Railway Localization with Magnetic Field Distortions: Calibration, Sensor Fusion, and Evaluation

Benjamin Siebler, Andreas Lehner, Benedikt Merk, and Stephan Sand, *German Aerospace Center (DLR)*
Uwe D. Hanebeck, *Karlsruhe Institute of Technology (KIT)*

BIOGRAPHY

Benjamin Siebler received his M.Sc. in electrical engineering and information technology from the KIT in 2014 and is currently pursuing his Ph.D. at the Intelligent Sensor-Actuator-Systems Laboratory of the KIT. Since 2014 he is a researcher at the DLR Institute of Communications and Navigation with focus on multisensor and magnetic field-based localization.

Andreas Lehner received his Ph.D. in electrical engineering from University Erlangen-Nuremberg in 2008. Since 2002 he is a researcher at DLR focusing on GNSS multipath characterization, robust multi-sensor localization, and decentralized wireless communication for land-mobile applications. Andreas has authored and co-authored more than 100 publications and co-founded Intelligence on Wheels GmbH in 2012.

Benedikt Merk received his M.Sc. degree in electrical engineering and information technology from Ulm University in 2023. He is a junior researcher at the DLR Institute of Communications and Navigation since 2024. His current focus of research is on train localization with magnetic field sensors.

Stephan Sand received his Ph.D. from ETH Zurich, Switzerland in 2010. Since 2002 Stephan has researched wireless communications and multi-sensor navigation at DLR. Since 2014 he is leading the Vehicular Applications Group researching robust navigation and wireless communications technologies for terrestrial applications. Stephan has authored and co-authored more than 100 publications.

Uwe D. Hanebeck is full professor of Computer Science at KIT and director of the Intelligent Sensor-Actuator-Systems Laboratory (ISAS). He obtained his Ph.D. degree in 1997 and his habilitation degree in 2003, both in Electrical Engineering from the Technical University in Munich, Germany. His research interests are in information fusion, nonlinear state estimation, stochastic modeling, system identification, and control. He is author and coauthor of more than 600 publications and an IEEE Fellow.

ABSTRACT

This paper shows how a train can be localized with meter-level accuracy using only speed and magnetic field measurements. The approach presented in the paper uses position-dependent distortions in the magnetic field of railway tracks. These distortions are a result of the interaction between the Earth magnetic field and magnetic material in the railway tracks and in their surroundings. In particular, the paper proposes a Kalman filter that fuses wheel encoder speed measurements with position estimates from a magnetic field-based simultaneous localization and calibration algorithm. The proposed filter is evaluated with real train measurements recorded with the advanced TrainLab of Deutsche Bahn during multiple runs on a track network in Berlin. The evaluation results demonstrate the feasibility of the magnetic field-based localization system and show that a position RMSE below $0.5\,\mathrm{m}$ is achievable.

I. INTRODUCTION

Accurate and reliable train localization is a requirement for the automation of railway traffic. The automation comes with several benefits, e.g., the capacity of existing tracks can be increased by decreasing the headway between consecutive trains on the same track. In the railway domain, current localization systems heavily rely on dedicated track-side equipment. In the case of the European train control system (ETCS) the localization uses so called Eurobalises which are essentially RF transponders that send a telegram to trains driving over them. From the received telegram the train then can deduce its current position. Between two balises the localization is based on dead reckoning with an on-board odometry. In ETCS the odometry performance requirements are defined in SUBSET-041 “Performance Requirements for Interoperability” [ERA, 2015] and the position error must not exceed the limit $\pm(5\,\mathrm{m} + 5\,\% \cdot d)$, where d is the traveled distance since the last balise. This means that in the worst case scenario and a balise spacing of $1\,\mathrm{km}$ errors up to $55\,\mathrm{m}$ are to be expected. Thus, in the ETCS framework a meter-level accurate localization requires a balise roughly every $100\,\mathrm{m}$. This makes accurate localization with ETCS costly both in time and money since the balises have to be purchased, installed, and maintained. Furthermore, relying heavily on dedicated track-side infrastructure makes the whole system prone to vandalism and theft. To avoid the shortcomings of track-side infrastructure-based localization, different research initiatives and projects investigate the potential use of global navigation satellite systems (GNSS) in the railway domain. For example, the R2DATO project, part of the Europe’s Rail Joint Undertaking (ERJU), has a work package dedicated to the development of absolute safe train positioning (ASTP) demonstrators that use different sensor modalities including GNSS

[ERJU, 2026]. While GNSS will be able to address the localization problem in parts of a track network, there will be always parts in which GNSS is inaccurate or unavailable, e.g., in tunnels and urban canyons. Furthermore, GNSS is susceptible to jamming and spoofing. When GNSS is not available, the localization system has to fallback to using track-side infrastructure components or other sensor modalities have to be used. This paper follows the latter approach and proposes to use magnetometer measurements for absolute localization.

