Robots going round the bend—a comparative study of estimators for anticipating river meanders

Kai Qin[1] and Dylan A. Shell[2]

***Abstract*— Marine robots and unmanned surface vehicles will increasingly be deployed in rivers and riverine environments. The structure produced by flowing waters may be exploited for purposes of estimation, planning, and control. This paper adopts a widely acknowledged model for the geometry of watercourse channels, namely sine-generated curves, as a basis for estimators that predict the shape of the yet unseen portion of the river. Predictions of this sort help a robot anticipate the future, for example, in throttling speeds as it rounds a bend. After examining how to reparameterize standard filters to incorporate this model, we compare the performance of three Gaussian filters and show that nonideality and theoretical challenges (of non-linearity, multi-modality/periodicity) degrade the performance of standard Kalman filters severely, but can be successfully mitigated by imposing an interval constraint. Thereafter, we present results of a constrained interval Kalman filter on data from three natural rivers. The results we report show the effectiveness of our method on the estimation of meander parameters. The results we report, including data from simulation, from maps, and from GPS tracks of a boat on the Colorado river, show the effectiveness of our method on the estimation of meander parameters.**

I. Introduction

Autonomous surface vehicles are increasingly being used on rivers. Riverine environments present many interesting and important opportunities because they are arteries carrying fresh water—a precious and all-too-scarce resource; because they are sites of ecological diversity—including unique fauna and flora; and because they are corridors of commerce—several rivers are of especial historical and cultural significance. But streams and rivers pose significant difficulties too. As Synder et al. [7] observe "prior map information on water hazards and obstacles is not dependable and does not have the accuracy needed for precision navigation and sensor directed reconnaissance." Since occlusion increases uncertainty, meanders, in particular, present challenges because they hamper long-distance observation. Inevitably, watercraft navigating a river for the first time have only limited understanding of stretches of water lying ahead.

Fortunately, meandering rivers possess considerable regularity. Meanders are well-characterized by sine-generated curves, where the angular direction at any point is a sine function of the distance measured along the channel [3]. This work explores how this model can serve as a foundation for estimators that fuse observations to make predictions of the shape of unseen portions of the river (see Fig. 1). Essentially, the model provides a parsimonious state space over which our filters operate and this representational economy translates into efficiency. We envision that the estimate could be valuable for making control decisions, e.g., in selecting reference trajectories or paths for a controller to track, or, more simply, in slowing the vehicle as it approaches a bend. Predictions of river geometry can provide rich information especially when combined with other domain knowledge. Much that is useful for navigation can be gleaned, as illustrated by the following passage:

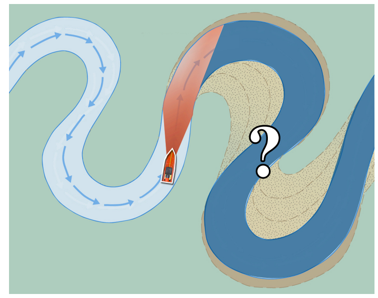

Fig. 1. This paper formulates and compares estimators that model river meanders to predict the shape of the channel ahead of the robot. Such estimates could be helpful in picking trajectories for the controller to track.

> "River boatmen navigating upstream on a large river face the problem that the deepest water, which they usually prefer, tends to coincide with the streamline of highest velocity. Their solution is to follow the *thalweg* (the deepest part of the river, from the German for 'valley way') where it crosses over the center line of the channel as the channel changes its direction of curvature but to cut as close to the convex bank as possible in order to avoid the highest velocity near the concave bank." [3]

As readers may anticipate, the periodic nature of the model presents some challenges for a standard extended Kalman filter. Indeed this *is* the case and, as will be clearly shown, these issues affect performance in practice. First we adopt a parameterization of meander geometry (Section III) and then to examine how classical filters behave using the model. This analysis, which we report in Section IV-A, involved comparison of the performance of three Gaussian filters in tracking and predicting the river's centerline. It led us to recognize the importance of constraining the estimate to a single period in order to guarantee unimodality. Thereafter, we investigated the performance of the superior, constrained filter on data from three natural rivers (Sections IV-B–V). But next we discuss relationships to the existing literature.

[1]Kai Qin (kai.qin@nxp.com) is with NXP Semiconductors.
[2]Dylan A. Shell (dshell@cse.tamu.edu) is with the Department of Computer Science and Engineering, Texas A&M University.

II. RELATED WORK

A. Models of river meanders

Langbein and Leopold [3] first proposed the sine-generated curve model of river meanders, formalized as:

$$\theta(s) = A \sin\left(\frac{2\pi s}{M}\right), \tag{1}$$

where $\theta(s)$ is the angle in radians between the direction of flow and the mean down-valley direction. The latter direction is a reference axis, pointing downstream and oriented along the centerline of the meandering pattern; as the name indicates, it reflects the broad slope of the land. Note that $\theta(s)$ is a function of the maximum angle A in radians, spatial period per meander M in meters and s, which is the distance along river from the apex of a left-hand bend in meters. This form was proposed because it minimizes the sum of squares of change in direction and also total work in bending.

