Wave-Based Terrain Navigation: A Feasibility Study for GNSS-Free Aerial Localization

James Theal Stewart
Dept. of Systems Design Engineering
University of Waterloo
Waterloo, Canada
jtstewart@uwaterloo.ca

***Abstract*—Terrain-referenced navigation (TRN) localizes a vehicle by matching depth observations against a bathymetric chart, typically using active sonar on underwater platforms. We evaluate whether passive, wave-derived depth estimates from a drone's downward camera can support the same task over the nearshore without GNSS. Using real bathymetric surveys and video-derived depth products from three coastal sites (Duck, NC; Scheveningen, NL; Porthtowan, UK), we simulate particle-filter TRN across five observation scenarios with 30 Monte Carlo seeds each, separating the effects of observation bias, random noise, and chart error.**

Systematic bias, not noise, is the dominant obstacle: a $+0.52$ m near-constant depth bias (drawn from UAV wave-inversion results [7]) degrades localization by 2.5–$5.9\times$, but a bias-augmented particle filter that jointly estimates position and bias recovers to median 5.9 m at Duck and 6.9 m at Porthtowan. Depth from a second pipeline (COCOS, based on dynamic mode decomposition) is harder to correct because its errors are depth-dependent and spatially correlated rather than constant. This error-structure distinction reflects the algorithm and integration window rather than the platform, since both pipelines are video-derived and one COCOS site is itself UAV video. Site morphology governs feasibility: distinctive bar-trough structure enables sub-7 m localization, while featureless slopes resist correction regardless of observation quality. Under the assumed motion model, these results are of similar order of magnitude to reported land-based aerial TRN, despite using a passive optical sensor and a single scalar depth observation per step.

***Index Terms*—terrain-referenced navigation, particle filter, GNSS-denied, UAV, wave-derived bathymetry, marine inspection**

I. Introduction

Aerial drones operating over the coast need reliable localization, but GNSS can be unavailable in contested environments, near multipath-inducing structures, or under emissions constraints. Terrain-referenced navigation (TRN) offers an alternative by matching depth observations against a prior bathymetric chart—an approach proven for submarines and cruise missiles [1], [3] but unexplored for aerial platforms operating over water. Over land, recent visual TRN methods achieve sub-11 m accuracy by matching LiDAR or camera imagery against elevation models [10], [11]. Over water, these methods fail because the surface is transient and featureless for visual matching.

The key insight enabling aerial TRN over water is that coastal depth can be inferred passively from video. A drone's downward camera captures propagating waves, and wave-based bathymetry algorithms invert observed wave celerity to estimate depth. Recent methods achieve sub-meter accuracy from UAV video [7] and 1–3 m from shore-based or UAV wave video processed through the COCOS pipeline [6]. This raises a direct question: can a drone estimate the depth beneath it from its own camera and use that to fix its position on a chart?

We study this through simulation using real data. Ground-truth bathymetric surveys provide reference charts, and observation errors are drawn from published UAV depth accuracy [7] or from video-derived depth products produced by the open-source COCOS pipeline [6]. Drone trajectories and particle-filter localization are synthetic; the underlying terrain and error characteristics are empirical. We evaluate five scenarios across three morphologically distinct coastal sites to separate the effects of observation bias, noise, and chart error.

The contributions are: (1) a feasibility analysis of aerial TRN using wave-derived bathymetry; (2) identification of systematic bias as the dominant failure mode when the observation pipeline produces a near-constant offset, recoverable via joint position-bias estimation; and (3) evidence that morphology and observation-error structure govern feasibility more than nominal depth RMSE.

II. Background

TRN originated with TERCOM for cruise missile guidance [1]; modern Bayesian formulations use particle filters to handle the multimodal posteriors arising in nonlinear terrain maps [2], [3]. Underwater TRN with active sonar is mature [4], [12], [13]. Over land, aerial platforms achieve < 11 m RMSE by matching LiDAR heightmaps [10] and < 7 m MAE via terrain-weighted image matching [11]. All such methods rely on persistent visual or elevation features visible from above. Over water the surface lacks them: waves, foam, and glint change on sub-second timescales. We apply the same particle-filter framework using a different terrain modality: the stable seafloor, observed indirectly through wave propagation.