The idea behind magnetic field-based localization is to use distortions in the Earth magnetic field introduced by magnetic material in the railway environment. These distortions are static as long as the magnetic material, mostly steel in the railway case, is not moving. The static distortions show a strong dependency on the along-track position and therefore can serve as a source of absolute position information. In order to use this information for localization, a map of the magnetic field has to be created that relates the along-track position to the distorted magnetic vector field. Once the map is created, the position can be estimated by comparing the measurements of a train-mounted magnetometer to the map. However, the comparison to the map only works when the magnetometer is calibrated or when the map is recorded with the same uncalibrated sensor. While this sounds like a nuisance, in practice this poses a major problem since commonly used calibration methods require a rotation of the magnetometer and the platform on which it is mounted in a homogeneous field. For a train that weighs several tons, this is hardly feasible. In [Siebler et al., 2021, Siebler et al., 2023] and [Siebler et al., 2025] we therefore proposed a Bayesian filter and a batch estimator for simultaneous localization and calibration (SLAC) that estimate the magnetometer calibration parameters online and that can use a magnetic field map recorded with an uncalibrated sensor. The use of the magnetic field is not limited to railways but can deliver accurate position estimates also for pedestrians [Haverinen and Kemppainen, 2009, Frassl et al., 2013, Solin et al., 2016], road vehicles [Shockley and Raquet, 2014], and even airplanes [Canciani and Raquet, 2017].

In this paper, we incorporate our SLAC batch estimator from [Siebler et al., 2025] into a linear Kalman filter. The filter fuses the SLAC along-track position estimates with the speed measurements of a wheel encoder to ensure a highly available localization solution. Furthermore, the filter is combined with a two-stage fault detection and exclusion (FDE) scheme that reliably rejects the outliers in the SLAC position estimates before they are handed to the filter update step. The proposed approach will be evaluated with real train measurements recorded with the advanced TrainLab (ATL) of Deutsche Bahn during multiple runs on a track network in Berlin.

II. MAGNETIC FIELD-BASED LOCALIZATION

Magnetic localization uses position-dependent distortions of the magnetic field along railway tracks and is in essence a fingerprinting approach. Position estimation requires a "database" of fingerprints to which measurements of a train-mounted magnetometer are compared to estimate the most likely train position. In the following it is assumed that the fingerprint "database" is given in the form of a magnetic field map. The map here is a function of the along-track position that returns for each position the magnetic vector field. For the comparison between the map and the magnetometer measurements different approaches are possible dependent on the set of available measurements, e.g., if only the magnetometer measurements are available the particle filter based approach from [Siebler et al., 2023] showed promising results. If in addition to the magnetometer an odometry sensor is available the particle filter can be replaced by a batch estimator and a simple linear Kalman filter. The latter combination has the advantage that the accuracy of prior information required for a cold start of the localization system is considerably reduced. The batch estimator only needs kilometer-level accuracy for its prior, while the particle filter solution needs a prior in the order of meters.

1. Position Estimation with Maximum Likelihood

To explain the idea behind the batch SLAC algorithm we start with a simple example that introduces the idea of maximum likelihood (ML) position estimation based on a single magnetometer measurement $\mathbf{z}^{\text{mag}}$ under the assumption of a calibrated magnetometer and a given magnetic map. For a calibrated magnetometer the measurement model is

$$\mathbf{z}_k^{\text{mag}} = \mathbf{m}(s_k) + \mathbf{n}_k^{\text{mag}} \,, \tag{1}$$

with time step k, the along-track position s, magnetic map $\mathbf{m}(\cdot)$, and the Gaussian sensor noise $\mathbf{n}^{\text{mag}} \sim \mathcal{N}(0, \mathbf{I}_{3\times3}\sigma_{\text{mag}}^2)$. The along-track position is one-dimensional and limited to the interval $s \in [0, L]$ with $L \in \mathbb{R}^+$ being the track length. The map is a simple function that maps from $[0, L]$ to the magnetic vector field along the track in $\mathbb{R}^3$. Based on the measurement model and the noise density, the likelihood is given by

$$\text{p}(\mathbf{z}_k^{\text{mag}}|s_k) = \mathcal{N}(\mathbf{m}(s_k), \mathbf{I}_{3\times3}\sigma_{\text{mag}}^2) \,. \tag{2}$$