Thakur and Scheidegger [9] examined the statistical distribution of angles of deviation with the mean down-valley direction a few years later. They confirmed that the angles of deviation in rivers are normally distributed. Their evidence provides support for the sine-generated curve model and also hints toward its aptness as a representation for estimation.

Much more recently, Mecklenburg and Jayakaran [4], to sidestep the highly nonlinear sine-generated curve, proposed a new arc-and-line meander pattern that represents the meander pattern with connecting arcs and lines; such a model, while perhaps easier to fit to geological data, does not provide an obvious state space description—unfortunately making it rather more complicated for our purposes.

B. Estimators from a constrained vs. unconstrained optimization perspective

A vast panoply of Bayesian filters have been proposed for parameter estimation. The present paper, being the first we are aware of to study the meander problem, begins by applying standard estimation techniques. The Kalman filter, a parametric recursive estimator for systems with Gaussian uncertainty, seems like a good choice, especially given that a Gaussian distribution is reported in [9]. However, as the meander problem is not linear, we must turn to various progeny of the classical Kalman filter.

For nonlinear systems, the extended Kalman filter (EKF) and the unscented Kalman filter (UKF) are the most commonly used Gaussian filters. At each iteration, both filters are two-step estimators that include a prediction step and a measurement update step. Compared to the EKF, the UKF has comparable computational cost and has been shown to yield more accurate results, at least to the second order of the Taylor series [1]. The measurement update steps of both filters, though rarely described this way, may be regarded as the solution of an unconstrained optimization problem [12].

Additional information about the system can be integrated into the design of filters by introducing constraints, restricting parts of the state space. Reviews of the state constraint extensions to Kalman filters can be found in [6] and [8].

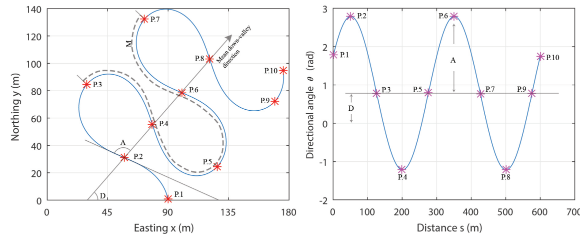

Fig. 2. A visualization of parameters A, D and M in (2).

Constraint Kalman filters can be classified into linear and nonlinear types according to the system transition function and measurement function. There are equality and inequality constraint filters based on the types of constraints.

The unscented recursive nonlinear dynamic data reconciliation (URNDDR) filter [11] improves the way that the EKF obtains the updated state estimates by solving (numerically) a constrained optimization problem and then updating the state covariance by selecting sigma points and weights much like the UKF. But, additionally, the URNDDR solves a constrained optimization problem to ensure that any state inequality constraints are satisfied in the sigma point updates. Taken together, both the updated state estimate and error covariance will satisfy state space constraint. Unfortunately this comes at a cost: URNDDR involves substantially more computation than the UKF and is sensitive to the performance of the constrained optimization problem solver. For a four-dimensional state space, at each iteration, the solver is run nine times; moreover, estimation terminates if the solver fails to find a solution.

A practical solution, and one which we adopt, is proffered by the constraint interval unscented Kalman filter (CIUKF) of Teixeira et al. [8]. It solves the constrained optimization problem only for the sampled mean of the state distribution (cf. (11) below). But the CIUKF opts to weaken the requirement for variances, using the standard UKF method to obtain the state covariance. (Due to limited space, we withhold details, but the description appears in full in [2].)

III. FORMULATION AND APPROACH

Estimating meander parameters using the model in (1) requires the robot measure the mean down-valley direction and start at the apex of a left-hand bend. These are unreasonable requirements for an autonomous vehicle. Instead, by including offset and scaling parameters, we proposed a new function for the river's centerline:

$$\theta(s) = A \sin(Bs + C) + D, \tag{2}$$

where s is the distance along the river from the robot's initial location; $\theta(s)$ is the angle between the direction of flow and magnetic East in radians; A is the maximum angle in radians; B is the spatial frequency in radians per unit length; C is the phase shift in radians; D is the angle between the mean down-valley direction and magnetic East. In addition to the spatial frequency B, the spatial period M is computed as:

$$M = \frac{2\pi}{B}. \tag{3}$$

Fig. 2 illustrates the parameters A, D and M.