Ocean surface gravity waves obey the dispersion relation $\omega^2 = gk\tanh(kh)$; in the nearshore ($kh < 0.3$) this simplifies to $h \approx c^2/g$, so depth depends only on phase celerity and not on wave period. Algorithms extract c from video via spectral analysis (cBathy [5], [14]), temporal cross-correlation [7],

or dynamic mode decomposition (DMD), as implemented in the COCOS pipeline [6]; satellite platforms extend the same physics to regional scales [16]. Reported accuracies span 0.38–2.6 m RMSE across cBathy v2 at Duck [14], UAV-hover video [7], and COCOS at 32 s–5 min windows [6], degrading above $H_s = 1.2$ m [15]. Unlike sonar noise, these errors are biased, spatially correlated, and depth-dependent—properties that, as we show, matter more for TRN than the headline RMSE.

III. Methodology

A. Observation Model

A drone over a bathymetric chart $h(\mathbf{x})$ obtains depth observations:

$$z_t = h(\mathbf{x}_t) + b + v_t, \quad v_t \sim \mathcal{N}(0, \sigma^2_{\text{obs}}) \tag{1}$$

We evaluate five scenarios (Table I) contrasting two different wave-inversion pipelines. S1 is a sonar-like baseline with no bias and moderate noise. S2/S3 parameterize a Gaussian observation model from Klotz et al.'s [7] temporal-correlation UAV results: near-constant bias $b = +0.52$ m and $\sigma = 0.59$ m, decomposed from their cross-shore profile analysis (overall RMSE $= 0.74$ m). S3 additionally estimates b online. S4 ingests Dynamic Mode Decomposition outputs directly from the COCOS pipeline [6] without added noise, preserving their depth-dependent and spatially correlated error structure. S5 inverts the chart-vs-observation question: the reference chart is the COCOS product while the observation is sonar-like, isolating the effect of chart error. Both observation sources are video-derived. Klotz is an offshore hovering UAV, and the public COCOS dataset is Argus-tower video at Duck and Porthtowan and hovering UAV video at Scheveningen. The S2-vs-S4 contrast reflects algorithm (temporal correlation vs. DMD) and integration window (14-min hover vs. 32 s–5 min) rather than platform. For S4 we set the filter's σ_{obs} to the residual RMSE of the COCOS map against ground-truth (GT). This is a natural default when no characterization of the spatial error structure is available. We then examine its sensitivity in §IV-C. All data are from the public COCOS dataset[1].

TABLE I
Experimental scenarios

#	Name	Chart	Observations
1	GT + Sonar	GT	$h + \mathcal{N}(0, 0.5^2)$
2	GT + Klotz	GT	$h + 0.52 + \mathcal{N}(0, 0.59^2)$
3	GT + Klotz + Bias	GT	Same; PF estimates b
4	GT + COCOS	GT	COCOS depth (no added noise)
5	COCOS + Sonar	COCOS	$h + \mathcal{N}(0, 0.5^2)$

[1]https://github.com/MatthijsGawehn/COCOS

B. Particle Filter

Position is estimated with a standard SIR particle filter [8] ($N = 1000$, systematic resampling [19] at ESS $< N/2$ with Gaussian roughening $\sigma_{\text{rough}} = 5.0$ m). Particles are propagated with the dead-reckoned displacement $\Delta\mathbf{x}_t$ plus Gaussian noise $\mathcal{N}(\mathbf{0}, \sigma^2_{\text{motion}}\mathbf{I})$. Over-water visual odometry is not characterized in the literature (published accuracy is from feature-rich land [9]), so we treat σ_{motion} as a parameter to be swept rather than a specific quality; 3.0 m/step is the operating point, with sensitivity across $[0.2, 5.0]$ m/step reported in §IV-B.

C. Bias-Augmented Filter

For Scenario 3, the state extends to $[x, y, b]$. Bias evolves as a random walk ($\sigma_{\text{diff}} = 0.03$ m/step) from a uniform prior $[-2, 2]$ m. The likelihood becomes:

$$w_t^{(i)} \propto w_{t-1}^{(i)} \exp\left[-\frac{(z_t - h(\mathbf{s}_t^{(i)}) - b_t^{(i)})^2}{2\sigma^2_{\text{obs}}}\right] \tag{2}$$

A single observation cannot separate position from bias; a *sequence* constrains both because depth-profile shape is invariant to constant additive bias.