In this simple example the ML estimate $\hat{s}_k$ of the along-track position s_k at time step k is obtained from

$$\hat{s}_k = \arg\max_{s_k} p(\mathbf{z}_k^{\text{mag}}|s_k) \,. \tag{3}$$

The ML estimate therefore is a one-dimensional maximization problem that can be easily solved even for kilometer long tracks by just evaluating the likelihood at a dense position grid, e.g., in the evaluation a $0.1\,\mathrm{m}$ spacing between two evaluated positions is used. Unfortunately, in practice things are more complicated and the likelihood of a single magnetometer measurement will have several local maxima due to ambiguities in the magnetic field. Furthermore, the concrete estimate can strongly depend on the noise realization, rendering the ML estimate (3) unsuitable for accurate train localization. In practice, this problem can be resolved by using a sequence of measurements that contains not only the magnetic field at a single position but at least the field recorded on a short part of the track. To obtain such a sequence, either a magnetometer array with sensors distributed along the train is required or the train has to move during the measurements when a single magnetometer is used. Here we only look at the latter case of single magnetometer since mounting a magnetometer array inside a train might be difficult to realize in practice especially if it has to be retrofitted to an already existing train. For the case of a single magnetometer on a moving train, the ML estimation has now N magnetometer measurements available but since the train moves also the N positions of the measurements have to be estimated. For the sequence $\mathbf{z}^{\text{mag}}_{k,N} = \mathbf{z}^{\text{mag}}_{k-(N-1)}, \ldots, \mathbf{z}^{\text{mag}}_{k}$ of N magnetometer measurements the ML estimate is given by

$$\hat{s}_{k,N} = \arg\max_{s_{k,N}} p(\mathbf{z}^{\text{mag}}_{k,N} | s_{k,N}) = \arg\max_{s_{1:N}} \prod_{i=k-(N-1)}^{k} p(\mathbf{z}^{\text{mag}}_{i} | s_i) \; , \tag{4}$$

where $s_{k,N}$ is the sequence of the N along-track positions and where we assume that the different magnetometer measurements are independent given the train positions. This problem is not that simple to solve anymore with brute force since even for small N the number of sequences $s_{k,N}$ for which the likelihood has to be evaluated becomes quickly unfeasible in a practical setting. Fortunately, this problem can be easily addressed by incorporating the measurements of an odometer. Modern trains typically are equipped with wheel encoders, Doppler radars, or a combination of both. These sensors provide accurate speed measurements and thus can be used to get a measurement of the distance the train has traveled between two magnetometer measurements. This is exploited in the next step to reduce the number of estimated along-track potion in (4) from N to one. This is done by relating the positions in $s_{k,N}$ to the position s_k of the magnetometer measurement at the current time step k

$$s_i = s_k - d_i, i \in (k-(N-1), \ldots, k) \; , \tag{5}$$

where d_i is the traveled distance obtained from integrating the measured train speed between the time steps i and k. For ML estimation it is assumed that the sequence of distances $d_{k,N} = d_{k-(N-1)}, \ldots, d_k$ is either accurately obtained from the speed measurements or is set to zero in the case of d_k. Thus, (4) can be reformulated to

$$\hat{s}_{k} = \arg\max_{s_{k}} p(\mathbf{z}^{\text{mag}}_{k,N} | s_{k}; d_{k,N}) = \arg\max_{s_{k}} \prod_{i=k-(N-1)}^{k} p(\mathbf{z}^{\text{mag}}_{i} | s_k - d_i) \; , \tag{6}$$

where the semicolon in the likelihood indicates that $d_{k,N}$ is treated as a known deterministic parameter, which is not part of the optimized variables. In essence, we assume here that the measurements are obtained from an virtual array of calibrated magnetometers with known sensor distances. This is a simplification because $d_{k,N}$ is obtained from measurements and therefore inherently contains random errors. However, one could argue that the simplification is valid as long as the error does not severely change the magnetic field returned from the magnetic map. For short measurement sequences covering $< 100\,\mathrm{m}$ of track this assumption is roughly met due to the spatial frequency of the magnetic field in the order of multiple meters. With the distances $d_{k,N}$ the optimization is again reduced to a one-dimensional optimization problem. As for the single measurement case (3), the optimization can be solved by evaluating the likelihood along a dense position grid along the track.