A. Problem Formulation

Consider a river meander centerline that is well-characterized by the following sine-generated curve in a Cartesian coordinate system:

$$y_1(s) = \int_0^s \cos(\theta(\tau))\, d\tau + y_1(0), \quad (4)$$

$$y_2(s) = \int_0^s \sin(\theta(\tau))\, d\tau + y_2(0), \quad (5)$$

$$\theta(s) = A \sin(Bs + C) + D. \quad (6)$$

Collecting A, B, C and D into a single parameter vector to be estimated, we define the state describing the river as:

$$x = \begin{bmatrix} A & B & C & D \end{bmatrix}^\mathsf{T}. \quad (7)$$

We assume that the robot is equipped with sensors that can measure the coordinates $(y_1(k), y_2(k))$ of locations along the river's centerline, a distance s_k along river from an initial point $(y_1(0), y_2(0))$, and the angle $\theta_k \in [-\pi, \pi)$ between the direction of flow and magnetic East, where the subscript k denotes the k^{th} measurement. To simplify this problem, the measurements of s_k (distances along the river) are treated as perfect and, therefore, $s_k \leq s_{k+1}$. We are concerned with θ at s_k, but of which only an imperfect observation, denoted $\bar{\theta}_k$, can be made; we assume that its error, v_k, is normally distributed with zero mean and variance R_k. Since we are assuming measurements of the watercourse centerline, nothing need be assumed about the river width.

We desire an estimate of the parameters x_k at a point $(y_1(k), y_2(k))$ given spatially discrete sensor readings. With these definitions in place, at each state k the river meander estimation problem is formulated as follows.

PROBLEM 1: **River Meander Estimation**
Input: Prior belief of state $\mathcal{N}(\hat{x}_{k-1},\ P_{k-1})$
Input: An observation $\mathcal{N}(\bar{\theta}_k,\ R_k)$
Input: Distance s_k along the centerline from initial point
Output: Posterior belief of state $\mathcal{N}(\hat{x}_k,\ P_k)$

B. Filter Design

The sine-generated curve model exhibits nonlinearity in three of the four parameters $[A\ B\ C\ D]^\mathsf{T}$, none of which are directly observable. Nevertheless, we wish to enable the robot to estimate these parameters in real-time. The nonlinearity precludes a standard Kalman filter, so we began by implementing EKF and UKF solutions.

The prior studies by geologists treat the sinusoidal parameters as fixed constants over the region of the river under study. We expect that over long distances these parameters may drift but, as we have no *a priori* transition model for any of the four variables, we assume constant parameters for each river. Of course, if other information is known it can be incorporated too; we have

$$x_k = T x_{k-1} + \mu_k, \quad (8)$$

where the transition matrix T is the 4×4 identity matrix I. To account for gradual drift in the values, it is prudent to add system process noise $\mu_k \sim \mathcal{N}(0, Q_k)$ to the state transition equation, where Q_k is the process-noise covariance matrix. (We have Q_k as a diagonal matrix, because error of each parameter is assumed to be independent.)

For the measurement update step of the filter, we choose the meander direction angle $\bar{\theta}_k$ as the (sole) observed variable because it has been confirmed to be normally distributed [9]. The obvious alternative, using the Cartesian coordinates of points on the meander centerline, does not have this statistical property and, moreover, the measurement model for the coordinates (see (4)–(5)) is complicated. In contrast our measurement function is given by:

$$\bar{\theta}_k = h(x_k, s_k) = A \sin(Bs_k + C) + D + v_k, \quad (9)$$

where $v_k \sim \mathcal{N}(0,\ R_k)$ is the measurement error. The Jacobian matrix H_k for the observation model is given as:

$$\begin{aligned} H_k = &\left. \frac{\partial \theta(s)}{\partial x_k} \right|_{x_k = \hat{x}_{k|k-1}, s = s_k} \\ &= [\sin(Bs_k + C) \qquad AB\cos(Bs_k + C) \\ &\quad A\cos(Bs_k + C) \quad 1]. \end{aligned} \quad (10)$$

Since Kalman filters, along with various extensions thereto, represent belief over their state space with a multivariate normal distribution, they can do poorly (even breaking down) when the distribution is not unimodal [10]. Regrettably, the state space distribution of meander parameters is not unimodal. The periodic nature of sine functions poses a problem: even exact observations could confirm an infinite number of values. For example, parameters $[A\ B\ C\ D]^\mathsf{T}$ and $[A\ B\ (C{+}2n\pi)\ D]^\mathsf{T}$ represent the same sine-generated curve for any integer number n. The probabilistic analogue, thus, has multiple modes. These infinite modes make it impossible to approximate the state distribution as a single Gaussian or a Gaussian Mixture [5]. We approach this problem by using the CIUKF and constraining parameters to lie inside a single period, thereby ensuring unimodality. (A proof that, given the constraints, we have a unimodal distribution appears as an Appendix in [2].)

We implement a CIUKF based on the algorithm given in [8] and the system dynamics in (8) and (9). The *posteriori* state estimate $\hat{x}_k$ is computed by solving the constrained optimization problem described in the following equation numerically:

$$\begin{aligned} \hat{x}_k = \arg\min_{\{x_k\}} \Big[&(\bar{\theta}_k - h(x_k, s_k))^\mathsf{T} (R_k)^{-1} (\bar{\theta}_k - h(x_k, s_k)) + \\ &(x_k - \hat{x}_{k|k-1})^\mathsf{T} (P^{xx}_{k|k-1})^{-1} (x_k - \hat{x}_{k|k-1}) \Big] \\ \text{subject to:} \quad & x_L < x_k \leq x_U, \end{aligned} \quad (11)$$

where the observation model is given in (9).