D. Study Sites and Evaluation

Three coastal sites span a useful range of bathymetric distinctiveness following the Wright and Short [17] morphodynamic classification (Fig. 1). All data (ground-truth surveys, wave video, and COCOS depth products) are from the publicly available COCOS dataset[2]; we did not conduct field campaigns.

Duck, NC (US Army Corps FRF) features a prominent nearshore sandbar creating a non-monotonic cross-shore depth profile [18]. The bar's depth signature is nearly unique at each cross-shore position but approximately uniform along-shore. **Scheveningen, NL** presents the hardest case: a featureless dissipative slope with weak undulations where many positions produce similar depth. **Porthtowan, UK** has complex 3D bar morphology with strong local gradients but along-shore ambiguity.

To ensure fair comparison, all trajectories use shared spatial bounds defined by the intersection of GT and COCOS coverage (internal NaN cells within those bounds are filled with the site-mean depth so the PF interpolator returns uninformative but well-defined likelihoods there). Cross-shore path lengths are 300 m (Duck), 500 m (Scheveningen), and 330 m (Porthtowan). Four trajectory types evaluate terrain observability: cross-shore (30 steps, 6 along-shore positions), along-shore (30 steps), zigzag (48 steps, 4-leg lawnmower), and diagonal (30 steps). Each scenario-trajectory pair runs 30 times with independent random seeds varying PF initialization, resampling draws, and measurement noise; the COCOS maps used in S4/S5 are fixed (regenerated deterministically; see §IV-C). Results are reported as medians with interquartile range (IQR). MAE is computed over the converged second half of each trajectory.

[2]https://github.com/MatthijsGawehn/COCOS

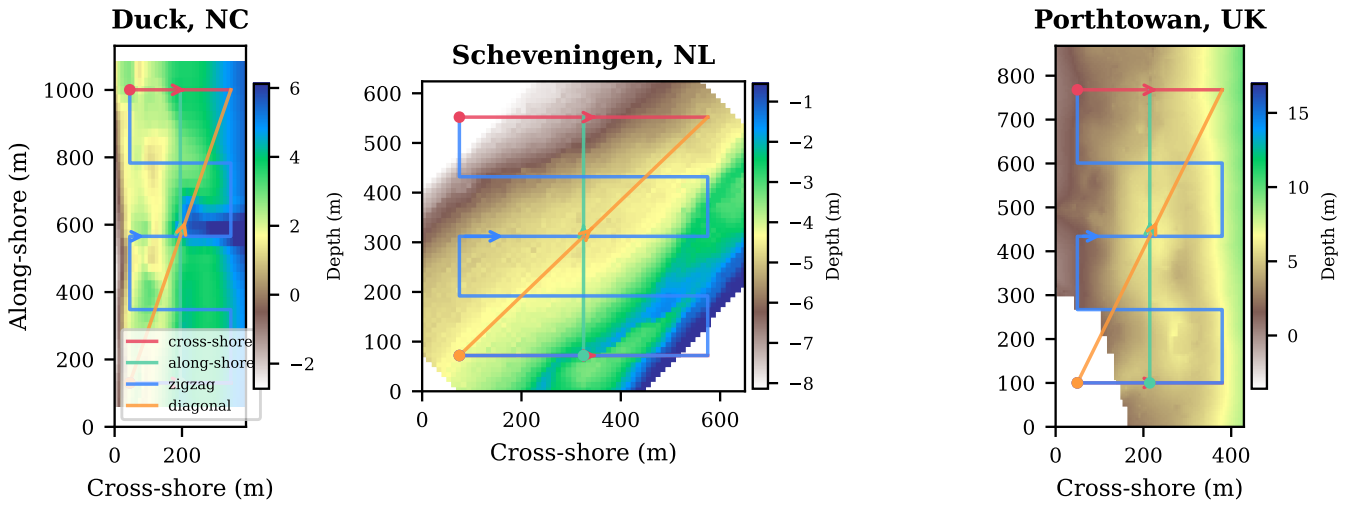

Fig. 1. Study sites with trajectories overlaid. Duck (left): bar-trough structure. Scheveningen (center): featureless slope. Porthtowan (right): complex 3D morphology.