2. Simultaneous Localization and Calibration (SLAC)

Up to now only calibrated magnetometers were considered. In practice this assumption is rarely fulfilled since the train is made of steel and potentially other magnetic material that effects the magnetometer measurements. Thus, even if a magnetometer is accurately calibrated before it is mounted inside a train it will be no longer calibrated once it is mounted. In our prior work [Siebler et al., 2025] we therefore proposed a ML batch estimator that extends the ML estimation from (6) with the possibility for simultaneous calibration. The approach from [Siebler et al., 2025] is based on the linear sensor model

$$\mathbf{z}^{\text{mag}}_{k} = \mathbf{C}\,\mathbf{m}(s_k) + \mathbf{b} + \mathbf{n}^{\text{mag}}_{k} \; , \tag{7}$$

commonly used for magnetometer calibration, cf. [Kok and Schön, 2016, Renaudin et al., 2010]. In (7), the matrix $\mathbf{C} \in \mathbb{R}^{3\times 3}$ accounts for soft-iron effects and $\mathbf{b}$ for biases caused, e.g., by permanent magnetized material inside the train. For calibration it is useful to reformulate (7) in terms of a single parameter vector $\boldsymbol{\theta} \in \mathbb{R}^{12}$ that contains the nine elements of $\mathbf{C}$ and the three elements of $\mathbf{b}$

$$\boldsymbol{\theta} = \left[[\mathbf{C}]_{1:} \quad [\mathbf{b}]_1 \quad [\mathbf{C}]_{2:} \quad [\mathbf{b}]_2 \quad [\mathbf{C}]_{3:} \quad [\mathbf{b}]_3 \right]^{\top} \; , \tag{8}$$

with the i-th row $[\mathbf{C}]_{i:}$ of $\mathbf{C}$ and the i-th element $[\mathbf{b}]_i$ of $\mathbf{b}$. With parameter vector $\boldsymbol{\theta}$ sensor model (7) becomes

$$\mathbf{z}_k^{\text{mag}} = \mathbf{H}(s_k)\boldsymbol{\theta} + \mathbf{n}_k^{\text{mag}} \ , \tag{9}$$

where the matrix valued function $\mathbf{H}(\cdot)$ contains the magnetic map evaluated at the current magnetometer position

$$\mathbf{H}(s_k) = \begin{bmatrix} \begin{bmatrix}\mathbf{m}^\top(s_k) & 1\end{bmatrix} & \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} \\ \mathbf{0}_{1\times4} & \begin{bmatrix}\mathbf{m}^\top(s_k) & 1\end{bmatrix} & \mathbf{0}_{1\times4} \\ \mathbf{0}_{1\times4} & \mathbf{0}_{1\times4} & \begin{bmatrix}\mathbf{m}^\top(s_k) & 1\end{bmatrix}^\top \end{bmatrix}. \tag{10}$$

For known positions the calibration is trivial and can be solved with with a linear least squares (LS) estimator, which in this particular case is equal to the ML estimate. To do this one stacks N measurement equations and solves for parameter vector $\boldsymbol{\theta}$

$$\hat{\boldsymbol{\theta}}_{\text{LS}}(s_{k,N}) = \left(\mathbf{H}(s_{k,N})^\top \mathbf{H}(s_{k,N})\right)^{-1} \mathbf{H}(s_{k,N})^\top \mathbf{z}_{k,N}^{\text{mag}} \ , \tag{11}$$

where $\mathbf{H}(s_{k,N})$ contains the stacked matrices $\mathbf{H}(\cdot)$ at the different positions in $s_{k,N}$ and $\mathbf{z}_{k,N}^{\text{mag}}$ is interpreted here as a column vector in which the N magnetometer measurements are stacked on top of each other.

For SLAC now the position estimation (6) and the calibration parameter estimation (11) must be combined into a single optimization problem. The naive approach would be to optimize of the joint space s_k and $\boldsymbol{\theta}$

$$\hat{s}_k, \hat{\boldsymbol{\theta}} = \underset{s_k, \boldsymbol{\theta}}{\arg\max} \ \mathrm{p}(\mathbf{z}_{k,N}^{\text{mag}} | s_k, \boldsymbol{\theta}; d_{k,N}) \ . \tag{12}$$

This leads to a optimization problem over an 13-dimensional space. Furthermore, the problem in non-convex due to the nonlinear relation of the position and the magnetic vector field. Without prior knowledge a naive approach would get certainly stuck in a local minimum more often than not. In [Siebler et al., 2025] we therefore proposed to exploit the closed form solution (11) for the ML estimate of $\boldsymbol{\theta}$. With (11) the optimization problem (12) is divided into two nested separate optimizations

$$\hat{s}_k, \hat{\boldsymbol{\theta}} = \underset{s_k}{\arg\max} \left[\underset{\boldsymbol{\theta}}{\arg\max} \ \mathrm{p}(\mathbf{z}_{k,N}^{\text{mag}} | s_k, \boldsymbol{\theta}; d_{k,N}) \right] \ , \tag{13}$$

where the inner optimization w.r.t. to $\boldsymbol{\theta}$ has the closed form solution given by (11). Thus, (13) can be replaced with

$$\hat{s}_k = \underset{s_k}{\arg\max} \ \mathrm{p}(\mathbf{z}_{k,N}^{\text{mag}} | s_k, \hat{\boldsymbol{\theta}}_{\text{LS}}(s_k); d_{k,N}) \ . \tag{14}$$