IV. Results

We report measures of estimator performance in three separate evaluations. We first consider simulations (Section IV-A) where the ground truth is both known and is a true sine-generated curve. Since the true parameters are known, we can determine the error of the estimate exactly. Even in these simplified circumstances, the EKF and UKF leave much to be

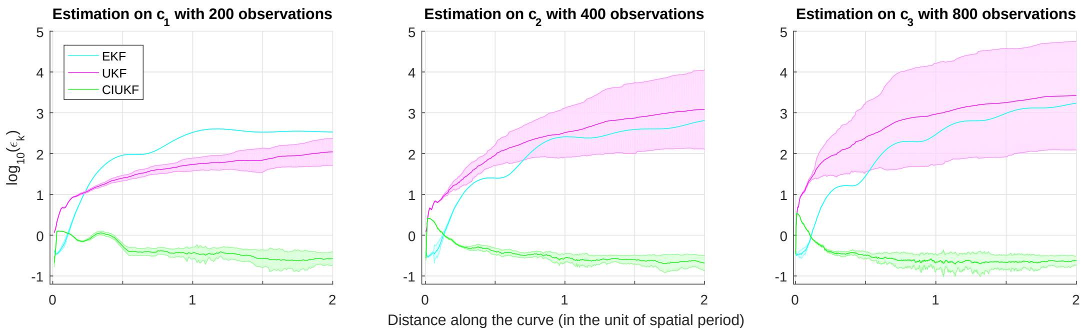

Fig. 3. State weighted error of constraint interval unscented Kalman filter *vs.* Number of measurements. The EKF and UKF produce estimates that are divergent. Plots show estimates of mean and variance computed from 30 independent simulations for each filter. Measurements are made every 5 m, though the three curves are of different scales. The horizontal axis is in units of the ground truth curve's spatial period.

desired. The second and third evaluations involve estimation on river meanders using map-based data (Section IV-B) and GPS positions collected from a boat (Section IV-C). Though these meanders are only approximately characterized by the sine-generated curve model, the data show that the CIUKF is able to provide useful predictions for the robot nevertheless.

A. Evaluation on sine-generated curves

For real rivers, one determines the scale of a meander from its spatial period M. The longer the spatial period, the larger the scale of the meander. Leopold and Langbein [3] presented the data collected from two meanders of the Mississippi River near Greenville, Mississippi (USA) and the Blackrock Creek in Wyoming (USA) to illustrate their model. Comparing the meander of the Mississippi River has spatial period of about 20 miles, while Blackrock Creek is of much smaller scale with a period of 700 feet. Here, we choose three curves with different different scales to examine the filter performance. The sine-generated curves c_1, c_2, c_3, have parameters:

$$x_{c_1} = \begin{bmatrix} 1 & \frac{2\pi}{500} & \frac{2\pi}{3} & \frac{\pi}{4} \end{bmatrix}^{\mathsf{T}}, \quad (12)$$

$$x_{c_2} = \begin{bmatrix} 1 & \frac{2\pi}{1000} & \frac{2\pi}{3} & \frac{\pi}{4} \end{bmatrix}^{\mathsf{T}}, \quad (13)$$

$$x_{c_3} = \begin{bmatrix} 1 & \frac{2\pi}{2000} & \frac{2\pi}{3} & \frac{\pi}{4} \end{bmatrix}^{\mathsf{T}}. \quad (14)$$

The initial state and its covariance matrix are given as

$$\beta_0 = \begin{bmatrix} \pi/2 \\ \frac{2\pi}{1500} \\ \pi/2 \\ \pi/2 \end{bmatrix}, \qquad P_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.01 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (15)$$

The measurement covariance matrix R, the process covariance matrix Q, and the sampling distance Δs, i.e., the distance between two sequential measurements, were

$$R = \frac{\pi}{6}, \qquad \Delta s = 5\,\mathrm{m}, \quad \text{and} \quad Q = 0.$$

The CIUKF lower and upper limits, x_L and x_U, were

$$x_L = [0\ 0\ 0\ 0]^{\mathsf{T}} \quad \text{and} \quad x_U = [2.2\ 0.1\ 2\pi\ 2\pi]^{\mathsf{T}}. \quad (16)$$

The upper limit for parameter A was selected to be 2.2 rad because the sine-generated curve model generates meanders

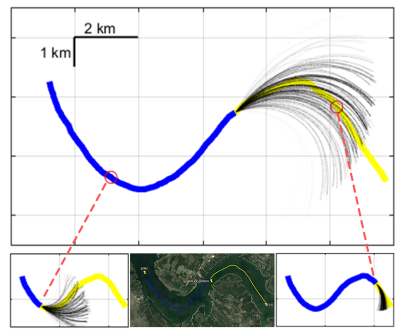

Fig. 4. After taking measurements (shown in blue) part of the way along the Brazos River, 100 samples are drawn randomly from the CIUKF estimator's current state distribution. Main image takes samples up to 8995 m from the designated start, lower-left up to 3975 m, and lower-right takes 12 957 m. These are plotted forward from this point to show predictions for the still unseen portion of the river. The curve in yellow is the river's actual centerline that the robot has not observed yet. Transparency corresponds to normalized probability of the sample.

with closed loops for values of A above approximately 2.2 rad, cf. [4]. And the upper limit for B is set to 0.1 m/rad, since we ignore the meanders with spatial period less than 62.83 m. Error corrupted observation is introduced to the filter by adding zero mean Gaussian error with standard deviation of $\pi/18$ rad to the true measurements.

