

Relaxing the Planar Assumption: 3D State Estimation for an Autonomous Surface Vessel

Gregory Hitz¹, François Pomerleau^{1,2}, Francis Colas^{1,3}, and Roland Siegwart¹

¹ETH Zurich, Autonomous Systems Lab, Switzerland.

²University of Toronto, Autonomous Space Robotics Lab, Canada.

³Inria, Villers-lès-Nancy, F-54600, France.

³CNRS, Loria, UMR 7503, Vandœuvre-lès-Nancy, F-54500, France.

³Université de Lorraine, Loria, UMR 7503, Vandœuvre-lès-Nancy, F-54500, France.

July 9, 2015

Abstract

Autonomous Surface Vessels (ASVs) are increasingly proposed as tools to automatize environmental data collection, bathymetric mapping and shoreline monitoring. For many applications it can be assumed that the boat operates on a 2D plane. However, with the involvement of exteroceptive sensors like cameras or laser rangefinders, knowing the 3D pose of the boat becomes critical. In this paper, we formulate three different algorithms based on 3D extended Kalman filter (EKF) state estimation for ASVs localization. We compare them using field testing results with ground truth measurements, and demonstrate that the best performance is achieved with a model-based solution in combination with a complementary filter for attitude estimation. Furthermore, we present a parameter identification methodology and show that it also yields accurate results when used with inexpensive sensors. Finally, we present a long-term series (i.e., over a full year) of shoreline monitoring data sets and discuss the need for map maintenance routines based on a variant of the Iterative Closest Point (ICP) algorithm.

1 Introduction

Autonomous Surface Vessels (ASVs) are most often deployed to help with environmental data collection [Grinham et al., 2011, Hitz et al., 2012]. Besides that, more and more sensors and applications have been considered, for instance bathymetry [Chen et al., 2008], and shoreline monitoring or mapping [Scherer et al., 2012]. In many such applications, the operation space of the vessel is assumed to be two-dimensional (2D), as it operates in open spaces where oscillations in attitude do not affect the primary application (e.g., way-point and trajectory following [Dhari-

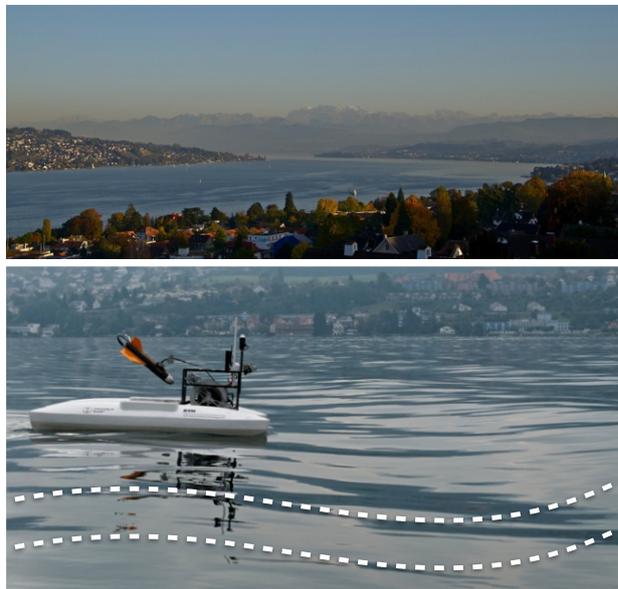


Figure 1: State estimation needs on a lake at different scales. *Top*: Overview of Lake Zurich, for which 2D state estimation fulfills the needs of most application cases. *Bottom*: Zoom on the ASV Lizhbeth with low frequency waves requiring 3D state estimation to observe small scale features.

wal and Sukhatme, 2007, Bibuli et al., 2009, Feemster and Esposito, 2011], environmental data collection [Dunbabin et al., 2009, Grinham et al., 2011, Hitz et al., 2012] or visual shoreline monitoring [Griffith et al., 2014]). Although the water surface is flat on average (Figure 1-*top*), waves can induce a significant three-dimensional (3D) excitation (Figure 1-*bottom*). Distal measurements relying on bearing, such as these from cameras or laser sensors, are particularly sensitive to this problem and prevent two-dimensional (2D) state estimation to be sufficient.

Given a relatively long time to generate a point cloud with a tilting sensor (e.g., approximately 2 s in our case), the pose of the ASV has usually significantly changed in terms of both position and attitude. If not taken into account it yields distortion, or motion blur, in the point cloud. For example, typical 5° pitch or roll oscillations can introduce a displacement of up to 0.8 m to a feature at a distance of 10 m.

A deeper investigation of pose computation, relying on both position and attitude estimates, is required to tackle the problem of local perturbations impacting on ASV localization. Several approaches for 3D state estimation are proposed in the literature. In Section 3 of this article, we formulate three different extended Kalman filter (EKF) approaches adapted for our ASV equipped with exteroceptive sensors affected by motion-blur. We discuss the advantages and disadvantages of using model-based EKFs. Even though our use-case involves the use of a laser rangefinder, which could potentially provide useful measurements for the estimation of the pose of the boat, we aim at providing a solution based on the very common and basic set of sensors used on ASVs (GPS, compass and Inertial Measurement Unit (IMU)). This restriction enables us to provide a generic solution for surface vessels which could be extended when additional sensors are present.

Some of the parameters of the model-based estimators that we present can be derived from hardware specifications, but others still have to be estimated. In order to fit these estimated parameters, we present a generic methodology which, while benefiting from ground truth position, is also deployable using simple GPS measurements (Section 4). We did extensive field testing using well-defined metrics and ground truth positioning, and thus were able to compare the three approaches in Section 5. In particular, the collected ground truth measurements allow us also to highlight problems which arise from non-Gaussian error components on GPS measurements. The developed state estimation solution enables the registration of undistorted point clouds into a 3D map of the shore. This article presents the analysis and the evolution of this map across a year. By discriminating between static and dynamic points, we are able to both maintain a clean representation of the environment and highlight changes in the vegetation and mooring position of boats (Section 6).

The core contributions of this work can be summarized as follows:

1. comparison of three approaches for 3D state estimation for ASVs,
2. parameter identification scheme for the dynamic models of the filters.

A shorter version of this work was published in the proceedings of the *International Symposium on Experimental Robotics (ISER)* [Hitz et al., 2014b]. Here, we present the

state estimation approaches and their differences in more detail. We also complete the description of the model-based estimator with a discussion on the required parameters and the methods to identify them for the deployed ASV. Furthermore, the experimental evaluation is considerably extended with additional data sets, including more accurate attitude results. Finally, we lengthen the long-term shoreline monitoring data set to span an entire year and detail environmental changes.

2 Related Work

As mentioned above, ASVs are increasingly deployed as autonomous information gathering tools in aquatic environments. Besides taking measurements within the water column below the boat, such as in the work of Brown et al. [2011] and Dunbabin et al. [2009], or also prior work with the present ASV [Hitz et al., 2014a], ASVs have been used to take measurements of features located above the surface. Hollinger and Sukhatme [2014] have presented an application of an ASV on a lake to estimate the scalar field of wireless signal strength in the context of planning efficient measuring paths. In contrast to the aim of estimating continuous scalar fields, Subramanian et al. [2006] have deployed an omni-directional camera and proposed a water-segmentation approach to extract 2D maps of the shoreline. The work of Griffith et al. [2014] is also camera-based, but aims at long-term monitoring of shoreline vegetation at a high level of integration and autonomy. As cameras have high shutter speeds, roll and pitch motion can be handled in post-processing. In a slightly different context, Scherer et al. [2012] have developed solutions to map river banks with a rotating 2D laser rangefinder. Their work is similar to ours, despite the fact that they aimed at using quadrotors instead of surface vessels. However, their solution depends on visual odometry, as it is often used in Unmanned Aerial Vehicle (UAV) applications.

State estimation solutions for ASVs have been investigated on a variety of platforms, such as the Unmanned Surface Vessel named *Charlie USV* [Bibuli et al., 2009], an over-actuated single-hull ASV [Feemster and Esposito, 2011] and small airboats [Dhariwal and Sukhatme, 2007]. However, these works aimed at deploying platforms in open spaces. Thus, the operation space of the vessels is assumed to be two-dimensional, neglecting roll and pitch motions. Fossen [2011] presents a comprehensive point-mass model for surface vessels, which is adopted both in our approach and in the aforementioned 2D approaches. Some works mention the use of commercial state estimation devices on ASVs [Subramanian et al., 2006, Neal et al., 2012, Almeida et al., 2010], which produce attitude estimates. However, it is not clear what the measured quantities are, and only little information on the performance of such solutions is provided. While many ASV systems have been

presented throughout the last two decades, there has only been little focus on wave induced measurement distortions. To the knowledge of the authors, there has been no detailed work on roll and pitch estimation for small ASVs.

Other than in works related to ASVs, full 3D state estimation has been widely investigated for UAVs. The use of the IMU-centric indirect EKF [Trawny and Roumeliotis, 2005] has become very popular, as it avoids the formulation of dynamic models. Especially for quadrotors, its combination with visual-inertial solutions [Weiss et al., 2013] has been used frequently because models of these systems tend to be very complex. For large-scale aerial applications with fixed-wing vehicles, Leutenegger and Siegwart [2012] reported a state estimator that does not use visual odometry. Despite the different operation regime, one of the approaches we compare in the following is very similar to theirs as they deal with a similar set of sensors. The Complementary filter (CF) presented by Mahony et al. [2008] aims at estimating only the attitude of an UAV given measurements of known directions. This last work lay down the basis for the third approach the we used in our current analysis.

Finally, range measurements can be used to monitor shoreline vegetation, but can also facilitate an additional localization method for the ASV. The registration of subsequent point clouds provides relative position updates and, when going a step further, the comparison against a known map links the position information to a global frame. Rectifying laser scan measurements via state prediction during the swiping has been done for other types of robots (e.g., ground robots [Kubelka et al., 2014] and quadrotors [Michael et al., 2012]). Another approach is to include the laser information directly in the state estimation [Bosse and Zlot, 2009]. This solution works well when the surrounding environment properly constrains the scan registration, but sensor maximum range coupled with water reflectivity can reduce the number of returned points to an unstable level. However, only few ASVs feature exteroceptive sensors and there is a large variety of sensors, which could potentially be deployed on ASVs. As it is our goal to provide a baseline for comparisons that is applicable for many systems and various applications, we concentrate this work on using GPS, compass and IMU data only.

3 State Estimation Approaches

The goal of using state estimation procedures is to estimate the state of a system in a consistent and robust manner, given measurements from a set of noisy sensors. In our application, we want to estimate the position and attitude of the body fixed frame $\{B\}$ with respect to an inertial world frame $\{W\}$. For this purpose, we use a Kalman filters (KFs), which uses a prediction model to assess how probable measurements are when they are fed to the filter.

For the sake of brevity, the KF concept, which can be found in many good textbooks, will not be explained here. Therefore, this section only describes the specificities of the filter variants under investigation.

As for many ASVs in research projects, ours features three basic sensors (GPS, IMU, compass) that provide data to the state estimation procedure. The GPS receiver measures the position of the boat with respect to $\{W\}$. A digital compass provides the north vector in $\{B\}$. Our ASV uses a simple one-dimensional compass commonly found on boats. Furthermore, the IMU measures the accelerations and the angular velocities in $\{B\}$. We assume that all the sensor measurements are corrupted by additive white Gaussian noise. We also assume that the accelerometers and the gyroscopes of the IMU are additionally distorted by time-varying, additive biases, which need to be estimated along with the states of the system.