In (14), the parameter vector estimate $\hat{\boldsymbol{\theta}}_{\text{LS}}(s_k)$ only depends on s_k since it is assumed that the remaining positions in $s_{k,N}$ can be once more obtained from the train speed measurements $d_{k,N}$. Consequentially, (14) is again a one-dimensional optimization problem. In contrast to the localization with a calibrated magnetometer, here every time the likelihood is evaluated also the LS estimate $\hat{\boldsymbol{\theta}}_{\text{LS}}(s_k)$ has to be obtained. Fortunately, this can be done efficiently by applying a few simply tricks that allow the offline computation of the pseudo inverse of $\mathbf{H}(s_{k,N})$ in (11). Therefore, estimating $\hat{\boldsymbol{\theta}}_{\text{LS}}(s_k)$ is reduced to a matrix vector multiplication. With this simplification it is possible to perform a global optimization for (14) by evaluating the likelihood on a dense position grid even for kilometer long tracks, as will be discussed also in the evaluation section. The exact details about the efficient implementation of the SLAC algorithm can be found in [Siebler et al., 2025].

III. SENSOR FUSION ALGORITHM AND OUTLIER DETECTION

The magnetic field-based localization approach introduced in the previous section provides only snapshot estimates and does not apply any filtering. To improve the overall localization solution, this section introduces a filter algorithm that fuses SLAC estimates with train speed measurements. Furthermore, a two-stage FDE scheme is proposed that reliably detects outliers in the SLAC estimates.

1. Filter Algorithm

The proposed filter is updated with SLAC along-track position estimates $\hat{s}_k$ and train speed measurements $\tilde{\dot{s}}_k$. To keep the filter as simple as possible, a linear two-dimensional motion model with state vector

$$\mathbf{x}_k = [s_k \quad \dot{s}_k]^\top \tag{15}$$

is used that contains the one-dimensional along-track train position s and its time derivative $\dot{s}$, i.e. the train speed. The temporal dynamics of the state vector are described by the linear model

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k \ , \tag{16}$$

with process noise $\mathbf{w}_k$. The process noise $\mathbf{w}_k$ is assumed to be white, zero mean, and Gaussian with a covariance matrix according to a discrete white noise acceleration model

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{4}T^4 & \frac{1}{2}T^3 \\ \frac{1}{2}T^3 & T^2 \end{bmatrix} \sigma_{\mathbf{w}}^2 \ , \tag{17}$$

where the standard deviation $\sigma_{\mathbf{w}}$ is a tuning parameter and should be in the range of the maximum acceleration of the train [Bar-Shalom et al., 2002]. The measurement model is trivial here since we have direct observations of both states, the position s and the speed $\dot{s}$. Therefore, the measurements are given by the linear equations

$$\hat{s}_k = [1 \quad 0]\, \mathbf{x}_k + n_k^{\mathrm{SLAC}} \quad \text{and} \quad \tilde{\dot{s}}_k = [0 \quad 1]\, \mathbf{x}_k + n_k^{\mathrm{odo}} \ . \tag{18}$$

Similar to the process noise term, the measurement noise is assumed to be Gaussian and zero mean with the distributions $n_k^{\mathrm{SLAC}} \sim \mathcal{N}(0, \sigma_{\mathrm{SLAC}}^2)$ and $n_k^{\mathrm{odo}} \sim \mathcal{N}(0, \sigma_{\mathrm{odo}}^2)$. The standard deviation of both noise terms is treated as a tuning parameter to get a conservative covariance for the estimated state variables. On the basis of (15)–(18) the well known Kalman filter equations are used to predict and update the state estimate.