For all results reported in this paper, the filters were initialized as described in this scenario, except that for real data, no additional Gaussian noise was added to the measurements, nor is the sampling distance treated as fixed.

A standard measure of performance in literature on the Kalman filters and its extensions [6], [8] is the root-mean-square error (RMSE) of each state. We found it necessary to adopt a different indicator because the spatial frequency parameter, B, has a much smaller order than the other parameters, especially for large-scale meanders, yet it affects the

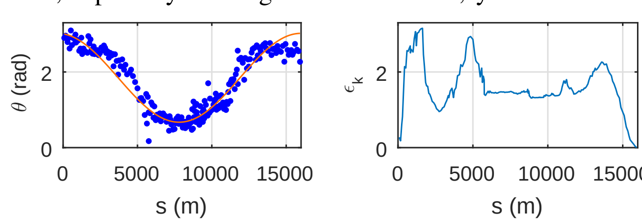

Fig. 5. Estimation quality of CIUKF on the Brazos River data. Left: (a) Blue dots are data points and the solid orange line shows the estimator's fit after 225 measurements. Right: (b) State weighted error calculated as data arrive (error estimated with all measurements used as ground truth).

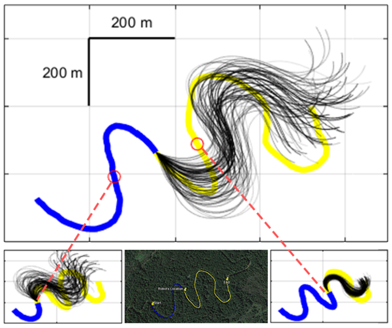

Fig. 6. An analogous plot to that of Fig. 4 but for the L'Anguille River. It is a substantially smaller watercourse. In the main image the samples are drawn after 647 m, for the lower-left after 369 m, and for the lower-right after 991 m.

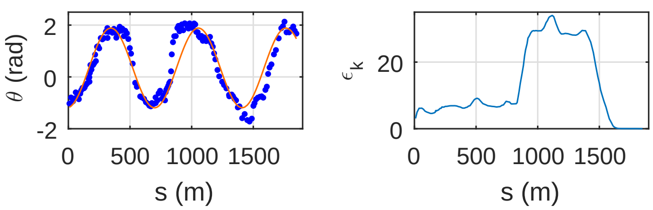

Fig. 7. An analogous plot to that Fig. 5 but for the L'Anguille. Left: (a) Blue dots are data points and the solid orange line shows the estimator's fit after 159 measurements. Right: (b) State weighted error calculated as data arrive (error estimated with all measurements used as ground truth).

quality of predictions acutely. Instead, we use the weighted error of the estimated parameters, defined as

$$\epsilon_k = (\hat{x}_k - x_{c_i})^{\mathsf{T}} (P_k)^{-1} (\hat{x}_k - x_{c_i}), \tag{17}$$

where x_{c_i} denotes the ground truth for the sine-generated curves. The similarity between (11) and (17) is worth clarifying: the measurement update steps of all three filters can be regarded as the solution of an optimization problem, for which the objective function is simply weighted least squares [12]. The first term in (11) is concerned with the immediate sensor reading; we discard its contribution and the second term forms the basis of the state weighted error metric, as it uses the same matrix, P_k^{-1}, for the weights.

Representative results using the three filters on the curves appear in Fig. 3, which shows mean and variances summarizing 30 independent simulations per filter. The vertical axis is the logarithm (base 10) of weighted error of estimated parameters at each state. The horizontal axis shows the observation data set for each curve, for twice the spatial period of each curve, with measurements with an inter-sample distance of 5 m. The convergence of CIUKF's stands in clear contrast to the degradation of the estimates of both the EKF and UKF.

B. *Evaluation on map-based data*

The verisimilitude of the sine-generated model will always be imperfect for real rivers. In a sense, the preceding evaluation gives best possible conditions for an estimator and, thus, gives ample justification to dispense with the EKF and UKF further.

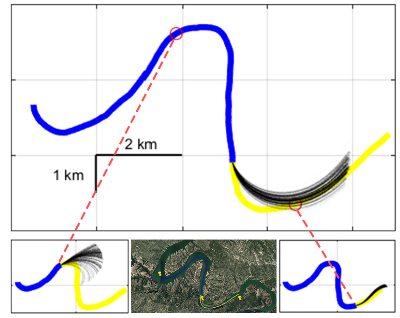

Fig. 8. Plots analogous to those in Figs. 4 & 6 for data collected from GPS in a boat on the Colorado River. In the main image the samples are drawn after 5020 m, for the lower-left after 9651 m, and for the lower-right after 12 030 m.