IV. Results

A. Baseline and Wave-Inversion Bias

Under ideal conditions (S1), zigzag trajectories achieve median cross-shore MAE of 4.4 m at Duck (IQR 3–6), 16.9 m at Scheveningen (IQR 14–20), and 4.8 m at Porthtowan (IQR 4–6). Zigzag consistently outperforms single-axis trajectories because it samples depth gradients in both cross-shore and along-shore directions. Cross-shore trajectories achieve comparable cross-shore MAE but poor along-shore localization (2–12× worse), confirming that along-shore localization is the fundamental bottleneck across these three sites.

Introducing the Klotz wave-inversion bias ($b = +0.52$ m, $\sigma = 0.59$ m, S2) degrades zigzag performance by 2.5–5.9× (Table II). The bias shifts the PF's depth-matching point: at Duck, the estimated sandbar position moves; at Porthtowan, false matches arise among irregular bars.

TABLE II
Median zigzag MAE_x / MAE_y (m) across scenarios, 30 seeds. COCOS maps are regenerated deterministically at seed 42 (§IV-C).

	S1	S2	S3	S4	S5
Duck	4/9	26/43	**6**/24	7/241	6/259
Schev.	17/18	42/47	**22**/24	151/67	25/75
Porth.	5/8	15/4	**7**/14	21/151	64/189

B. Bias Recovery

The bias-augmented PF (S3) recovers to within 1.3–1.4× of the sonar baseline at all sites using the same 1000 particles (Fig. 2). At Duck, median zigzag MAE_x drops from 26.1 m to 5.9 m (IQR 3.8–10.7). At Scheveningen, from 42.4 m to 21.6 m (IQR 12.7–41.5). At Porthtowan, from 14.9 m to 6.9 m (IQR 4.8–9.1). The wider IQR at Scheveningen reflects the

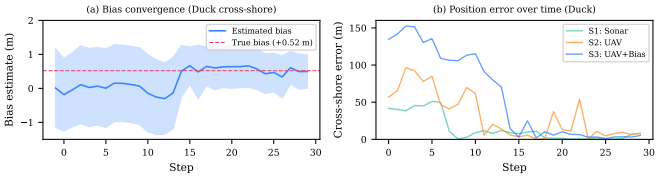

Fig. 2. (a) Bias estimate convergence at Duck. (b) Position error: S1–S3 comparison showing bias estimation (S3) recovers accuracy.

difficulty of jointly estimating position and bias on a featureless slope where many (x, y, b) triples locally explain the same observations. The result is not sensitive to the motion-model assumption: across $\sigma_{\text{motion}} \in \{0.2, 0.5, 1.0, 3.0, 5.0\}$ m/step, the S3/S1 zigzag MAE_x ratio stays within $[1.1\times, 1.7\times]$ at all three sites.

Recovery works because the depth-profile shape along a trajectory is invariant to a constant additive bias. The PF identifies position from the shape and attributes the residual offset to the bias state. The bias estimate typically converges within 10–20 steps, stabilizing near $\hat{b} \approx +0.5$–0.6 m, close to the true value of $+0.52$ m. Multi-directional trajectories are important: zigzag paths sample depth gradients in orthogonal directions, breaking the position-bias degeneracy more effectively than single-axis traverses. On cross-shore-only trajectories, bias recovery is less reliable because the along-shore position remains unconstrained.

C. COCOS Observations and Chart Error

COCOS observations (S4) produce worse localization than the Klotz-profile model (S2) at Scheveningen (3.6×) and Porthtowan (1.4×), despite comparable effective RMSE at Duck (0.79 m COCOS vs. $\sqrt{b^2 + \sigma^2} = 0.79$ m for the Klotz model). The difference lies in error structure: Klotz's approximately constant bias can be absorbed by the bias-