2. Fault Detection and Exclusion

The quality of the obtained filter solution strongly depends on the quality of the measurements used in the update step. For example, undetected outliers have a strong negative effect on the filtered solution, thus a reliable outlier detection is essential. This is especially true for the SLAC batch estimator used in this paper because its estimates can suffer from outliers with errors in the range of kilometers. To reliably detect outliers in the SLAC estimates, the proposed localization system uses a two-stage FDE scheme. The first stage performs outlier detection based only on the speed measurements and the SLAC estimates. The idea of the first stage is that consecutive SLAC estimates must be self-consistent if the movement of the train between them is corrected accordingly. To correct the movement of the train between consecutive SLAC estimates, the last M SLAC estimates $\hat{s}$ starting with the newest estimate at time step k are stored in a buffer alongside all speed measurements $\tilde{\dot{s}}$ recorded between the time steps k and $k-(M-1)$. Based on the values in the buffer, all SLAC estimates are corrected for the train movement by integrating the speed measurements between their respective time step and a reference time step. We selected time step $k-(M-1)$ of the oldest estimate in the buffer as reference, but the actual choice does not matter, e.g., we could also pick the current time step k as reference. After applying the corrections, the M corrected SLAC estimates are collected in the set $\mathcal{S} = \{\hat{s}_{k-(M-1)}^{\mathrm{corr}}, \ldots, \hat{s}_k^{\mathrm{corr}}\}$.

In the ideal case, all corrected estimates in $\mathcal{S}$ have the same value and their sample variance $\mathrm{Var}(\mathcal{S})$ should be zero. In practice this will never be fulfilled due to the limited accuracy of the SLAC estimates, the noise in the speed measurements, and the numerical integration of the speed to obtain the travel distance of the train. However, for accurate SLAC estimates without outliers we still can expect that all estimates are close to each other with a small variance $\mathrm{Var}(\mathcal{S})$. In our experiments this method works reliably and can be used also as a standalone FDE scheme, but using only this scheme comes with a trade-off between finding all outliers and the amount of false alarms raised. This is not surprising since this kind of trade-off emerges in most FDE schemes. Here in particular the trade-off comes down to selecting a concrete value for the buffer length M. We found that the SLAC outliers are fairly random with large errors and thus already short buffer lengths are sufficient to find most outliers. Unfortunately, there are also some cases in which the outliers are highly correlated and longer buffers are required such that at least one estimate is contained that is not consistent with the correlated outliers. In the evaluation, we decided to use only a short buffer with $M=3$. The first FDE stage thus cannot capture all outliers and has to be combined with a second stage of FDE.

The second FDE stage utilizes the innovation of the Kalman filter and its Mahalanobis distance to find outliers in the SLAC estimate. This is a common approach to detect outliers in the update step of a Kalman filter and is easily computed based on the predicted state vector and its covariance. Here only an outlier detection on the position state s_k is performed since we don't expect large outliers in the speed measurements, at least we didn't observed any during our experiments. As test statistic for the second FDE stage we simply calculate the Mahalanobis distance

$$d_k^{\mathrm{maha}} = \sqrt{\frac{(\hat{s}_k - s_k^-)^2}{[\mathbf{P}_k^-]_{1,1} + \sigma_{\mathrm{SLAC}}^2}} \ , \tag{19}$$

where s_k^- is the predicted position at time step k and $[\mathbf{P}_k^-]_{1,1}$ is the corresponding element of the predicted state covariance matrix $\mathbf{P}_k^-$. Once d_k^{maha} is calculated it is compared to a threshold to detect the SLAC outliers. If all the assumptions of the Kalman filter

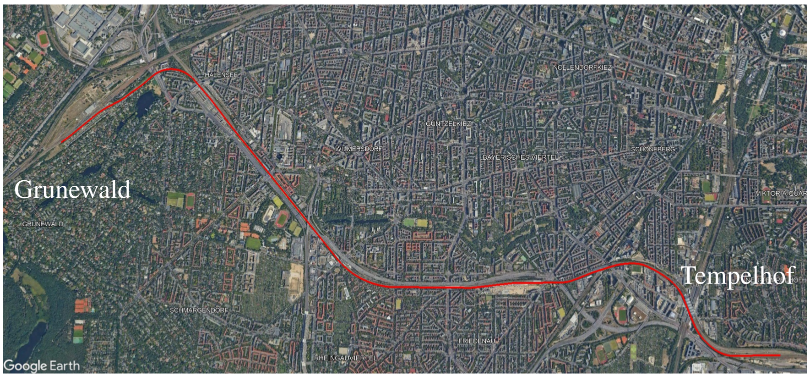

Figure 1: Test tracks between Berlin Grunewald and Tempelhof. Source: Google Earth.

