels assume that correlation is uniform in all directions and constant over the domain. In practice, this causes them to over-smooth sharp, anisotropic transitions or under-smooth fine local structure, obscuring localized variability near hotspots and producing nearly uniform uncertainty belts around sampling tracks.

Cao and Low address this by proposing entropy- and mutual-information–based path planning algorithms that condition only on a fixed window of nearby observations, thereby exploiting directional correlation in anisotropic fields to achieve near-optimal uncertainty reduction while keeping computation time linear in the planning horizon length (14).

Likewise, the Multi-robot Adaptive Sampling Problem (MASP) framework formulates exploration as a dynamic program over Gaussian and log-Gaussian processes, and proves that increasing adaptivity—i.e., responding to newly observed, localized hotspots—monotonically reduces mapping uncertainty. This demonstrates the practical benefit of allowing spatially varying correlation lengths that can shrink in high-variance regions (15).

D. Non-Stationary Kernel Formulation

To capture spatially varying length scales and anisotropy, we employ the non-stationary Matérn kernel (11):

$$k_{\mathrm{NS}}(\mathbf{x},\mathbf{x}') = \sigma_f^2 \,\frac{|\Sigma(\mathbf{x})|^{1/4}\,|\Sigma(\mathbf{x}')|^{1/4}}{\big|\tfrac{1}{2}(\Sigma(\mathbf{x})+\Sigma(\mathbf{x}'))\big|^{1/2}}\,\frac{z^\nu}{2^{\nu-1}\Gamma(\nu)}\,K_\nu(z),$$

with $\Sigma(\mathbf{x})$ encoding local length scales and $z^2 = (\mathbf{x}-\mathbf{x}')^\top[(\Sigma(\mathbf{x})+\Sigma(\mathbf{x}'))/2]^{-1}(\mathbf{x}-\mathbf{x}')$. The prefactor normalizes and adapts covariance amplitude to local variability, sharpening uncertainty contours around temperature patches.

E. Log Marginal Likelihood & Cholesky

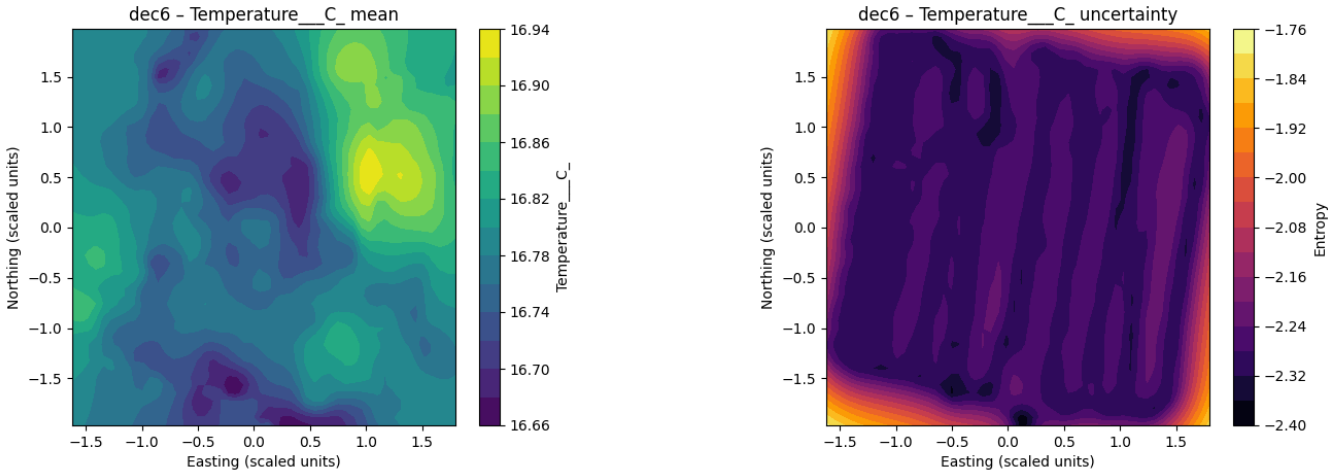
Hyperparameters $\theta = \{\sigma_f, \ell_0, \alpha, \rho, \sigma_n\}$ are learned by maximizing the log marginal likelihood (16) of $\mathbf{y}$ under covariance $K = K(X,X;\theta) + \sigma_n^2 I$. We compute $K = LL^\top$ via Cholesky so that inverses and determinants are efficiently and stably obtained.