In order to provide realistic river data as input to the CIUKF, we manually labeled rivers on maps within Google Earth. Longitudes and latitudes of points on two centerlines of the Brazos River near Lake Whitney, Texas (USA) and the L'Anguille River near Caldwell, Arkansas (USA) were collected and processed to provide measurements for the filter. Satellite photographs of these two meanders are shown in lower-center insets in Fig. 4 and Fig. 6.

As before, we purposefully used input data at different scales. The total distances traversed in the datasets for the rivers are 15 929 m and 1849 m, respectively. There are 225 labeled points with average sampling distance of 70 m for the meander centerline of the Brazos River, and 159 labeled points with average sampling distance of 11 m for the meander centerline of the L'Anguille River. Fig. 5(a) and Fig. 7(a) show the measured and the CIUKF estimated directional angles of both meanders, where the angles are computed using (2) and red line is plotted using the estimated parameters after the final measurement. In order to track the performance of the filter over the distance along the flow, we have constructed the weighted error at each state, see Fig. 5(b) and Fig. 7(b), using (17), where here parameters estimated after all measurements is taken to be the ground truth in lieu of any alternative.

A more direct and perhaps a more meaningful visualization of the estimate is to produce sine-generated curves using sampled parameters from the estimator, and convert those into Universal Transverse Mercator (UTM) coordinates. Fig. 4 and Fig. 6 show predictions of the river forward of where the robot has traveled. Fig. 4 shows the visualization of estimates at 3975 m (lower-left inset), 8995 m (main figure) and 12 957 m (lower-right inset). Fig. 6 shows the visualization of estimates at 369 m (lower-left inset), 647 m (main figure) and 991 m (lower-right inset). The transparency of predicted meanders indicates the normalized probability densities of the samples.

C. *Evaluation using data collected on a boat*

To further get a sense of how capable the estimator would be for use on a robot, we used data collected *in situ* from a boat navigating an extended stretch of river. We hired a ski-boat (along with an experienced pilot) and collected GPS

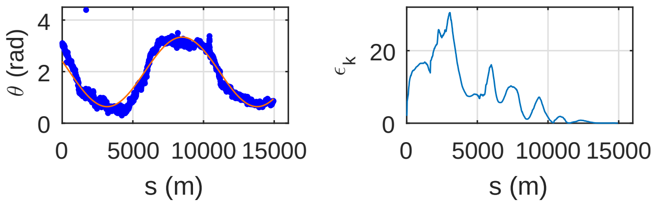

Fig. 9. Plots analogous to those in Figs. 5 & 7 for GPS data from the boat on the Colorado River. Left: (a) Blue dots are data points and the solid orange line shows the estimator's fit after 1370 measurements. Right: (b) State weighted error calculated as data arrive (error estimated with all measurements used as ground truth).

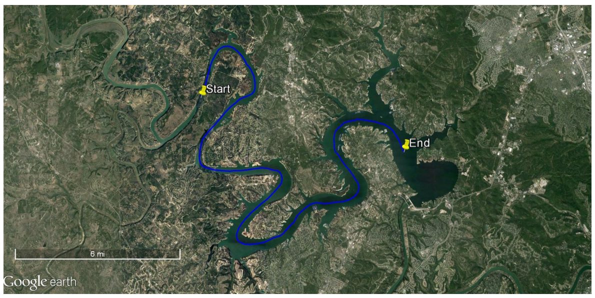

Fig. 10. The trajectory the boat travelled on the Colorado River, TX (USA).

positions of the boat trajectory on the Colorado River starting from noon, 4th Sept. 2016. To get an overall sense of the entire dataset, Fig. 10 shows the trajectory tracked on our homeward leg, returning from the furthest point reached back to the dock, with a distance of $40\,593\,\mathrm{m}$. The entire trip was twice this one-way length, for which we made 8366 measurements with an average sampling distance of $10\,\mathrm{m}$. The measurement frequency for the GPS sensor was $10\,\mathrm{Hz}$. The average speed for the boat was $10\,\mathrm{m/s}$, though the boat is capable of a maximum speed roughly double that.

Turning first to data collected on the outward journey, we applied CIUKF using parameters identical to before on the first stretch of $14\,900\,\mathrm{m}$ for which there are 1370 measurements with average sampling distance of $10\,\mathrm{m}$. Fig. 8 provides a visualization of estimates at $5020\,\mathrm{m}$ (lower-left inset), $9651\,\mathrm{m}$ (main figure) and $12\,030\,\mathrm{m}$ (lower-right inset) on the outward journey. The filter fits a sinusoidal curve to the measurements in a manner comparable to previous data. Fig. 8 shows additional detail on convergence for this part of the journey. However, in the next the section, we apply the filter over a longer stretch of the Colorado River, with a far less satisfactory outcome.