Figure 2: (left) ATL on a track in Berlin. (right) Magnetometer array with five elements mounted below a wagon of the ATL. One of the magnetometers (green circle) is used only to create the magnetic map and another sensor (red circle) is used exclusively for localization.

are fulfilled, e.g. all noise terms are uncorrelated and Gaussian, the threshold can be selected theoretically based on same desired properties of the fault detection. Here these assumptions are not perfectly fulfilled since the errors on the SLAC estimates cannot be considered Gaussian. Thus, the threshold in the evaluation was selected empirically such that all outliers are detected.

IV. EVALUATION

In this section, the proposed localization system consisting of the combination of batch SLAC, Kalman filter, and FDE is evaluated on the basis of measurements recorded with the ATL of Deutsche Bahn during multiple runs on a track network in Berlin.

1. Measurement Campaign

The measurements were conducted in 2021 during two days of measurements in the city of Berlin between the neighborhoods Grunewald and Tempelhof. During the measurements the train was taking different routes between Tempelhof an Grunewald, an example run is shown in Fig. 1. For the evaluation we look at a total of 23 runs, where 4 runs are used for mapping and the remaining 19 are used to evaluate the proposed localization system. The average length of each run was around $9.4\,\mathrm{km}$ during which the ATL was driving with an average speed of $8.26\,\mathrm{m\,s^{-1}}$.

During all runs the ATL was recording multiple magnetometers, GNSS receivers, and a wheel encoder. For the evaluation two low cost Kionix KMX62 magnetometers with a measurement rate of $200\,\mathrm{Hz}$ are used. Both sensors are mounted below the train as shown in Fig. 2. One of the sensors was used exclusively to create the magnetic map and the other one exclusively to test the localization system. Note, both magnetometers are not calibrated and therefore also the map is not calibrated. For the proposed SLAC algorithm this is not an issue because in the railway domain the linear magnetometer model (7) is still valid even if the map is not calibrated, a detailed explanation why this is the case can be found in [Siebler et al., 2023]. The resulting calibration parameters are relative and specific for the used map and applying them to the magnetometer data used for localization does not recover the true physical field dimensions. For localization purposes this is not important since the calibration parameters can be considered as mere nuisance parameters. In addition to the magnetometer measurements, the SLAC algorithm uses measurements of the axle mounted DEUTA wheel encoder which are recorded at a rate of $1\,\mathrm{Hz}$.

The error of the proposed magnetic field-based localization system is calculated w.r.t. a ground truth obtained from a postprocessed smoothed RTK solution obtained with the NovAtel Inertial Explorer. The smoothed RTK solution is matched to a geometric map of the track network to obtain a one-dimensional along-track position. The map-matched position is then passed to a Rauch–Tung–Striebel smoother, which combines the matched position with the wheel encoder speed measurements. The last

Table 1: Position error statistics after processing all 19 runs through Berlin.

Approach	RMSE	Position / m q_{95}	q_{99}	|max|
unaided odometry	4.86	10.62	13.96	14.99
Kalman filter	0.49	0.91	2.05	4.10

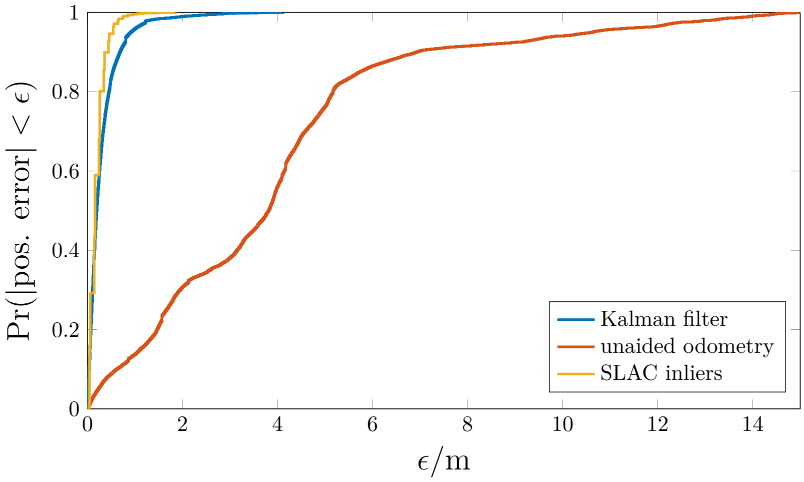

Figure 3: Absolute position error CDF for the Kalman filter output, the unaided odometry, and the batch SLAC inliers.

step became necessary because the RTK solution has larger errors and gaps when the train approaches and passes through track segments covered by road bridges. For the GNSS ground truth recording a Septentrio PolaRx5TR receiver and Antonics OmPlecs-TOP 200 AMR 1500-B L1/L2 H antennas were used.