F. Posterior Prediction

Given test inputs X_*, the predictive mean and covariance follow standard GP expressions, and prediction entropy guides adaptive sampling.

V. RESULTS AND DISCUSSION

Figure 6 compares the stationary-kernel GP reconstructions on Dec. 6 data. In the top row, the mean prediction maps recover the broad temperature gradient but completely smooth over smaller hotspots. The corresponding uncertainty maps in the bottom row remain nearly uniform around the survey track, masking areas where the field is poorly constrained.



(a) Mean

(b) Uncertainty

Fig. 7: Non-stationary GP results on Dec. 6 data: mean and uncertainty.

By contrast, Figure 7 shows results from the non-stationary GP. Its mean field not only reproduces the large-scale structure but also preserves sharper local peaks, while the uncertainty map adapts to spatial heterogeneity—contracting in well–sampled regions and highlighting variance near temperature patches.

We further validate these improvements on an analytic test field (Fig. 8). Here the non-stationary kernel resolves both the central Gaussian bump and anisotropic sinusoidal patterns more faithfully than any stationary kernel, and it concentrates uncertainty where gradients are steep.

In summary, stationary kernels smooth out fine-scale anomalies in their mean predictions and enforce a uniform uncertainty structure that can hide local variability. The non-stationary GP, by contrast, dynamically modulates both mean and variance, making it far better suited to mapping heterogeneous aquatic environments.

VI. Conclusions

We demonstrated that a low-cost autonomous surface vessel, R/V Karin Valentine, when coupled with Gaussian Process regression, can produce high-fidelity reconstructions of aquatic scalar fields. Our systematic comparison of stationary kernels (Exponential, Squared Exponential, Matérn) against a non-stationary Matérn kernel revealed that only the non-stationary formulation adapts its uncertainty and correlation structure to local anisotropy, sharpening mean estimates and concentrating variance around true hotspots. By leveraging Cholesky-based hyperparameter optimization and efficient posterior inference, our pipeline delivers numerically stable, interactive 2D/3D field maps suitable for decision-support. These results confirm that non-stationary covariance models substantially enhance environmental monitoring accuracy over traditional stationary alternatives.

VII. Future Work

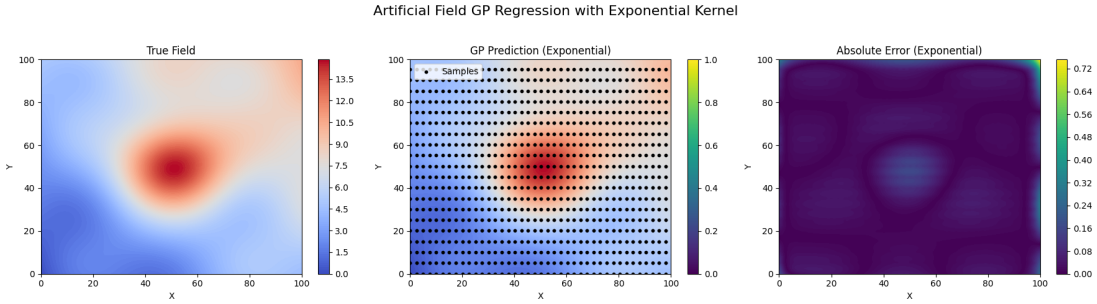
A secondary mission with an informative path can be estimated after the first lawnmower survey, by maximizing an utility function. Additionally, multi-robot mapping can be carried out with a second autonomous vessel t(under construction) to both increase the coverage area of the experiment, and reduce mapping time.

We aim to evaluate sampling efficiency by measuring information gain (e.g., reduction in posterior variance) per unit distance or energy, across both smoothly varying and patchy (e.g., algal bloom) environments, in order to empirically validate when adaptive, non-stationary–kernel sampling delivers the greatest modeling improvements for a given resource budget.