We are also interested in situations where the boat does not to follow the centerline exactly owing to uncertainty. To simulate this situation we corrupted the GPS positions with noise through the addition of random displacements. Vectors v_d, in polar coordinates (r_{v_d}, θ_{v_d}), are randomly drawn from distributions $r_{v_d} \sim \mathcal{N}(20, 25)$ and $\theta_{v_d} \sim \mathcal{U}(0, 2\pi)$. The mean value of the displacement magnitude, $20\,\mathrm{m}$, is inappreciable compared to the length and width of the meander. Nevertheless, Fig. 11 (a) shows that the added displacement vectors introduce non-negligible errors into the estimate of direction. As shown in Fig. 11 (b), it also takes more steps for the CIUKF to converge. In our experiments, the boat's

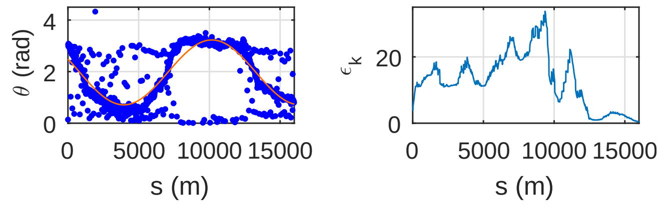

Fig. 11. Plots analogous to those in Figs. 5, 7 & 9 for GPS data from the boat on the Colorado River, but with noise added. Left: (a) Blue dots are data points and the solid orange line shows the estimator's fit after 1370 measurements. Right: (b) State weighted error calculated as data arrive (error estimated with all measurements used as ground truth).

skipper may have deviated from the true river centerline, but one would expect that the error introduced by an experienced pilot would be less than the noise introduced synthetically in the preceding evaluation.

V. Further Analysis & Outlook

A. *Assessing sensitivity to initial conditions*

We examined the sensitivity of the estimates to the choice of initial covariance matrices (i.e., in (15)). Details are in [2], but to summarize: neither the predictions made by the EKF, nor the CIUKF, show sensitivity to the initial conditions.

B. *Quantifying confidence in the predictions*

The curves in Figs. 4, 6, and 8 show that predictions generally improve with additional measurements, but also that there are parts of river bends that deviate from the sine-generated curve model. In judging curves by sight alone one can easily be mislead. What is desired is a measure of precision with an obvious interpretation: we wish to have a general sense of how useful the resulting predictions will be for subsequent planning. If the estimator's output is to be used by a robot to select its actions, for example, in bounding throttling speeds as the boat rounds a bend, a metric that relates to some notion of risk as a function of distance ahead of the boat seems prudent.

We introduce a measure, that we call the *prediction confidence*, which serves to quantify that proportion of predictions (weighted by the probability) falls inside the river. The idea is that one can tolerate some imprecision, but misclassifying riverbank for water is crucial mistake. This confidence is

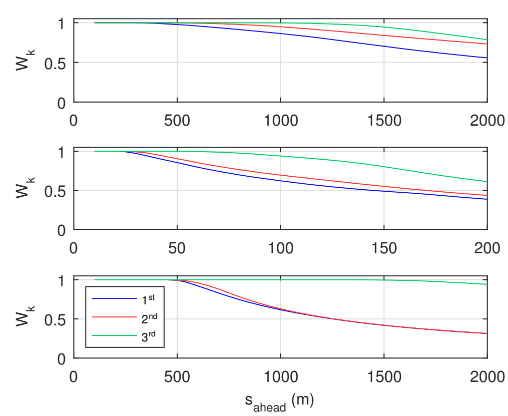

Fig. 12. Prediction confidence *vs.* look-ahead distance for the (top) Brazos, (center) L'Anguille, and (bottom) Colorado Rivers. The 1st, 2nd, and 3rd refer to the three positions on meanders shown in Figs. 4, 6 and 8, respectively.

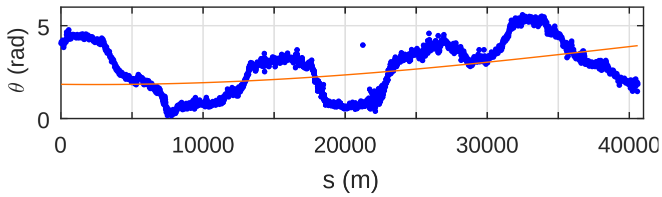

Fig. 13. The estimator clearly fails to converge on this long stretch of data, in contrast to successes on local portions.After fusing 4183 measurements, it appears to ignore the 'high-frequency' structure in the angular data.

most meaningful when thought of as a curve that falls off a one looks further up the river. We define W_k, the prediction confidence at state k, as:

$$W_k(s_{\text{ahead}}) = \int_{-\infty}^{+\infty} \min\left(\frac{s_{\text{in}}}{s_{\text{ahead}}}, 1\right) \Pr\left(x \mid \hat{x}_k, P_k\right) \, \mathrm{d}x, \quad (18)$$

where $\Pr\left(x \mid \hat{x}_k, P_k\right)$ is the probability density of the normal distribution with mean $\hat{x}_k$ and variance P_k, s_{ahead} is the look-ahead distance and s_{in} is how far the look-ahead distance remains inside the riverbanks. The value of s_{in} is computed starting from the Cartesian coordinates of the position at s_k. (Note that, up to this point, the paper has considered only the centerline of any river; computation of s_{in} also requires knowledge of the river's breadth.)