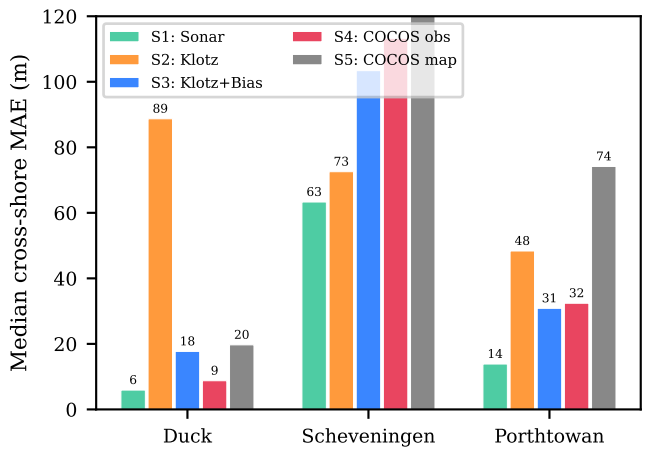

Fig. 3. Median cross-shore MAE on cross-shore trajectories, across scenarios and sites.

augmented PF, while COCOS's DMD-pipeline residuals vary with depth and position (cf. Gawehn et al.'s own error characterization [6]), so no single bias term can correct them. The same correlated error structure appears at all three COCOS sites even though Scheveningen's COCOS input is itself UAV video, which confirms that the S2-vs-S4 contrast reflects the algorithm and integration window rather than the UAV-vs-tower platform distinction. At Duck, the COCOS error is small and spatially smooth, and the cross-shore axis recovers to $\mathrm{MAE}_x = 7\,\mathrm{m}$ on zigzag—near the S1 baseline (4 m).

Comparing S4 (COCOS observations, GT chart) with S5 (COCOS chart, sonar observations) reveals an asymmetry: a biased chart is more survivable than biased observations when the chart preserves the depth profile's shape. At Duck, S5 achieves $\mathrm{MAE}_x = 6\,\mathrm{m}$ (on zigzag at the S1 baseline) because the COCOS chart distortion is smooth. At Porthtowan, S5 is much worse ($\mathrm{MAE}_x = 64\,\mathrm{m}$ on zigzag; best-trajectory cross-shore is $42\,\mathrm{m}$) because $\sim$50% of the COCOS chart is NaN, degrading the reference map beyond recovery. Fig. 3 bars median cross-shore MAE across all five scenarios at each site, showing where the bias-augmented filter (S3) comes close to but does not fully recover the sonar baseline (S1).

S4 performance is partly sensitive to the filter's σ_{obs}: widening it above the fitted residual RMSE on Scheveningen reduces MAE_x by $\sim 35\%$ because a broader likelihood tolerates spatially correlated residuals, but the effect is modest or adverse at the other sites. Matching the likelihood to the spatial correlation structure of the observation error, rather than tuning a scalar, is the principled next step.

V. Discussion

Comparison to visual TRN. Over land, visual TRN achieves $< 11\,\mathrm{m}$ by matching LiDAR or camera imagery against elevation models [10], [11]; over water these methods have no persistent features to match. Wave-derived bathymetry supplies the missing terrain layer: a stable seafloor signature accessed through wave physics. The 4.4–6.9 m median accuracy on favorable morphology is of comparable magnitude to land-based results, despite using a passive single-point depth observation rather than a dense LiDAR scan.

Spatially extended observations. This paper models depth as a single scalar per step, and this is the most conservative assumption. A drone camera physically observes a 2D wave field, and depth-inversion algorithms produce local depth maps. A preliminary 3×3 patch at 15 m spacing drops zigzag MAE at Scheveningen from 16.9 m to 4.6 m under independent-noise assumptions; however, wave-derived estimates at 15 m spacing share most of their analysis window ($\sim$60–80 m radius), so the independent-noise assumption overstates the gain. The true benefit lies between single-point and independent-patch, and a likelihood that respects the spatial correlation of the observation field is future work.

What remains unshown. Three prerequisites separate this result from a deployable system. First, over-water visual odometry is not characterized in the literature; published VIO numbers are from land and the comparison of our accuracy to land-based aerial TRN is contingent on equivalent dead-reckoning being achievable over water. Second, all observations here are from stationary/hovering video; wave inversion under platform motion has, to our knowledge, no published velocity-vs-accuracy characterization at drone speeds (1–20 m/s), and is a prerequisite for the hover-and-fly deployment envisioned. Third, field validation, such as flying a known trajectory and processing wave video in real time, would confirm that the simulated error models are representative of an actual flight.