This set of sensors is most suited for operation in open spaces where GPS errors of up to a few meters are negligible. When operating with distal sensors in the vicinity of the shoreline it might be beneficial to integrate observations of features on the shore, as these might provide position information with higher accuracy than GPS. However, we refrain from feeding distal measurements to the state estimator for two reasons. First, the aforementioned basic set of sensors is available on many small ASVs, thus the evaluation we present in this paper is useful for a broader community of researchers. Second, it provides a baseline for the comparison of more complex and specialized state estimators, which take into account measurements from additional sensors.

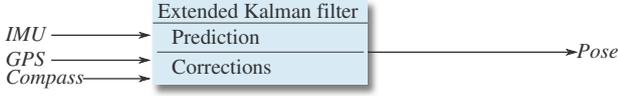
The world frame $\{W\}$ is defined as a metric North-East-Down frame (NED). In our specific case, we use the Swiss grid,¹ which is based on a Mercator projection. However, any other metric coordinates could be used (e.g., UTM). We directly convert GPS measurements from WGS84 to Swiss grid coordinates before feeding them to the state estimator. The body frame $\{B\}$ of the boat is defined similarly to the world frame $\{W\}$ with the x axis pointing forwards, the y axis to the right and z downwards.

We compare three different formulations of EKFs. Figure 2 gives a basic overview of their working principles. The IMU-centric (IC) implementation makes use of the IMU measurements in the prediction step and treats only the GPS and compass measurements as updates. The Model-Based version (MB3D) implements regular updates with the IMU measurements and uses a dynamic model for the state propagation. Thus, the thrust inputs of the motors are also required as inputs to the filter. The third version, Model-Based with Complementary Filter (MBCF), is a hybrid version, which combines the model-based approach

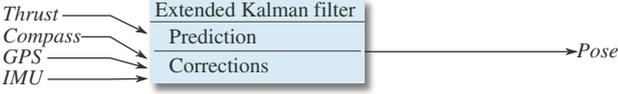
¹Swiss map projections, Federal Office of Topography, <http://www.swisstopo.admin.ch/internet/swisstopo/en/home/topics/survey/sys/refsys/projections.html> accessed May 18, 2014

with a purely sensor-driven estimation of the attitude. In the following, we will describe all three implementations in more details.

IMU-centric version (IC):



Model-based version (MB3D):



Model-based version with Complementary filter (MBCF):

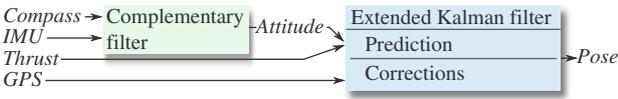


Figure 2: The three filtering schemes used for state estimation with their different inputs. From top to bottom: IC, MB3D and MBCF.

3.1 IMU-centric Extended Kalman Filter

The IMU-centric filter implements the state-of-the-art EKF formulation that is mostly used in UAV applications. It is favored for complex systems since it circumvents the use of dynamical models by directly integrating the IMU measurements in the prediction steps, treating them as a system input rather than actual sensor measurements. This simplifies the prediction model to simple kinematic equations, which can be formulated generically for rigid bodies and thus can be applied to various robotic systems. The measurements from the GPS receiver and the compass are treated as regular updates in the EKF. Our formulation of the IMU-centric EKF follows the one of [Leutenegger and Siegwart \[2012\]](#). The state vector is defined as:

$$x = [p, q, v, b_g, b_a]^T$$

where $p \in \mathbb{R}^3$ defines the position in the world frame $\{W\}$, $q \in SO(3)$ is the attitude represented as a unit quaternion, $v \in \mathbb{R}^3$ is the linear velocity of the boat in the world frame $\{W\}$, $b_g \in \mathbb{R}^3$ and $b_a \in \mathbb{R}^3$ describe the biases of the gyroscopes and accelerometers respectively. The attitude quaternion $q = [q_1, \rho^T]^T$ defines a rotation of an angle α around an axis $a \in \mathbb{R}^3$, with $q_1 = \cos(\alpha/2)$ and $\rho = a \sin(\alpha)$. The state dynamics are then described as

follows:

$$\begin{aligned} \dot{p} &= v \\ \dot{q} &= \frac{1}{2} \Xi(q) ({}_B \tilde{\omega} - b_g) \\ \dot{v} &= R_{WB}(q) ({}_B \tilde{a} - b_a) + {}_W g \\ \dot{b}_g &= 0 \\ \dot{b}_a &= -\frac{1}{\tau} b_a \end{aligned}$$

where ${}_B \tilde{\omega}$ and ${}_B \tilde{a}$ are the gyroscope and accelerometer measurements, ${}_W g$ is the gravity vector, $R_{WB}(q)$ is the rotation matrix from the $\{B\}$ to $\{W\}$ and

$$\Xi(q) = \begin{bmatrix} \rho^T \\ q_1 \mathbb{I}_{3 \times 3} + [\rho]_{\times} \end{bmatrix} \in \mathbb{R}^{4 \times 3}.$$

Here, $[\cdot]_{\times}$ denotes the cross-product operator, which is defined such that $[a]_{\times} b = a \times b$ for $a, b \in \mathbb{R}^3$:

$$[a]_{\times} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \quad (1)$$

The covariance matrix of the quaternion representing the attitude is not well defined due to its unit length constraint [[Trawny and Roumeliotis, 2005](#)]. In order to circumvent this problem, the filter is designed to operate on the error state Δx . For the attitude components, the error state is defined multiplicatively: $\hat{q} = \delta q \otimes q$, whereas for the other states it is additive: $\hat{x}_{\setminus q} = x_{\setminus q} + \Delta x_{\setminus q}$. Here, $x_{\setminus q}$ denotes the components of the state vector x except the attitude quaternion q . The resulting error quaternion δq can be reduced to a 3D representation using small angle approximations [[Trawny and Roumeliotis, 2005](#)], which renders the covariance matrix non-singular again. The full derivation of the Jacobian matrix is not provided here, as it is provided in detail by [Leutenegger and Siegwart \[2012\]](#).

3.2 Model-based Extended Kalman Filter

The model-based EKF implementation uses a dynamic model, which describes the dynamics of the system based on the commands sent to the motors. It makes use of specific knowledge about the systems dynamics rather than only using sensor information. For the derivation of the model, we follow the work of [Fossen \[2011\]](#). The model is defined by a kinematic part (Equation 2) and a dynamic part (Equation 3):

$$\dot{\eta} = J(\eta) v \quad (2)$$

$$M \dot{v} + C(v) v + D(v) v + g(\eta) = \tau_m + \tau_e \quad (3)$$

where $\eta \in \mathbb{R}^3 \times SO(3)$ is the position and attitude of the boat defined in the world frame $\{W\}$, $v \in \mathbb{R}^6$ the boat's linear and rotational velocities in the body frame $\{B\}$, M

the inertia matrix, C the Coriolis matrix, D the damping matrix, g a vector of hydrostatic restoring forces, τ_m the motor forces as input and τ_e unknown external perturbations. $J(\eta)$ defines the transformation matrix from body fixed velocities to a world frame representation. In this generic formulation, the model has a large number of parameters which can be difficult to identify. In general, all matrices of the system dynamics (Equation 3) have dimension $\mathbb{R}^{6 \times 6}$. Whereas the inertia matrix and the Coriolis matrix can be derived mathematically (given exact knowledge of the mass distribution of the system), other parts are more difficult to identify. In particular, the damping matrix $D(v)$ has 36 parameters when formulated linearly ($D(v) = D_l v$, with $D_l \in \mathbb{R}^{6 \times 6}$), but potentially even more when respecting non-linear effects.

This highlights the need for pragmatic simplifications to the model, as the number of parameters prohibits an efficient model identification process. We explain the simplifying assumptions that we made for the model of our ASV along with its derivation. We use a set of Euler angles to represent the attitude of the boat, which simplifies the modeling of the hydrodynamic restoring forces, as we will explain in the following. Thus, we define η as follows:

$$\eta = [p^\top, \Theta^\top]^\top$$

where p denotes the position of the boat in $\{W\}$ and $\Theta = [\phi, \theta, \psi]$ is a set of Euler angles. Note that the well-known singularity at 90° pitch is inherently not problematic for a surface vessel. We define the model (Equation 3) around the center of gravity of the body. Therefore, the inertia matrix can be formulated as follows:

$$M = \begin{bmatrix} m \mathbb{I}_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & I_g \end{bmatrix}$$

where m is the mass of the boat, $\mathbb{I}_{3 \times 3}$ is the 3×3 identity matrix, and $I_g \in \mathbb{R}^{3 \times 3}$ is the rotational inertia matrix. As a further simplification, we assume that $I_g = \text{diag}(I_\phi, I_\theta, I_\psi)$. The Coriolis matrix is defined as follows:

$$C(v) = \begin{bmatrix} m[v_R]_\times & 0_{3 \times 3} \\ 0_{3 \times 3} & -[I_g v_R]_\times \end{bmatrix}$$

where $v_R \in \mathbb{R}^3$ denotes the rotational part of v and $[\cdot]_\times$ the cross-product operator as defined in Equation 1.

Simple linear damping forces have proven to yield satisfactory results, at least in the velocity regimes that our ASV operates in (i.e., up to 1.5 m/s). The simplest model would assume fully decoupled dimensions allowing us to use a diagonal damping matrix D . This assumption has shown to be adequate, except for one non-negligible coupling between the surge (x) and the yaw (ψ) components, which produce the off-diagonal element named $D_{x\psi}$. The coupling of these two dimensions reflects the fact that the catamaran configuration of the boat induces higher yaw damping forces when moving forwards. Thus the damping ma-

trix takes this non-diagonal form:

$$D(v) = \begin{bmatrix} D_x & 0 & 0 & 0 & 0 & 0 \\ 0 & D_y & 0 & 0 & 0 & 0 \\ 0 & 0 & D_z & 0 & 0 & 0 \\ 0 & 0 & 0 & D_\phi & 0 & 0 \\ 0 & 0 & 0 & 0 & D_\theta & 0 \\ D_{x\psi} & 0 & 0 & 0 & 0 & D_\psi \end{bmatrix}.$$

Note, that the damping matrix is not symmetric in general [Fossen, 2011], and therefore the simplification of setting $D_{x\psi}$ to 0 is valid.

The hydrostatic restoring forces are the most difficult to model. The buoyancy of a vessel is defined by the mass of water that it displaces, which depends on the (unknown) shape of the local water surface surrounding the boat. Theoretically, each wave changes the buoyancy force that acts on the vessel. Even when assuming a flat water surface, the buoyancy depends on the geometry of the boat. While the geometry is usually known, it is less straightforward to define the derivatives thereof, which are necessary to formulate the Jacobian matrix of the entire system. Therefore, we adopt the simplified formulation by Fossen [2011], which assumes small roll and pitch angles and box-shaped vessels. It decouples roll and pitch motion and results in linear terms:

$$g(\eta) = G\eta, \quad \text{where } G = \text{diag}(0, 0, G_z, G_\phi, G_\theta, 0). \quad (4)$$

Using this simplification, the resulting dynamics are described by a second order linear system.