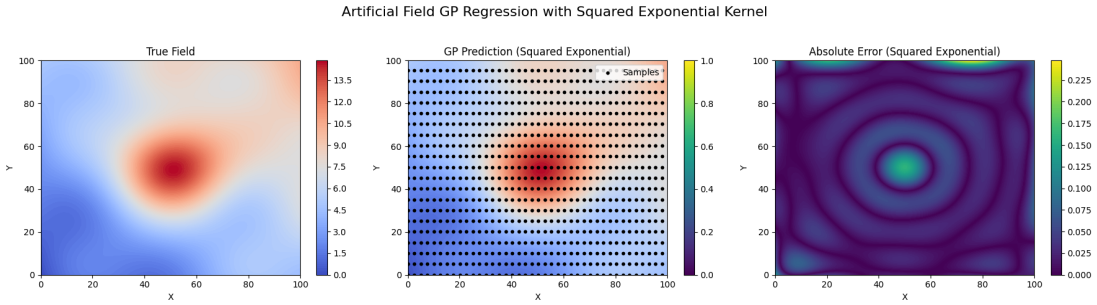
Continuing with the escalation of the experimental design, as the winched system embarks on its limnological analysis along the water column, there exists hope to assess the geospatial, bathymetric, and biogeochemical data as a function of time (thus making a 4D model of the waters' behavior). This concept necessitates a GP too. Additionally, current research is assessing the usefulness of hydrological flow models (such as CE-QUAL-W2) in predicting the characteristics of the water.

Acknowledgements

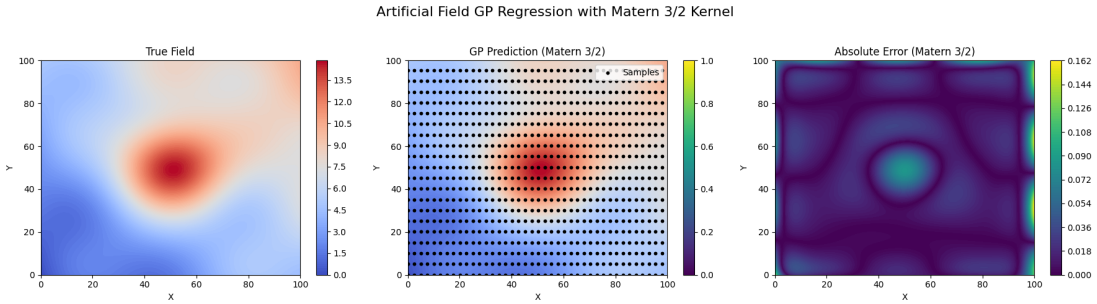
We would like to thank Swaraj Akurathi for his assistance with the field experiments.