VI. Conclusion

A GNSS-denied drone can, in simulation, localize over the nearshore by matching wave-derived depth against a bathymetric chart, provided the terrain is distinctive and observation bias is modeled. Over 30 Monte Carlo seeds, a bias-augmented particle filter recovers to within 1.3–$1.4\times$ of the sonar baseline: 5.9 m at Duck and 6.9 m at Porthtowan on zigzag trajectories, with Scheveningen's featureless slope remaining the hard case. The core lesson is that, for this modality, the structure of the observation error matters more than its nominal magnitude: the approximately constant offset that characterizes the Klotz temporal-correlation profile is recoverable by jointly estimating position and bias, while the depth-dependent and spatially correlated residuals produced by the COCOS DMD pipeline are not. This is the case even at comparable headline RMSE, and even when both inversions run on UAV video. The cleanest remaining gap between these simulation results and deployment is not the particle filter, but the production of navigation-usable depth estimates from a translating camera in real time; over-water VIO characterization and spatially extended observations are the natural next steps.

References

[1] J. P. Golden, "Terrain contour matching (TERCOM): A cruise missile guidance aid," in *Proc. SPIE*, vol. 238, 1980, pp. 10–18.

[2] F. Gustafsson *et al.*, "Particle filters for positioning, navigation, and tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 425–437, 2002.

[3] K. B. Anonsen and O. K. Hallingstad, "Terrain aided underwater navigation using point mass and particle filters," in *Proc. IEEE/ION PLANS*, 2006, pp. 1027–1035.

[4] J. Melo and A. Matos, "Survey on advances on terrain based navigation for autonomous underwater vehicles," *Ocean Eng.*, vol. 139, pp. 68–87, 2017.

[5] R. Holman, N. Plant, and T. Holland, "cBathy: A robust algorithm for estimating nearshore bathymetry," *J. Geophys. Res. Oceans*, vol. 118, no. 5, pp. 2595–2609, 2013.

[6] M. Gawehn, S. de Vries, and S. Aarninkhof, "A self-adaptive method for mapping coastal bathymetry on-the-fly from wave field video," *Remote Sens.*, vol. 13, no. 23, p. 4742, 2021.

[7] A. N. Klotz *et al.*, "Deriving nearshore bathymetry and waves characteristics from a single UAV video," *Coastal Eng.*, vol. 202, p. 104820, 2025.

[8] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proc. F*, vol. 140, no. 2, pp. 107–113, 1993.

[9] C. Forster *et al.*, "SVO: Semidirect visual odometry for monocular and multicamera systems," *IEEE Trans. Robot.*, vol. 33, no. 2, pp. 249–265, 2017.

[10] M. Werner *et al.*, "Kilometer-scale GNSS-denied UAV navigation via heightmap gradients," *arXiv:2510.01348*, 2025.

[11] F. Yao *et al.*, "GNSS-denied geolocalization of UAVs using terrain-weighted constraint optimization," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 135, p. 104277, 2024.

[12] L. Paull, S. Saeedi, M. Seto, and H. Li, "AUV navigation and localization: A review," *IEEE J. Oceanic Eng.*, vol. 39, no. 1, pp. 131–149, 2014.

[13] G. Salavasidis *et al.*, "Terrain-aided navigation for long-endurance and deep-rated autonomous underwater vehicles," *J. Field Robot.*, vol. 36, no. 2, pp. 447–474, 2019.

[14] R. Holman and E. W. J. Bergsma, "Updates to and performance of the cBathy algorithm for estimating nearshore bathymetry from remote sensing imagery," *Remote Sens.*, vol. 13, no. 19, p. 3996, 2021.

[15] K. L. Brodie *et al.*, "Evaluation of video-based linear depth inversion performance and applications using altimeters and hydrographic surveys in a wide range of environmental conditions," *Coastal Eng.*, vol. 136, pp. 147–160, 2018.

[16] R. Almar *et al.*, "Coastal topo-bathymetry from a single-pass satellite video," *Remote Sens.*, vol. 14, no. 7, p. 1529, 2022.

[17] L. D. Wright and A. D. Short, "Morphodynamic variability of surf zones and beaches: A synthesis," *Marine Geol.*, vol. 56, pp. 93–118, 1984.

[18] T. C. Lippmann and R. A. Holman, "The spatial and temporal variability of sand bar morphology," *J. Geophys. Res.*, vol. 95, no. C7, pp. 11575–11590, 1990.

[19] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.