The thruster commands t_{left} and t_{right} to the motors of the boat are encoded as the rates with respect to full power: $t_{\text{left/right}} \in (-1, 1)$. The resulting motor forces F_{left} and F_{right} have quadratic form:

$$\tau_m = [F_{\text{left}} + F_{\text{right}}, 0, 0, 0, 0, d(F_{\text{left}} - F_{\text{right}})]^\top$$

$$F_i(t_i) = \begin{cases} at_i^2 + bt_i & \text{if } t_i \geq 0 \\ \beta[-at_i^2 + bt_i] & \text{if } t_i < 0 \end{cases}$$

$$i \in \{\text{left}, \text{right}\}$$

where d is the lateral distance from the center of gravity to the motors, $\beta \in (0, 1)$ is a scaling factor accounting for the inefficiency of the thrusters in backwards direction,² and a and b are the parameters of the quadratic form. Since these two parameters do not have a clear physical meaning, we transform them linearly, such that the resulting parameters are easier to interpret:

$$f_{1.0} = F_i(1.0) = a + b$$

$$f_{0.5} = F_i(0.5) = a/4 + b/2$$

where $f_{1.0}$ and $f_{0.5}$ are respectively the motor force at full and half thrust.

²Most water propellers have asymmetric blade shapes optimized for forward propulsion.

The resulting model has 17 parameters, which are the mass of the boat m , three rotational inertia parameters I_ϕ, I_θ, I_ψ , seven damping coefficients in D , three parameters for the restoring forces and finally three parameters for the thruster forces. We will discuss the identification procedures for the parameters in the following section.

The state vector for the model-based filter is defined as $x = [\eta^\top, v^\top, b_g]^\top \in \mathbb{R}^3 \times SO(3) \times \mathbb{R}^6 \times \mathbb{R}^3$, where b_g denotes the additive biases of the gyroscopes. We discretize the state at constant time intervals Δt and use Euler forward integration to compute state transitions:

$$x_{k+1} = x_k + \Delta t \dot{x}_k, \quad \text{where } x_k = x(t_0 + k\Delta t).$$

The Jacobian matrix $H \in \mathbb{R}^{15 \times 15}$ can be computed analytically:

$$H = \frac{\partial x_{k+1}}{\partial x_k} = I + \Delta t \frac{\partial \dot{x}_k}{\partial x_k}.$$

We omit the full derivation of the Jacobian matrix here, as it is based on simple analytical derivatives.

While the updates of compass, GPS, and gyroscope measurements are straightforward, the measured accelerations can not be used directly. In the general case, the IMU is not situated in the center of gravity and thus the measured accelerations need to be translated to the center of gravity. Since this is done in a moving frame $\{B\}$, the translation depends not only on the rotational velocities, but also on the rotational accelerations. The rotational accelerations can not be measured directly and thus need to be inferred from the noisy and drifting gyroscope measurements. As this is unpractical, the other option would be to shift the center of the model equations to the IMU. This however, makes the model itself far more complicated. For these reasons, we do not use the measured accelerations in the model-based version.

3.3 Attitude Estimation with the Complementary Filter

In our third approach, we use a separate estimator for the attitude of the ASV. For this purpose, we use the Complementary filter (CF) [Mahony et al., 2008], which integrates the gyroscopic measurements ω and corrects them with measurements of known directions c_i . The state of the filter consists of the attitude represented by a unit quaternion q and a vector of additive gyroscope biases b_g . The CF is then defined as follows:

$$\begin{aligned} \omega_c &= -\text{vex} \left(\sum_{i=1}^N \frac{k_i}{2} (c_i \hat{c}_i^\top - \hat{c}_i c_i^\top) \right) \\ \dot{q} &= \frac{1}{2} \hat{q} \otimes (\omega - b_g + k_p \omega_c) \\ \dot{b}_g &= -k_f \omega_c \end{aligned}$$

where $\text{vex}(\cdot)$ defines the inverse operator of the cross-product operator: $\text{vex}([a]_\times) = a$ for $a \in \mathbb{R}^3$. The correction term ω_c depends on the measured directions c , on the estimates thereof \hat{c} , and on their respective weights k_i . k_p and k_f denote filter parameters. In our case, two measured directions are provided by the compass and the accelerometers. The compass measures the north vector and the accelerometers provide the direction of gravity. However, the gravity vector is affected by the accelerations, which are induced by the motors. Such distortions can be avoided by having the model described above estimate the forces caused by the motors and correct the measurements. Given the attitude estimate from the CF, we use a simplified version of the model-based state estimator to estimate the remaining states (positions p and linear velocities v). This can be achieved easily, by only using the first three dimensions of the model (Equations 2 and 3). Since the state vector of the simplified 2D model does not contain the attitude, we need to use q instead of η as an input to the transformation matrix J .

Table 1 provides an overview of the state estimation methods that were introduced above. The 2D version of the model based filter (MB2D) is listed along with the three evaluated 3D filters, as it was used as a baseline in the evaluation and has often been reported in literature (Bibuli et al. [2009], Feemster and Esposito [2011], Dhariwal and Sukhatme [2007]). The IMU-centric (IC) does not require a model as it directly uses the IMU measurements (a and ω) in the predictions. Both model-based solutions (MB3D and MBCF) require the motor inputs $t_{\text{left/right}}$ and a set of model parameters. The model-based version with the complementary filter (MBCF) has a smaller state vector as it does not contain the angular velocities ω .

3.4 Point Cloud Registration

Building upon the state estimation procedures presented above, we can assemble the range measurements of the laser scanner into undistorted point clouds. The goal is to be able to reuse the same map through a full year, thus establishing the foundations of a broader range of applications (e.g., plan safe paths to access the lake, dock the ASV in tight locations repetitively, collect geometric information about the evolution of the environment, connect a bathymetric map to a map generated from a ground vehicle). Using an optimized version of the Iterative Closest Point (ICP) algorithm based on the observations of Pomerleau et al. [2013], we register the point clouds to a global map. To apply this solution, the evaluated state estimation methods need to respect the following requirements:

- *State transitions should be smooth during one 3D swipe.* If not, large discontinuities will result in punctual deformations generating more local minima in the registration process.

Table 1: Summary of the proposed state estimation methods. On the last line, the 2D version of the model-based filter (MB2D) is added for comparison. The measurements p, ψ, a, ω refer to GPS, compass, accelerometer and gyroscope data, respectively. The initial covariance matrix of the filters are denoted by P_{init} and the process noise by Q (both are assumed to be diagonal).

Method	State size	State variables	Sensor data usage		Model parameters	Filter parameters
			Predictions	Updates		
IC	15	p, q, v, b_g, b_a	a, ω	p, ψ	–	$15(P_{\text{init}}) + 15(Q)$
MB3D	15	$p, \Theta, v, \omega, b_g$	$t_{\text{left/right}}$	p, ω, ψ	17	$15(P_{\text{init}}) + 15(Q)$
MBCF	11	p_{2D}, v_{2D}, q, b_g	$t_{\text{left/right}}$	p, a, ω, ψ	8	$4(P_{\text{init}}) + 4(Q) + 4(\text{CF})$
MB2D	7	$p_{2D}, \Psi, v_{2D}, \dot{\Psi}, b_{\Psi}$	$t_{\text{left/right}}$	p, ω, ψ	8	$4(P_{\text{init}}) + 4(Q)$

- *Noise on the 3D estimation should be unbiased.* Otherwise there might be systematic deformations that, observed from a different point of view, could lead to a mismatched model and eventually to discarding parts of the reconstruction.
- *State estimation errors should be within the convergence bound of ICP.* Scan matching registration based on ICP is precise but requires a good prior to converge on a global minimum alignment. The convergence depends on many factors (e.g., structure of the environment, speed of the platform, sensor field of view), some of which can not be controlled [Pomerleau et al., 2013].

State-of-the-art map management methods allow us to distinguish static and dynamic points [Pomerleau et al., 2014]. This brings several benefits. First, shorelines are usually not entirely static, for instance boats that are moored at a single buoy have different positions depending on wind conditions. Second, noisy points, which are introduced by state estimation errors, can be classified as dynamic points through time and can be removed. Finally, it allows to detect seasonal changes in the shoreline, which can be of interest for biological or geology studies, one of the scientific goals of our ASV.

4 Evaluation Methodology

We used an ASV named Lizabeth for the state estimation evaluation and shoreline mapping applications as depicted in Figure 1-*Bottom*. The boat has a catamaran configuration with two thrusters allowing operation without a rudder. It measures $2.5 \times 1.8 \times 0.9$ m and weighs approximately 155 kg. It is fully described in the work of Hitz et al. [2012]. We conducted multiple series of field tests with our ASV to evaluate the different state estimation approaches. The experiments also provided data for both estimating model parameters and assessing sensor noise levels.

4.1 Position Ground Truth

To go beyond comparing the results of the state estimation against its sensor input (GPS and compass measurements), we recorded ground truth positioning information with an external positioning device. We used a theodolite (Leica Total Station TS15), which is able to track a specialized prism up to a distance of 2 km with 2 mm accuracy and outputs the position of the prism at 7 Hz. Thus, we mounted the prism on the boat such that it could be tracked from the shore (see Figure 3). For practical reasons, the prism could not be mounted at the center of gravity of the boat. The mounting frame on the boat is shifted to the front (see Figure 1) and the prism had to be mounted at a certain minimal height to ensure its visibility at all times. The resulting offset of the prism from the center of gravity is $[0.54, 0.31, 1.16]^T$ m and causes significant horizontal offsets due to roll and pitch motions of the boat. These offsets can either be corrected using attitude estimates, or can be considered measurements noise. As the first option could potentially introduce additional errors, we chose the latter. Typical roll and pitch oscillations of up to 5° result in horizontal motions of the prism of ~ 0.1 m. Despite those errors, the theodolite provides measurements that are an order of magnitude more precise than the GPS receiver.

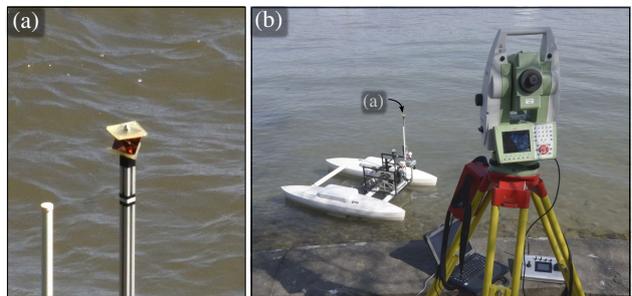


Figure 3: Material used for position ground truth. (a) Specialized prism used as target for automatic tracking. (b) A theodolite (Leica Total Station TS15) mounted on shore and tracking the ASV.

4.2 Attitude Ground Truth

Measuring the attitude of the ASV in a very precise manner is tedious and not often found in literature in outdoor evaluations. One approach, which was used for work on rovers for lunar missions, would be to track three or more reference points on the boat simultaneously, creating a setup similar to indoor tracking systems [Maimone et al., 2007]. Having only one theodolite available rendered the use of external observation methods infeasible. The use of visual markers on shore and a camera on the boat could be a good solution, but the distance between the shore and the boat would lead to unmanageable marker sizes and reduce greatly the accuracy of those measurements.

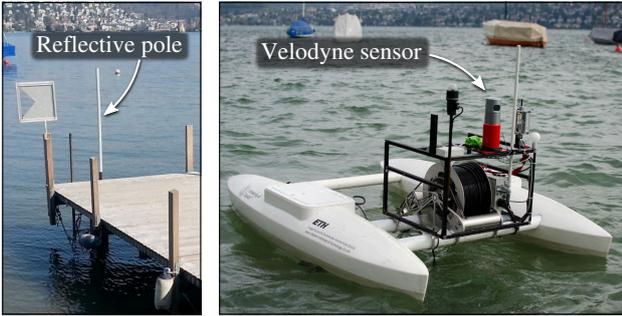


Figure 4: *Left*: The reflective pole used to assess the distortion of point clouds. *Right*: The temporarily mounted Velodyne sensor on the ASV.