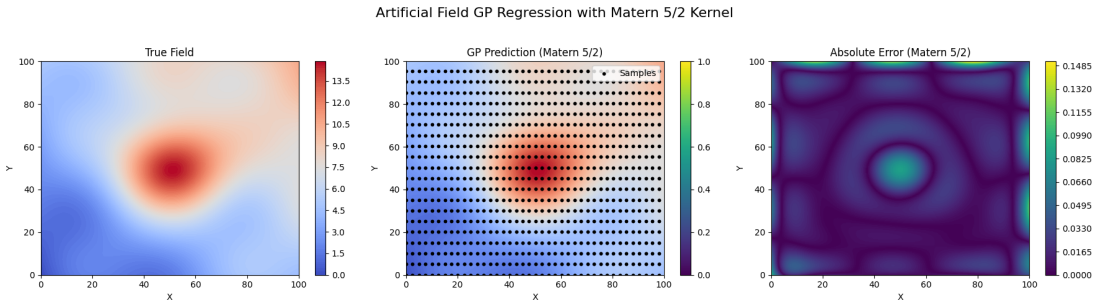
(a) Exponential



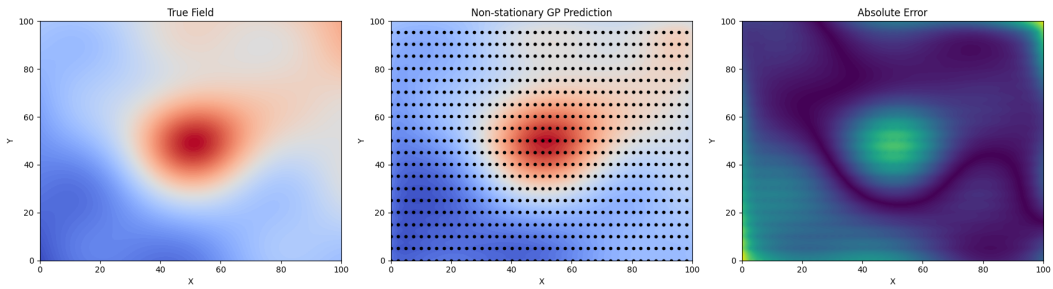
(b) Squared Exponential



(c) Matern (ν=1.5)



(d) Non-Stationary



(e) Matern (ν=2.5)

Fig. 8: Validation with synthetic scalar fields: Predictions for stationary and non-stationary kernels on an analytic test field, arranged vertically.

References

[1] B. Bayat, A. Crespi, and A. Ijspeert, "Envirobot: A bio-inspired environmental monitoring platform," in *2016 IEEE/OES Autonomous Underwater Vehicles (AUV)*, pp. 381–386, 2016.

[2] K. C. Cavanaugh, T. W. Bell, K. E. Aerni, J. E. Byrnes, S. McCammon, and M. M. Smith, "New technologies for monitoring coastal ecosystem dynamics," *Annual Review of Marine Science*, vol. 17, no. Volume 17, 2025, pp. 409–433, 2025.

[3] Williams, Christopher and Rasmussen, Carl, "Gaussian processes for regression," in *Advances in Neural Information Processing Systems* (D. Touretzky, M. Mozer, and M. Hasselmo, eds.), vol. 8, MIT Press, 1995.

[4] P. B. Jónsson, J. Wang, and J. Kim, "Scalar field reconstruction based on the Gaussian process and adaptive sampling," in *2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, pp. 442–445, 2017.

[5] Y. Wang, R. Tan, G. Xing, X. Tan, J. Wang, and R. Zhou, "Spatiotemporal aquatic field reconstruction using robotic sensor swarm," in *2012 IEEE 33rd Real-Time Systems Symposium*, pp. 205–214, IEEE, 2012.

[6] L. M. Díaz, S. Y. Luis, A. M. Barrionuevo, D. S. Diop, M. Perales, A. Casado, S. Toral, and D. Gutiérrez, "Towards an Autonomous Surface Vehicle Prototype for Artificial Intelligence Applications of Water Quality Monitoring," 2024.

[7] S. Manjanna, A. Q. Li, R. N. Smith, I. Rekleitis, and G. Dudek, "Heterogeneous multi-robot system for exploration and strategic water sampling," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4873–4880, 2018.

[8] A. Pereira, J. Das, and G. S. Sukhatme, "An experimental study of station keeping on an underactuated asv," in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3164–3171, 2008.

[9] F. Sotelo-Torres, L. V. Alvarez, and R. C. Roberts, "An unmanned surface vehicle (USV): Development of an autonomous boat with a sensor integration system for bathymetric surveys," *Sensors*, vol. 23, no. 9, p. 4420, 2023.

[10] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research*, vol. 9, no. 8, pp. 235–284, 2008.

[11] C. Paciorek and M. Schervish, "Nonstationary Covariance Functions for Gaussian Process Regression," in *Advances in Neural Information Processing Systems* (S. Thrun, L. Saul, and B. Schölkopf, eds.), vol. 16, MIT Press, 2003.

[12] P. D. Sampson and P. Guttorp, "Nonparametric Estimation of Nonstationary Spatial Covariance Structure," 1992.

[13] T. Beckers, "An Introduction to Gaussian Process Models," 2021.

[14] N. Cao, K. H. Low, and J. M. Dolan, "Multi-Robot Informative Path Planning for Active Sensing of Environmental Phenomena: A Tale of Two Algorithms," 2013.

[15] K. H. Low, J. Dolan, and P. Khosla, "Adaptive multirobot wide-area exploration and mapping," vol. 1, pp. 23–30, 01 2008.

[16] A. Schirru, S. Pampuri, G. Nicolao, and S. Mcloone, "Efficient Marginal Likelihood Computation for Gaussian Process Regression," 10 2011.