Fig. 12 shows the prediction confidence versus look-ahead distance for three positions on meanders shown in Figs. 4, 6, and 8. This quantifies the improvement in prediction as the filter processes more observations (from the first location to the second, and to the third), indicating how the robot can have more confidence in its estimation. Analyzing these curves allows one to understand the horizon over which a planner might safely and profitably operate.

C. Nonstationarity in the meander model

Following the evaluation reported in Section IV-C, we attempted to run CIUKF on the full homeward trajectory taken by the boat, shown in Fig. 10. The measured and the end-location CIUKF estimated angles are shown in Fig. 13 in a plot similar figures before. In this case, the failure of CIUKF to fit a sinusoidal curve to the measured directional angles is obvious. We posit that this failure is because the overall trajectory of the river is insufficiently well-characterized with a sine-generated curve for any choice of parameters.

One might have better success on this full trajectory by regarding it as a piecewise concatenation of multiple sine generated curves, each with local sine-generated curves. One may be able to reset the filter once a series of measurements indicate poor fit and converge (from the prior) to a good local characterization. (This is not an entirely speculation, as the data in Fig. 9 are very similar to a window in Fig. 13, one being part of the outward journey, the other being the full return.) Alternatively, with a better understanding of how non-stationarity is manifest in meanders, one may be able to capture this in (8). These are directions for future work.

VI. Conclusion

In this paper, we have shown how to use a simple but classical geological model of watercourses to parameterize estimators. The periodic and non-linear form of the model, while quite natural seeing as river meanders are themselves characteristically periodic phenomena, poses challenges for straightforward Kalman-based filters. Our results provide convincing evidence that imposing state space constraints to ensure unimodality improves the quality of prediction estimates, helping achieve convergence. The model of meanders has been shown to be applicable across an impressive range of scales, from small streams to cross-continental rivers. In our evaluation too, we examine diverse scales of river. When the observations are from a part of the river that is well-described by the meander model, no matter the particular scale of the river, the estimates are sufficient to aid a planner.

More broadly, the regularity induced by a flowing stream of water represents an important opportunity for the roboticist. Relatively little research has incorporated such structure, but the present paper provides only one example of a rich lode ready for exploitation.

Acknowledgements

Part of the research costs were covered by awards IIS-1302393 and IIS-1453652, support for which we thank the NSF.

We also wish to thank our skipper, John, who used his 30 years of expedience in piloting his trusty ski-boat (see Fig 14) along the Colorado and ensured we were able to collect good-quality data safely.

Fig. 14. The boat used for data collection, see Section IV-C.

References

[1] G. Chowdhary and R. Jategaonkar. Aerodynamic Parameter Estimation from Flight Data Applying Extended and Unscented Kalman Filter. *Aerospace Science and Technology*, 14(2):106–117, 2010.

[2] K. Qin and D. A. Shell. Estimators for anticipating river meanders: a comparative study. Technical Report 2017-2-1, Department of Computer Science & Engineering, Texas A&M University, 2017.

[3] L. B. Leopold and W. B. Langbein. River Meanders. *Scientific American*, 214:60–70, 1966.

[4] D. E. Mecklenburg and A. D. Jayakaran. Dimensioning the Sine-Generated Curve Meander Geometry. *Journal of the American Water Resources Association*, 48(3):635–642, February 2012.

[5] M. J. Quinlan and R. H. Middleton. Multiple Model Kalman Filters: A Localization Technique for RoboCup Soccer. In *Robot Soccer World Cup*, volume 5949, pages 276–287. Springer, LNCS, 2009.

[6] D. J. Simon. Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *IET Control Theory & Applications*, 4(8):1303–1318, August 2010.

[7] F. D. Snyder, D. D. Morris, P. H. Haley, R. T. Collins, and A. M. Okerholm. Autonomous River navigation. In *Proceedings of SPIE, Mobile Robots XVII*, volume 5609, pages 221–232, December 2004.

[8] B. O. S. Teixeira, L. A. B. Tôrres, L. A. Aguirre, and D. S. Bernstein. Unscented filtering for interval-constrained nonlinear systems. In *IEEE Conference on Decision and Control*, pages 5116–5121, 2008.

[9] T. R. Thakur and A. E. Scheidegger. A test of the statistical theory of meander formation. *Water Resources Research*, 4(2):317–329, 1968.

[10] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. MIT press, 2005.

[11] P. Vachhani, S. Narasimhan, and R. Rengaswamy. Robust and reliable estimation via Unscented Recursive Nonlinear Dynamic Data Reconciliation. *Journal of Process Control*, 16(10):1075–1086, 2006.

[12] P. Vachhani, R. Rengaswamy, V. Gangwal, and S. Narasimhan. Recursive estimation in constrained nonlinear dynamical systems. *American Institute of Chemical Engineers Journal*, 51(3):946–959, 2005.