To circumvent those challenges, we set up a specific experiment, referred to as *pole experiment* henceforth. A pole covered with a reflective material was mounted vertically on a pontoon and was then scanned repeatedly with a laser scanner on the boat. We mounted a Velodyne HDL-32 onto the ASV exclusively for this test (see Figure 4). This sensor provides full point clouds at 10 Hz and thus provides snapshot-like views of the pole without any noticeable motion-induced distortions in the point clouds. As the measured points on the pole have significantly higher return intensities, they can be extracted from the point cloud simply by applying a constant threshold. By fitting a line to these points, we can use its inclination as a measure to assess the attitude estimation accuracy. In comparison to a camera-based approach this method is not affected by lighting conditions and less sensitive to distance changes.

The Velodyne sensor is very precise, but is also relatively costly. In robotics, rotating 2D laser rangefinders are often used to provide similar 3D point clouds at considerably lower price (e.g., [Scherer et al., 2012, Kubelka et al., 2014]). In fact, for the shoreline mapping application which we describe below, we also used a nodding 2D laser rangefinder (see Figure 6). We repeated the *pole experiment* with this cheaper sensor setup to evaluate if it provides similar results as the use of the Velodyne sensor. As our nodding 2D scanner takes approximately 2 s to com-

plete a full scan and thus the distortion of the pole in the point cloud has to be taken into account. In addition to the inclination measure, the root mean squared error (RMSE) of the line fit provides a measure of distortion. To observe distortions, both induced from roll and pitch motions, the pole has to be positioned at varying bearing angles with respect to the boat. For example, when positioning the pole directly in front of the boat, pure pitch motion mainly distorts the pole along its axis. Such a distortion neither affects the inclination nor the RMSE of the fitted line and is therefore not detectable with these two error metrics. Similarly, positioning the pole either left or right of the boat is unfavorable for detecting errors which are induced by roll motion. Therefore, observations at various bearing angles are necessary.

4.3 State Estimation Evaluation

While the *pole experiment* is most informative about distortions that are induced by roll and pitch motion, the ground truth data from the theodolite enabled us to define the following error metrics:

$$e_{\text{pos},k} = \|p_k - p_{l,k}\|_2 \quad (5)$$

$$e_{\text{vel},k} = \|v_k - v_{l,k}\|_2 \quad (6)$$

$$d_{\text{pos},k} = \|p_{k+1} - p_k\|_2 \quad (7)$$

$$d_{\text{vel},k} = \|v_{k+1} - v_k\|_2 \quad (8)$$

$$d_{\text{att},k} = |\ln(q_{k+1} \otimes q_k^{-1})|. \quad (9)$$

The ground truth measurements from the theodolite are denoted p_l and the derived velocity measurements v_l . The index k refers to the value at time step t_k . Given the ground truth measurements the absolute error in position and velocity can be computed according to Equations 5 and 6. Furthermore, we evaluate the smoothness of an estimated state trajectory with a measure of discontinuity d which is defined for position (d_{pos}), velocity (d_{vel}) and attitude (d_{att}) in Equations 7 to 9. The smoothness measure reveals drastic update steps, which can occur whenever predicted measurements and actual sensor data do not match well in the update steps of the EKF.

For all the statistical significance tests in our evaluation we used the Mann-Whitney U test with $p < 0.001$ as significance threshold.

4.4 Model Parameter Identification

The model-based versions of the state estimators (MB3D and MBCF) make use of predictions, which are based on the model introduced in Section 3. The model has 17 parameters, of which especially the damping parameters are difficult to set. Manually tuned parameters can yield relatively good results when considering that we need the model to perform well only over short time spans in an

EKF (as frequent correction measurements are available). Within the typical field sensor suite that is available on our ASV (GPS, IMU and compass), the GPS receiver operates at the lowest update rate (1 Hz in our application). The non-Gaussian noise on the GPS measurements affects the linear position and velocity components of the state the most. For these reasons, we need the model to be most reliable for these components.

Besides manual tuning, we can apply optimization techniques to find the most suitable set of parameters, with respect to recorded field data. Mišković et al. [2011] presented a method to estimate model parameters *in situ* in an online fashion. Their method relies on exciting the system with self-oscillations and adjusting the parameters iteratively. In comparison to offline batch optimization methods, this is beneficial in the sense that the method acquires as much data as necessary to achieve a stable estimate of the parameters. On the other hand, these experiments need to be repeated as soon as the model is altered and can not be performed with recorded data. Therefore, we chose an offline method, which is less flexible but can be used for various models, error metrics and optimization methods.

For this, we can use either the precise measurements from the theodolite or the GPS and compass readings. Integrating the attitude ground truth of the *pole experiments* is feasible, yet restricts the operation space of the boat significantly, as the pole has to remain visible and close enough, such that it is hit by several of the laser beams of the Velodyne sensor. Therefore, we restrict the parameter optimization to the 2D sub-part of the model. To optimize the parameters, we have defined the following fitness function, which is the sum of the error position after N time steps normalized by the path length:

$$e(\Theta) = \sum_{k=0}^M \frac{\left\| \left[\sum_{i=kT}^{kT+N} \dot{p}_i(p_i, \Theta, \tau_{m,i}) \Delta t \right] - \text{gt}_{kT+N} \right\|}{\sum_{i=kT}^{kT+N-1} \|\text{gt}_{i+1} - \text{gt}_i\|} \quad (10)$$

where Θ is a vector of parameters, p_i the position components of the model state vector x_i at time step i , gt_i denotes the ground truth measurement at time step i and τ_m are the thruster inputs as introduced above. The model is initialized M times at intervals of T time steps using the ground truth information ($p_{kT} = \text{gt}_{kT}$ and $v_{kT} = \dot{\text{gt}}_{kT}$). The model is then used to “blindly” integrate motor commands over a time span of $N\Delta t$. At this point, the error is defined as the quotient of the offset of the model predicted position p_{kT+N} and the total path length over the time span. Figure 5 demonstrates this procedure along with the key variables.

The error metric (Equation 10) does not take heading errors into account. While this might be a shortcoming for certain applications, it has the advantage that the error has physical meaning since it is measured in meters. For non-holonomic systems such as the boat deployed here, errors in heading always lead to position errors as well and it is

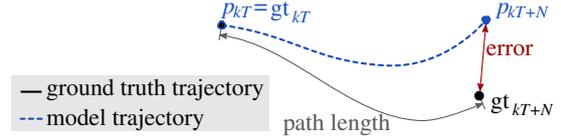


Figure 5: Illustration of the error generated from a wrong model. After integrating over N time steps only based on thruster inputs τ_m , the model trajectory deviates from the ground truth.

thus reasonable not to account for heading errors explicitly in the error metric.

For the optimization, we use the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [Hansen, 2006]. This sampling-based optimization routine does not require gradients and allows us to handle unfeasible parameter sets easily without explicitly defining constraints in the parameter space. Out of the 17 parameters of the model, the 2D sub-model requires nine, namely the mass m , the rotational inertia I_ϕ , damping coefficients $D_x, D_y, D_\psi, D_{x\psi}$ and thruster force parameters $f_{0.5}, f_{1.0}$ and β . We consider the mass to be fixed, as it can be measured *in situ*, thus our optimization space is in \mathbb{R}^8 .

The proposed parameter estimation scheme provides a generic method, which can be applied for many robotic systems. As mentioned before, the error metric can be adapted by adding rotational components if necessary. This will however lead to an additional weighting factor to balance the rotational and the translational components. The optimization routine can be replaced, for instance if explicit gradients are available.

4.5 ICP-Mapping Evaluation

The main focus of this article is to evaluate the state estimation algorithms. However, we can use this opportunity as a preliminary study for long-term mapping with an ASV. The long-term character of the recorded mapping data allows us to observe changes in the environment. However, the dense vegetation around the boathouse prohibited the collection of theodolite measurements, as a clear line of sight to the prism on the boat could not be maintained at all times. At this stage, qualitative results based on visual assessment of the representation will be used. To support our observation, we used a ground survey unit (Leica Nova MS50) to build a precise 3D representation of the boathouse. For the application of shoreline monitoring, we have mounted a nodding laser rangefinder (i.e., Hokuyo UTM-30LX-EW) on the ASV as depicted in Figure 6. This particular configuration allows us to tilt the laser up to 180 degrees, allowing the field of view to extend up to the rear of the boat.

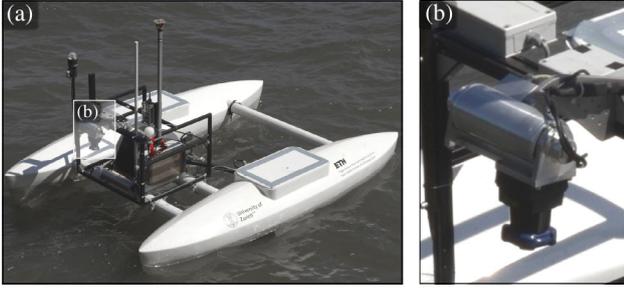


Figure 6: Material used for shore monitoring. (a) Overview of the ASV. (b) A detailed view of the custom tilting Hokuyo UTM-30LX-EW installed on the ASV.

5 Results

We have conducted a series of field experiments, during which we collected data sets consisting of GPS, compass, and IMU measurements, system specific information (such as thrust values of the motors) and external position information from a theodolite. We collected data in 10 one-day campaigns, during which the boat traveled an overall distance of 13.2 km. We encountered varying conditions, ranging from strong winds and rain to very calm days, resulting in data sets with different environmental influences.

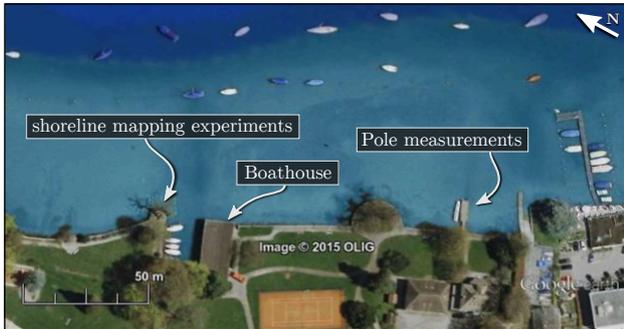


Figure 7: Overview of the testing area, with both structured and unstructured segments on the shoreline.

The collected data sets serve two separate purposes. First, we evaluate the state estimation approaches on dedicated data sets with ground truth information. These data sets do not contain any measurements for mapping purposes. The second type of data sets focused on the shoreline mapping application. They consist of measurements to perform state estimation and the range measurements from the laser scanner. As outlined in Figure 7, those data sets were taken in the area around the boathouse, which provides a rich variety of both structured and unstructured components. We have collected a total of seven shoreline mapping data sets with the laser rangefinder over the course of a full year (from October 2013 to September 2014).

5.1 Model Parameters

For the optimization of the model parameters, several different data sets from different days were used. Figure 8 shows one of the used data sets. The ground truth trajectory is shown in black and the model trajectories in blue. The data set has a total duration of 439 s and the entire path is 304 m long. It consists of different motion types (i.e., straight parts and also sharp turns with low forward speed). Note that the estimation of β does not rely on explicit backwards motion of the boat as turning motions require one of the motors to operate in backwards direction.

For the integration of the model, time steps of $\Delta t = 1/20$ s have proven to be small enough. We integrated the model over 20 s ($N = 400$) and spawned new model trajectories every 5 s ($T = 100$). M is then dependent on the length of the data set, in the example shown in Figure 8 there are $M = 85$ model trajectories. The population size for *CMA-ES* was set to 46, as a cluster computer with 46 cores was available. However, for a parameter space of 8 dimensions, a minimal number of 10 particles is sufficient [Hansen, 2006] and a standard desktop PC can be used. The optimization runs required approximately 150 iterations to converge, on data sets of comparable size as the example shown in Figure 8.

The top view graphs in Figure 8 highlight the improvements of the optimization. Especially the rotational components have lead to more stable results. While the initial parameters result in a median error of 33.1%, the optimized parameter set achieves 7.2% (compare Figure 8-Left). As highly accurate ground truth data might not be available for many systems, we also used the noisy GPS data as a reference. Using the GPS as a reference during the optimization, yields comparable results (9.5%), even when evaluated with respect to the more accurate ground truth measurements (10.2%).

Table 2: Initial and optimized model parameters, both with respect to ground truth (Leica) and GPS.

Parameter	Unit	Initial	Leica	GPS
$f_{1.0}$	N	100.00	52.82	41.06
$f_{0.5}$	N	30.00	20.43	15.85
β	-	0.30	0.37	0.43
D_x	kg s ⁻¹	80.00	38.41	28.22
D_y	kg s ⁻¹	1000.00	5280.05	5261.33
D_ψ	kg m ² s ⁻¹	200.00	104.25	103.19
I_ψ	kg s ⁻²	190.00	13.79	10.65
$D_{x\psi}$	kg s ⁻¹	10.00	24.18	17.69

Table 2 provides an overview of the parameter values. When comparing the parameters of the GPS-based optimization to the corresponding Leica-based values, an average difference of 17.2% can be observed. These differences are largest on the damping coefficient in x -direction

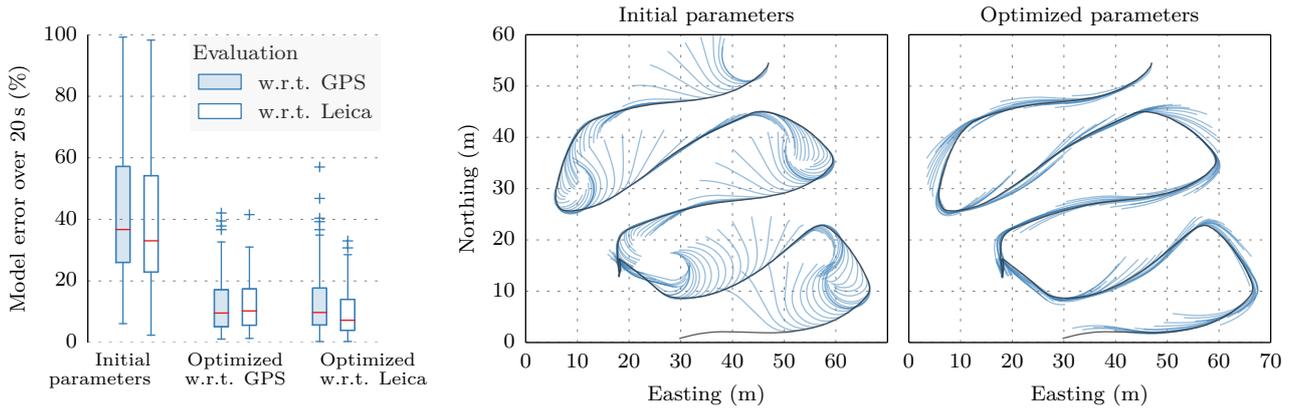


Figure 8: *Left*: Model errors for different sets of parameters. Leica refers to the theodolite ground truth measurements. *Middle*: Ground truth trajectory (black) and model trajectories (blue) over 20 s for the initial parameters. *Right*: trajectories after optimization.

(D_x) and on the thruster model ($f_{1.0}$ and $f_{0.5}$). However, these differences are counteracting and thus both parameter sets result in similar model performance (cf. Figure 8).

From these results, we can conclude that using accurate ground truth data is favorable, but also noisy GPS data provides good results. The difference of using ground truth or GPS data is not statistically significant.

5.2 State Estimation

Using the optimized parameters, a comparison of the proposed state estimation algorithms can be conducted based on field data. Figure 9 shows both the theodolite measurements used as ground truth (GT) and the raw GPS data points along an example of a short round trip path. The third EKF version (MBCF) is not shown in Figure 9, because this filter relies on the same motion model as MB3D and thus the resulting trajectory is almost identical. The graph highlights constant GPS offsets in a particular direction, which appears in the first inset (1), where the GPS measurements start diverging severely from the ground truth. The direction of that offset changes during the execution of the path as shown in inset (2). Inset (1) illustrates well the jitter caused by the GPS information not coherent with IMU measurements leading to a wrongly slow adjustment of the bias for the IC trajectory. As a result of this, the overall performance of all EKF implementations is globally stable (i.e., does not drift over time). However, in terms of the absolute position error, the performance is only slightly better than the raw GPS positioning.

More quantitatively, Figure 10 shows an evaluation of combined data from all five test days according to the metrics defined in equations 5 to 9. For the absolute errors in position (Figure 10a) and velocity (Figure 10b) the raw GPS is provided as a comparison. In the case of the velocity, this refers to the differentiation of the GPS points. Since the

discontinuity measures are not absolute, we can not compare them to the raw GPS input.

As already observed in Figure 9, there is only very little improvement in terms of the absolute position error e_{pos} with respect to the raw GPS measurements. Figure 10a clearly shows the limitations of the GPS measurements, with an average error of 1.9 m. The medians of the absolute position error is reduced by 4.1%, 9.1% and 9.4% for the IC, MB3D and MBCF version respectively compared with only GPS measurements. While these differences are statistically significant, the difference between MB3D and MBCF is not ($p = 0.24$). This correlates with the fact that these two versions are based on the same model for linear motion and thus yield very similar errors. Whereas the positioning information can not be improved significantly, the estimates of the velocity clearly improve with respect to the GPS baseline. The model-based versions (MB3D and MBCF) show the lowest velocity errors (73.5% and 73.2% better than GPS). The IMU-centric EKF has still quite some outliers and only improves by 56.3% with respect to the baseline, which is due to wrong estimates of the accelerometer biases. Since the GPS measurements drift over time, the corresponding position updates map the error to the accelerometer biases and velocity estimates. This effect also leads to less smooth state trajectories in the position and velocity space. Figures 10c and d show this in terms of the discontinuity measure (cf. Eq. 7-9). A high discontinuity measure indicates that the corresponding state was frequently corrected, leading to large differences between two subsequent states. Wrong estimates of the accelerometer biases distort the velocity estimates during the predictions, which then need to be corrected by sensor updates. Especially the discontinuity of the velocity estimates shows that the model-based versions (MB3D and MBCF) result in smoother trajectories than the IC version.

The discontinuity of the attitude estimates suggests that MB3D has higher inconsistencies, which relates to the fact

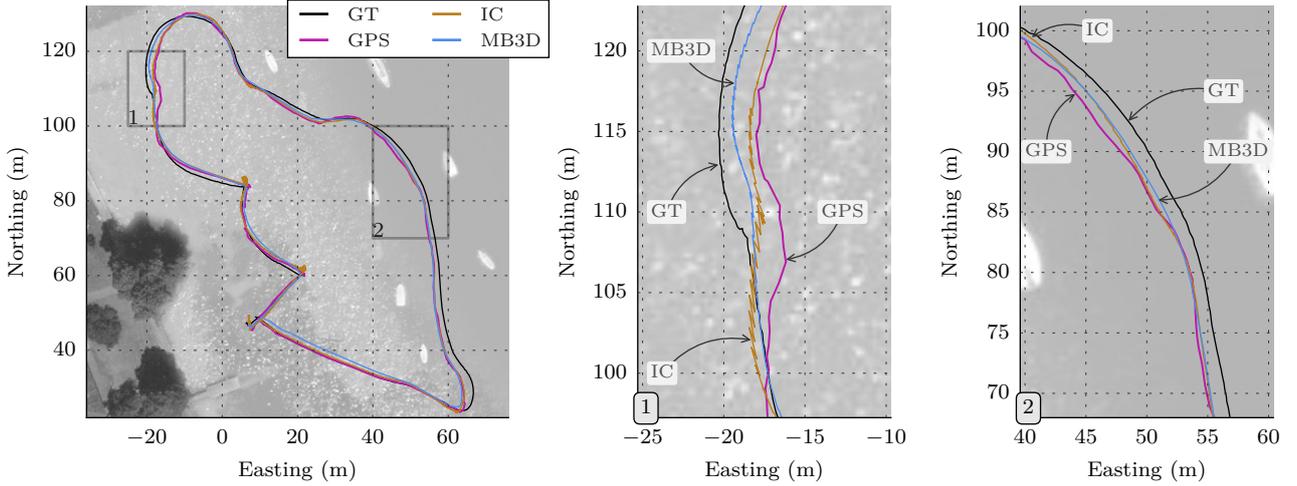


Figure 9: *Left*: Top view of resulting trajectories estimated by the IMU-centric approach (IC, brown line) and the model-based approach (MB3D, blue line). GT refers to the ground truth data from the theodolite (black line) and raw GPS measurement are given for comparison (purple line). The path is overlaid on an aerial map of the lake with the shoreline and trees visible on the lower left corner. *Middle and right*: Zoomed results highlighting GPS offset and its drifting effect on both IC and MB3D.

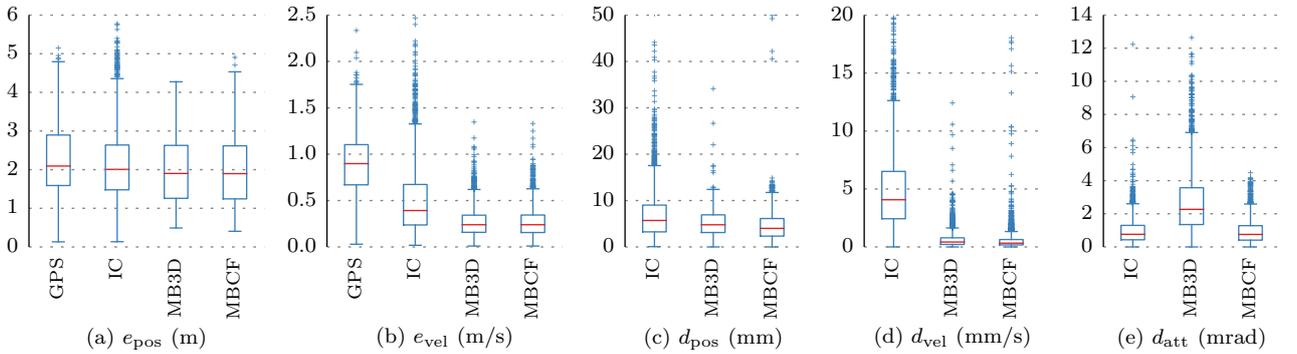


Figure 10: Results of the position evaluation based on theodolite ground truth. The metrics are: position error (e_{pos}), velocity error (e_{vel}), position discontinuity (d_{pos}), velocity discontinuity (d_{vel}), attitude discontinuity (d_{att}). For the first two panels, the raw GPS data is shown as a baseline.

that it relies on a parameter-based model. Even though we estimated these parameters with respect to field data, the linear formulation of the hydrostatic restoring forces (cf. Equation 4) might be too simple to accurately describe the motion. In relation to the MB3D version, the IC and MBCF implementations have lower discontinuities (improvements by 66.4% and 66.7%). Both versions achieve relatively smooth attitude trajectories as they directly integrate the gyroscope measurements. Even though there are quite some differences in their implementations, the resulting distributions of discontinuity measures do not differ significantly ($p=0.08$).

Focusing on the attitude evaluation of each filter, Figure 11 presents the results of both *pole experiments*, with the Velodyne and the nodding Hokuyo sensor. In order to create a baseline for the evaluation, we have used the model-based state estimator in 2D, in order to account for distortions,

which were induced by linear motions, while ignoring roll and pitch. This 2D state estimator version is denoted with MB2D. Figure 11d shows an example of a point cloud, which is distorted by roll and pitch motion. The extracted points from the pole, used as a vertical and straight reference, are highlighted in blue and a 1 m long segment of the fitted line is shown in light blue. The same point cloud is shown in Figure 11e, however it was assembled with the use of the MBCF state estimator. The pole is less distorted and resembles more a straight and vertical cylinder. One can note that such point cloud distortions only occur with the nodding laser rangefinder. The Velodyne sensor creates an almost instantaneous scan of the pole, in which it is then always straight.

Figure 11a shows the results for a data set of 390 point clouds, which were collected with the Velodyne sensor. As expected, all three state estimators perform better than

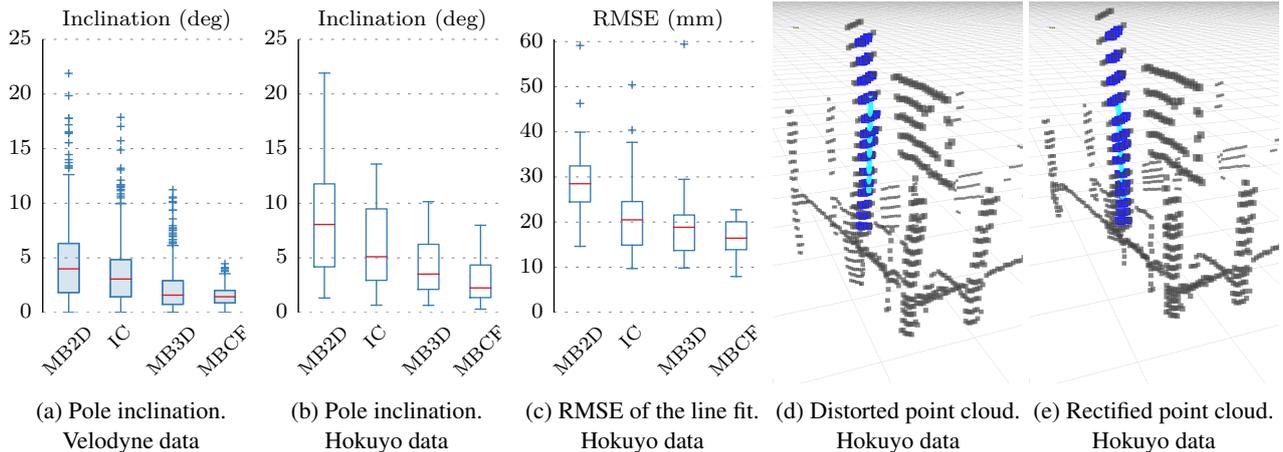


Figure 11: Results of the *pole experiments*. The resulting inclinations of the pole in the Velodyne data set are depicted in (a). The results of the line fitting for the Hokuyo data is shown in (c) and (b). An example of a point cloud (dark grey) with the extracted points on the pole (blue) and the fitted line (cyan) is shown in (d) and (e).

the 2D baseline (MB2D). The MBCF version outperforms the other variants, owing to its good model-based translational components and the direct integration of the gyroscope measurements in the complementary filter. In contrast to the discontinuity measure presented above (cf. Figure 10e), the model-based version MB3D performs relatively well. However, during the *pole experiments* the boat was not actuated. This potentially lead to simpler rocking motions, with less cross-couplings to the translational components of the model, which can be better represented by the simplistic roll/pitch model. Interestingly, the IMU-centric version performs worse. A possible explanation is that the tight coupling of translational and rotational components of the IMU-centric estimator maps erroneous accelerometer bias estimates to the estimated attitude.

The results of the *pole experiment* with the nodding Hokuyo sensor are shown in Figure 11b. The graph clearly shows that this sensor setup leads to the same conclusions as the results of the experiment with the Velodyne sensor. For the Hokuyo sensor setup we also use the RMS error of the line fit as a distortion measure (see Figure 11c). It can be observed that the RMSE and the inclination measure reveal a consistent performance ordering of the state estimation methods. In the results of the *pole experiment* with Velodyne sensor, the RMSE of the line fitting does not contain any comparative information, as the pole it recorded almost instantaneously and thus the RMSE is independent of the state estimation procedure.

There are some differences in the absolute values, which are due to the fact that the two experiments were not performed simultaneously. One could note that the lack of a ground truth measurement of the poles inclination causes a few degrees of uncertainty in the inclination measure in an absolute sense, but the relative performances between solutions remain the same.

This evaluation reveals that the MBCF implementation performs best at recovering the attitude of the pole. Furthermore, it shows that the cheaper sensor setup provides qualitatively similar results to the faster Velodyne sensor. We do not compare the results of the two tests statistically as they were not recorded simultaneously and thus have different inputs for the state estimation methods.

5.3 Point Cloud Mapping

For our application of shoreline monitoring, we have collected seven data sets throughout an entire year. The first map was realized in October 2013 and then all subsequent maps were built upon the previous one up to the final data acquisition in September 2014. Figure 12 shows a top view of the resulting final map. The mapped area contains a small harbor with a boathouse on the left and mooring places on the right. Through the year, different boats were parked and the foliage of trees changes. The boathouse is open towards the harbor, such that the laser rangefinder is able to measure points on the lower side of the roof. We mapped this area by controlling the robot into the harbor and around the willow tree manually.

The map was constructed according to the approach presented by Pomerleau et al. [2014], meaning that the same global map was reused through time leading to an evolving representation over the year. This leads to one combined map over all mapping missions, but comes with the disadvantage that the map does not “forget”. Naively implemented, falsely registered point clouds or environmental changes such as differences in the configuration of the moored boats cause inconsistencies in the map. Furthermore, the number of points in the map continuously grows, which eventually leads to large computational loads for the registration of new point clouds. While simple subsam-

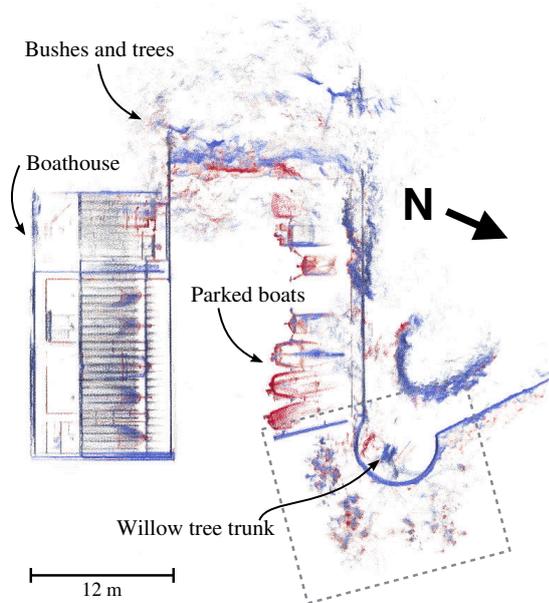


Figure 12: Top view of the resulting point cloud map. Red points have higher probability of being dynamic. The dashed rectangle corresponds to the position where a large willow tree sits and for which an expanded view is provided in Figure 14.

pling of the map would be sufficient for the latter issue, the classification of registered points into dynamic and static points as it is proposed by Pomerleau et al. [2014], provides a solution to both problems. Figure 14e shows the willow tree in the map. Points that are depicted in green have been classified as dynamic. Considering only the points which were labeled as static reveals a representation (see Figure 14b), which is less affected by environmental changes and has significantly fewer points. This reduced and more stable version of the map is then used for local registration using a point-to-plane variant of ICP.

We also provide a map realized from a ground survey laser rangefinder in Figure 13. The main complexity encountered during the survey process is that there was very little place where the Leica MS50 could be positioned to ensure similar view points. As for example, the roof of the boathouse and the brick wall (Figure 13-Right) can only be observed from the water level, where the installation of a tripod was impossible. This leads to two representations, although similar, with very little overlap in horizontal structures. This problem is amplified by the fact that the water surface was partially reflected in the Leica MS50 scans (Figure 13-Left) due to a high incidence angle with the water when the survey unit was positioned on the roof of the boathouse.

An example of stable features learned through the mapping process are the brick wall and trunks of trees as opposed to foliage. This can be observed in Figure 14, which is a de-

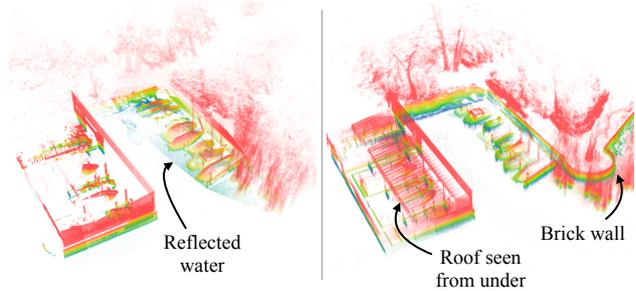


Figure 13: Comparison of a professional land survey unit and a low cost laser rangefinder mounted on a mobile platform. The area used for the experiment is a boathouse where the ASV is usually stored. *Left*: An assembly of 5 scans taken with a Leica MS50 on Sept. 3, 2014. *Right*: Reconstruction based on all points acquired from the robot Lizabeth. Colors represent the elevation, red being high and blue being low.

tailed view of a large tree over a brick wall, expanded from the dashed area of Figure 12. At the entrance of the harbor on the right, there is a large willow tree, which undergoes substantial seasonal changes in foliage (see Figure 14c and 14f).

Those results, although qualitative and preliminary are used in the discussion (Section 6) to highlight foreseen challenges in a more complete deployment for long-term shore monitoring.

6 Discussion

In the light of the above results, we present a deeper analysis around four main themes: 1) problems related to the use of GPS, 2) EKF comparison, 3) amelioration of the ASV model, and 4) challenges related to shore mapping. We also point out certain limitations of our experimental results and discuss potential improvements to our methodology for a deeper knowledge on state estimation for ASVs.

Relying on GPS The standard EKF formulation relies on the assumption that the process and observation noise can be accurately modeled by a zero mean multivariate Gaussian. This assumption is known to be wrong for accelerometers and gyroscopes, for which additive biases are tracked in the vehicle states in order to recenter their noise to zero. Those biases tend to mainly follow the internal temperature of the sensors, which depends on uncontrolled environmental conditions and varies with a low frequency. To be estimated properly, those biases need to be observable (directly or indirectly) using another sensor. In our case, this sensor is the GPS.

Looking at the 2D offset of all the proposed EKF trajectories presented in Figure 9-Left, one can observe that

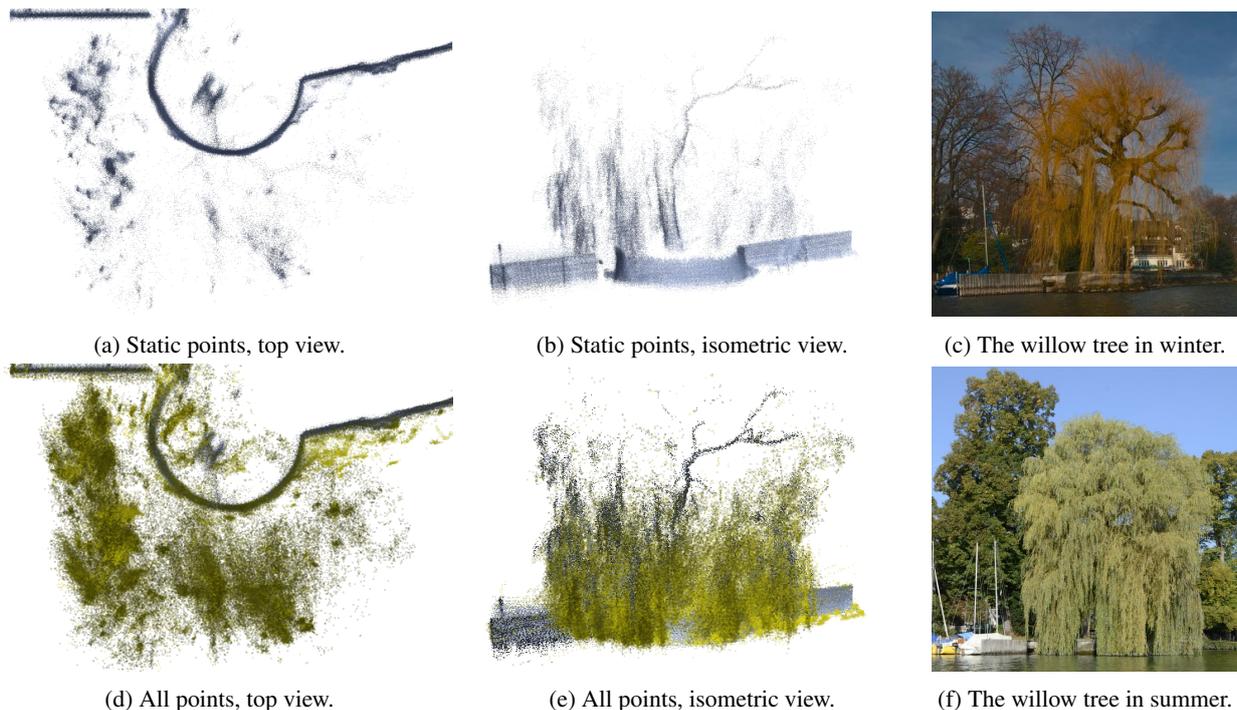


Figure 14: Detailed view of the willow tree in the point cloud map, both in winter (c) and summer (f). Figure (a) and (b) show the static components of the map. The dynamic points are highlighted in green in Figure (d) and (e).

both IC and MB3D follow the ground truth globally nicely. The two insets (1) and (2) show in more details the errors caused by the GPS offsets: as the EKFs assume Gaussian noise on the GPS signal, they eventually drift to reduce the relative error to the GPS signal (see Figure 9, inset (1)). As long as the offset on the GPS signal remains relatively constant (see Figure 9, inset (2)), the estimates from both EKFs cannot reduce the error with respect to the ground truth. This explains the high absolute position errors for all EKF versions as highlighted in Figure 10a.

To evaluate the noise on our specific GPS antenna, we conducted a small experiment with the ASV on the ground (fixed position) for 25 minutes with at least seven satellites in an open area with clear view of the sky (Figure 15-Top). Naively extracting the mean and evaluating the offset error, results in a standard deviation of 1.5 m (see histograms in Figure 15-Top). The temporal evolution of the GPS error is shown in Figure 15-Bottom and reveals strong correlations over time, which resembles more a bounded random walk. Occasionally, the signal shows large constant offsets for up to 3 minutes in perfect conditions. The fact that modeling GPS noise as Gaussian noise is insufficient was also observed in geodetic analysis [Amiri-Simkooei et al., 2007] but not so much in robotics. The addition of multi-path problems in cluttered environments even adds a spatial correlation of the GPS error. ASVs often operate in slow speed regimes (around 1 m/s), which makes it difficult to discriminate the GPS error from actual motion of the vessel. This leads to the conclusion that accelerometers

biases cannot be observed using GPS unless the platform is fast enough and only operates in clear sky conditions. This motivates the use of other exteroceptive measurements into the state estimators, which could observe a GPS bias at a rate faster than its drift.

Selecting the appropriate filter The experimental comparison of the three proposed state estimators shows that the MBCF method performed best for the system at hand. MBCF combines the good performance of the linear components of the model-based approach, with the simplicity of the IMU-centric method for the attitude estimation. We highlight its superior performance with two key figures: Figure 10c and Figure 11a. The first one shows that the model-based approaches (MB3D and MBCF) provide the significantly smoother velocity estimates than the IMU-centric version. The second figure demonstrates that the use of the CF recovers the attitude of the pole best. The simple formulation of the complementary filter directly integrates the IMU measurements, but is not affected by erroneous accelerometer biases as is the IC version. In contrast to the fully model-based approach (MB3D), it circumvents the need for formulating an accurate model of the hydrostatic restoring forces that would describe the rocking motion of the boat.

Our results also show that the IMU-centric version does not perform as well as one might expect, which is mostly due to the error on the GPS signal. In comparison with the work

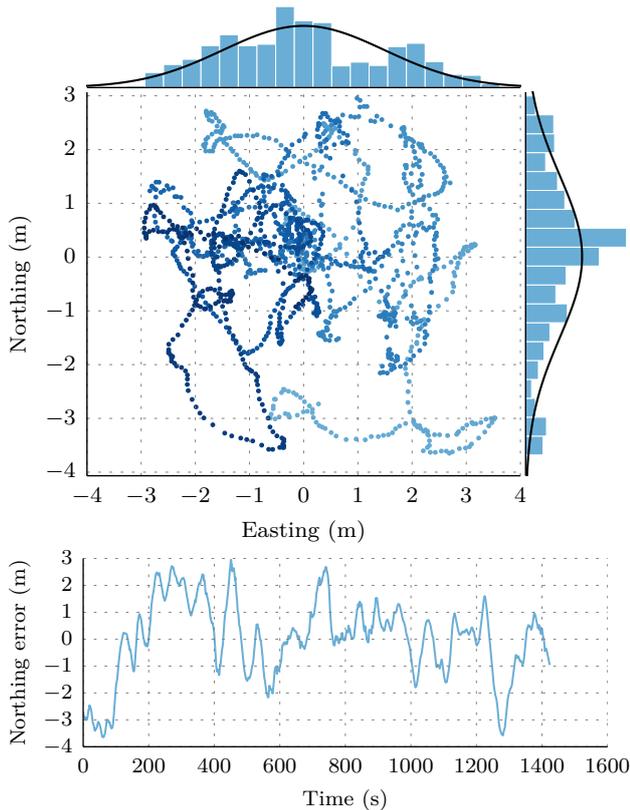


Figure 15: GPS noise analysis of a stationary point (on the shore) over 24 minutes. *Top*: 2D error of the GPS coordinates with respect to the mean. The coloring of the points corresponds to the temporal evolution. A fitted Gaussian distribution is overlaid onto the error histograms. *Bottom*: Evolution of the northing error through time.

of Leutenegger and Siegwart [2012], which explores state estimation for a fixed-wing aircraft, the motion of an ASV is in the same order of magnitude as the error level on the GPS readings. This fact added to the non-Gaussian characteristics of the noise lead to a lower performance when applied on an ASV.

In our evaluation we focused on solutions based on EKF and therefore rely on linearizations. We believe this to be a reasonable choice given the slow dynamics of an ASV. However, it might be interesting to compare our results with more complex filters, such as the unscented Kalman filter (UKF) or a particle filter, and analyse the error that the linearized propagation of the covariance matrix introduces.

Improving ASV models One of the main conclusion is that all filters using a model perform better. This indicates that simplified models can also contribute significant information to the state estimation process. Appropriate parameter identification is one of the bottlenecks of the modeling process. The use of an external tracking sys-

tem for this identification, like a theodolite installed on the shore, can be difficult, if not impossible. Some ASV deployments may require to be faraway from the shore. We compared the use of GPS and theodolite information and conclude that it is possible to use GPS measurements to optimize model parameters, while being aware of its limitations. The quality of the parameter identification will depend on environmental conditions, which might influence the GPS measurements, and on the nature of the data set. Also, when using GPS, the overall trajectory has to span a reasonably large area with respect to the GPS noise, to avoid over-fitting. The same conclusion was obtained from the attitude estimation using a more expensive sensor, the Velodyne, for which similar results could be achieved with a custom-made, slow-tilting laser rangefinder. However, the relatively slow nodding motion with which our laser rangefinder is actuated, might introduce unwanted effects. Due to the low nodding speed, full point clouds are only available at 0.5 Hz, and at this frequency the inclination signal might not be captured entirely. Those observations should motivate better model identification even without access to expensive equipment.

Monitoring shores through seasons Scanning shore elements from an ASV brings back the generic problem of reflections in foreground. In our current configuration, the sensor is placed at approximately 1 m above the waterline. This brings the advantage that most of the beams hitting the water are not reflected back to the sensor due to a low incidence angle. In rare cases, few points will be measured when a wave will have the appropriate angle and timing to generate a plane that is almost perpendicular to the laser beam. The problem gets more apparent at the interface between the shore and the water, where the beams that were reflected from the water surface hit the shore. Figure 16 shows this situation where the reflected points have a lower intensity than their direct counterparts. The reflection is even more apparent on the right side of the figure, where the terrain is expected to continue with the same slope in the water, but abruptly changes direction, due to the reflection on the water surface.

An easy way to fix this situation is to remove all points below the water level before sending them to the mapping module. This plane is defined in global coordinates, which highlights the utility of an accurate state estimation dealing with 3D attitude. Yet, the global elevation of the platform cannot be estimated only with GPS data and is set to an initial value relying on external information. We could observe the water level of the lake varying by up to 0.8 m during the year, as depicted in Figure 17. The water level of the lake was measured from a meteorological station situated on the east shore, at 3.5 km from the boathouse. We could estimate the average water level using ICP corrections when moving close to the boathouse. Its roof served as a global anchor point and provided hor-

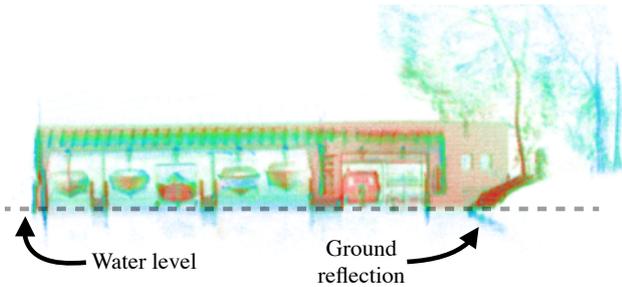


Figure 16: Front view of the boathouse. The dashed line represents the approximate position of the water level. The color of the point represents the laser intensity, red being high and blue low.

horizontal constraints. However, the fact that the boats were moored on wooden pillars, but floated on the water complexifies the problem. Given that the boats follow the water level, they will tend to compete with the surrounding environment because the registration will bring them back to the last observed position if not appropriately flagged as dynamic elements. For more complex deployments of mobile surveying units, one should keep in mind that the water level will need to be estimated to achieve a coherent representation of the scene throughout a year.

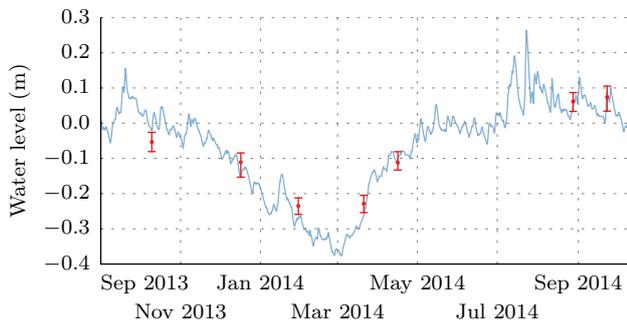


Figure 17: Comparison of the water level from the meteorological station (blue) and ICP-based estimation (red). The whiskers represent the quartiles and the dots the median values.

The water level could only be estimated close to the boathouse because this was the only place where there was enough points on horizontal structures. Indeed, shorelines close to human activities can be heavily transformed and, in our case, lead to vertical building facades and brick walls higher than the laser position. This problem prevents the use of automatic registration algorithms to compare our mapping results to a ground survey unit, for which horizontal structures could not be observed by both systems (recall Figure 13). Given those observations, one should foresee such challenges before deploying ground and water survey systems in urban environments. Another reason explaining why the correction based on point clouds was noisier outside the boathouse entry, is that the number of measured

points in each swipe is considerably reduced when the boat is facing the open water and could only partially see the large willow tree (unstructured element).

The Velodyne sensor, which we used for the attitude evaluation, might also be an appropriate choice for the application of shoreline mapping. However, it comes at a significantly higher price compared to conventional 2D laser rangefinders, and the static arrangement of lasers in vertical direction results in limited vertical resolution. For these reasons, we believe that the use of externally rotated 2D laser rangefinders is a good choice, and depending on the application, the fact that it can produce point clouds of higher density might be beneficial. Independent of cost, the range of the selected scanner is intrinsically linked to the capability of the ASV to navigate in shallow water, as shore can rapidly be out of sensor reach.

General Extensions The identification of the model parameters depends on data sets from the field. It is important that the recorded data properly span the state space of the model, i. e., it is necessary to have a homogeneous distribution of the primitive motion types in the data set. It might be beneficial to automate this process by, for instance, using feedforward-like scripts for the motor commands. Furthermore, the parameters N and T of the optimization routine have been selected manually. As they relate to physical quantities, meaningful values can be chosen by hand. However, finding optimal values might improve the optimization runtime.

Even so, we provided a relatively simple solution to acquire attitude ground truth, it implies unfavorable spatial restrictions. Such spatial restrictions can not be avoided when using laser rangefinders. An interesting alternative approach would be the utilization of an omnidirectional camera to track the motion of the horizon.

7 Conclusion

This article tackled the issue of state estimation for ASVs with exteroceptive sensors that can be affected by motion blur. We demonstrated that for such sensors, 2D state estimation is not enough because waves constantly rock the boat. Consequently, we presented three different filters for 3D state estimation based on a widely used set of basic sensors. The first one is an IMU-centric EKF inspired by state-of-the-art unmanned aerial vehicles. The second is a model-based EKF using the dynamics of our ASV. The final approach decouples the attitude estimation from the position by using a complementary filter on the attitude, and a 2D model-based EKF for position and bearing.

To compare these three state estimation algorithms, we recorded more than 13 km of data over a year. We used a theodolite to measure accurate positioning and a reflec-

tive pole coupled with an extra fast 3D scanner to assess the stability of the attitude estimation. Modeling the boat and water interaction leads to complex parameters that are hard to estimate. To cope with this problem, we also proposed an optimization procedure for the model parameters that can be applied even when only GPS position is available.

The results show that using a dynamic model of the boat is important to obtain a good and smooth estimation of linear velocity. On the other hand, IMU measurements are sufficient for the estimation of the roll and pitch angles. Therefore the best state estimation scheme combines a complementary filter for the attitude and a model-based EKF for the position and linear velocity.

We also presented a long-term series of shoreline monitoring data sets. With a classification between static and dynamic points, we highlighted environmental changes in the vegetation and the position of moored boats. Furthermore, we discussed the implications of specific challenges, such as reflections on the water surface and alternating water levels.

In conclusion, ASVs offer the opportunity to connect underwater measurements (e.g., bathymetry or temperature) with sensor data recorded above the water level. We believe that this opens a broad field of applications in the field of robotic environmental monitoring.

Acknowledgments

This work was funded by the Swiss National Science Fund (CR22I2-130023) and the EU FP7 projects NIFTi (ICT-247870) and Tradr (ICT-609763). This research was also partially funded by the Natural Sciences and Engineering Research Council of Canada (NSERC). We also thank M-E Garneau for her thorough review and the EUROPA2 project (FP7-610603) for the use of the Velodyne.

References

João Almeida, Carlos Silvestre, and Antonio Pascoal. Cooperative control of multiple surface vessels in the presence of ocean currents and parametric model uncertainty. *International Journal of Robust and Nonlinear Control*, 20(14):1549–1565, 2010.

Ali Reza Amiri-Simkooei, Christian C. J. M. Tiberius, and Peter J. G. Teunissen. Assessment of noise in GPS coordinate time series: Methodology and results. *Journal of Geophysical Research*, 112(B7):B07413, July 2007.

Marco Bibuli, Gabriele Bruzzone, Massimo Caccia, and Lionel Lapierre. Path-following Algorithms and Experiments for an Unmanned Surface Vehicle. *Journal of Field Robotics*, 26(8):669–688, 2009.

Michael Bosse and Robert Zlot. Continuous 3D scan-matching with a spinning 2D laser. In *2009 IEEE International Conference on Robotics and Automation*, pages 4312–4319, Kobe, Japan, 2009.

Jenna Brown, Chris Tuggle, Jamie MacMahan, and Ad Reniers. The use of autonomous vehicles for spatially measuring mean velocity profiles in rivers and estuaries. *Intelligent Service Robotics*, 4(4):233–244, July 2011.

Victor L. Chen, Maxim A. Batalin, William J. Kaiser, and Gaurav Sukhatme. Towards Spatial and Semantic Mapping in Aquatic Environments. In *IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 629–636, Pasadena, California, USA, 2008.

Amit Dhariwal and Gaurav S. Sukhatme. Experiments in Robotic Boat Localization. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pages 1702–1708, San Diego, California, USA, 2007.

Matthew Dunbabin, Alistair Grinham, and James Udy. An autonomous surface vehicle for water quality monitoring. In *Australasian Conference on Robotics and Automation (ACRA)*, page 13, 2009.

Matthew G. Feemster and Joel M. Esposito. Comprehensive framework for tracking control and thrust allocation for a highly overactuated autonomous surface vessel. *Journal of Field Robotics*, 28(1):80–100, 2011.

Thor I Fossen. *Handbook of Marine Craft Hydrodynamics and Motion Control*. Wiley, Chichester, 2011.

Shane Griffith, Paul Drews, and Cédric Pradalier. Towards Autonomous Lakeshore Monitoring. In *International Symposium on Experimental Robotics (ISER)*, Marakesh, 2014.

Alistair Grinham, Matthew Dunbabin, Deborah Gale, and James Udy. Quantification of ebullitive and diffusive methane release to atmosphere from a water storage. *Atmospheric Environment*, 45(39):7166–7173, 2011.

Nikolaus Hansen. The CMA Evolution Strategy: a Comparing Review. *Towards a new evolutionary computation - Advances on Estimation of Distribution Algorithms*, 102:75–102, 2006.

Gregory Hitz, François Pomerleau, Marie-Eve Garneau, Cédric Pradalier, Thomas Posch, Jakob Pernthaler, and Roland Y. Siegwart. Autonomous Inland Water Monitoring: Design and Application of a Surface Vessel. *Robotics & Automation Magazine, IEEE*, 19(1):62–72, 2012.

Gregory Hitz, Alkis Gotovos, François Pomerleau, Marie-Eve Garneau, Cédric Pradalier, Andreas Krause, and Roland Y. Siegwart. Fully autonomous focused exploration for robotic environmental monitoring. In

- Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 2658–2664, Hong Kong, China, 2014a.
- Gregory Hitz, François Pomerleau, Francis Colas, and Roland Siegwart. State Estimation for Shore Monitoring Using an Autonomous Surface Vessel. In *International Symposium on Experimental Robotics (ISER)*, Marakesh, 2014b.
- Geoffrey A. Hollinger and Gaurav S. Sukhatme. Sampling-based robotic information gathering algorithms. *The International Journal of Robotics Research*, 33(9):1271–1287, 2014.
- Vladimír Kubelka, Lorenz Oswald, François Pomerleau, Francis Colas, Tomáš Svoboda, and Michal Reinštein. Robust Data Fusion of Multi-modal Sensory Information for Mobile Robots. *Journal of Field Robotics*, 2014. To appear.
- Stefan Leutenegger and Roland Y. Siegwart. A Low-Cost and Fail-Safe Inertial Navigation System for Airplanes. In *IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 612–618, St. Paul, Minnesota, USA, 2012.
- Robert Mahony, Tarek Hamel, and JM Pflimlin. Nonlinear complementary filters on the special orthogonal group. *Automatic Control, IEEE Transactions on*, 53(5):1203–1218, 2008.
- Mark Maimone, Y Cheng, and L Matthies. Two Years of Visual Odometry on the Mars Exploration Rovers. *Journal of Field Robotics*, 24(3):169–186, 2007.
- Nathan Michael, Shaojie Shen, Kartik Mohta, Yash Mullaonkar, Vijay Kumar, Keiji Nagatani, Yoshito Okada, Seiga Kiribayashi, Kazuki Otake, Kazuya Yoshida, Kazunori Ohno, Eijiro Takeuchi, and Satoshi Tadokoro. Collaborative mapping of an earthquake-damaged building via ground and aerial robots. *Journal of Field Robotics*, 29(5), September 2012.
- Nikola Mišković, Zoran Vukić, Marco Bibuli, Gabriele Bruzzone, and Massimo Caccia. Fast In-Field Identification of Unmanned Marine Vehicles. *Journal of Field Robotics*, 28(1):101–120, 2011.
- Mark Neal, Tom Blanchard, Alun Hubbard, Nolwenn Chauché, Richard Bates, and John Woodward. A Hardware Proof of Concept for a Remote-Controlled Glacier-Surveying Boat. *Journal of Field Robotics*, 29(6):880–890, 2012.
- François Pomerleau, Philipp Krüsi, Francis Colas, Paul Furgale, and Roland Y. Siegwart. Long-term 3D map maintenance in dynamic environments. In *IEEE Int. Conf. on Robotics and Automation (ICRA)*, pages 3712–3719, Hong Kong, China, 2014.
- François Pomerleau, Francis Colas, Roland Siegwart, and Stéphane Magnenat. Comparing ICP Variants on Real-World Data Sets. *Autonomous Robots*, 34(3):133–148, 2013.
- Sebastian Scherer, Joern Rehder, Supreeth Achar, Hugh Cover, Andrew Chambers, Stephen Nuske, and Sanjiv Singh. River mapping from a flying robot: state estimation, river detection, and obstacle mapping. *Autonomous Robots*, 33(1-2):189–214, 2012.
- Anbumani Subramanian, Xiaojin Gong, J.N. Riggins, D.J. Stilwell, and C.L. Wyatt. Shoreline mapping using an omni-directional camera for autonomous surface vehicle applications. In *OCEANS 2006*, pages 1–6. IEEE, 2006.
- Nikolas Trawny and Stergios I Roumeliotis. Indirect Kalman Filter for 3D Attitude Estimation A Tutorial for Quaternion Algebra. Technical Report 612, University of Minnesota, Department of Computing Science and Engineering, 2005.
- Stephan Weiss, Markus W Achtelik, Simon Lynen, Michael C Achtelik, Laurent Kneip, Margarita Chli, and Roland Siegwart. Monocular Vision for Long-term Micro Aerial Vehicle State Estimation: A Compendium. *Journal of Field Robotics*, 30(5):803–831, 2013.