2. Parameterization and Initialization

In order to run the SLAC algorithm and the Kalman filter, several parameters have to be set and the filter state and covariance have to be initialized. The magnetic and geometric map of the track network were created with a spatial resolution of $0.1\,\mathrm{m}$. For position estimation the SLAC algorithm uses $50\,\mathrm{m}$ long signatures with a spatial resolution of $0.3\,\mathrm{m}$. Reducing the resolution of the magnetic signature to $0.3\,\mathrm{m}$ speeds up the SLAC postion estimation and does not have a severe impact on the localization performance. The global optimization performed by SLAC compares the magnetic signature to the magnetic map for each position along the complete $\approx 9.4\,\mathrm{km}$ long route with the full map resolution of $0.1\,\mathrm{m}$. With these settings a single SLAC estimate takes less than a second on a laptop CPU. To avoid strong correlations between consecutive SLAC estimates, a new estimate is triggered when the train has moved at least $10\,\mathrm{m}$ since the last estimate. The first stage of SLAC outlier detection uses $M = 3$ corrected SLAC estimates and the threshold for their standard deviation is set to $0.7\,\mathrm{m}$. For the second stage based on the Mahalanobis distance the threshold is set to $2\,\mathrm{m}$.

The Kalman filter position state is initialized with the ground truth position and the speed state is set to zero since all runs start from standstill. The initial covariance is set to $\mathrm{diag}([1\,\mathrm{m}^2, (0.15\,\mathrm{m\,s^{-1}})^2])$ and the process noise standard deviation is set to $\sigma_{\mathbf{w}} = 1\,\mathrm{m\,s^{-2}}$. The measurement noise standard deviations were set empirically to $\sigma_{\mathrm{SLAC}} = 1\,\mathrm{m}$ and $\sigma_{\mathrm{odo}} = 0.15\,\mathrm{m\,s^{-1}}$. The filter is updated with each new wheel encoder measurement at a rate of $1\,\mathrm{Hz}$. A SLAC update is only performed when the current estimate passes both FDE stages.

3. Results

This section presents the evaluation results of the proposed magnetic field-based localization system. As a baseline to show the benefit of the proposed system, the unaided odometry positioning solution is used, i.e. the integrated wheel encoder speed measurements. The overall position error statistics after processing all 19 runs are shown in Table 1. The different columns show the root mean squared error (RMSE), $95\,\%$ quantile q_{95}, $99\,\%$ quantile q_{99}, and the maximum absolute error |max|. All statistics in Table 1 show clearly that the proposed Kalman filter with SLAC updates outperforms the unaided odometry solution. This was expected since the SLAC updates in the Kalman filter render the position state observable and therefore each update “resets” the position error resulting from integrating the measured train speed. By using the SLAC position estimates, the position RMSE can be reduced from multiple meters to only $0.49\,\mathrm{m}$. Furthermore, both position error quantiles are significantly reduced compared to the odometer localization solution. This is also confirmed by the cumulative distribution function (cdf) of the absolute position

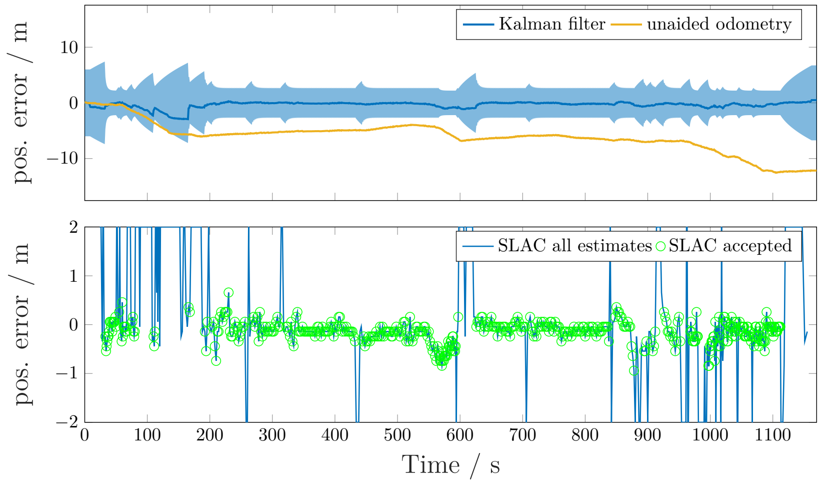

Figure 4: Position error over time for one full run from Berlin Grunewald to Tempelhof. (top) Position error of the Kalman filter (blue line). The blue shaded area indicates six times the position standard deviation estimated by the Kalman filter. As a baseline the position error of the unaided odometry (yellow line) obtained from integrating the wheel encoder speed is shown. (bottom) Position error of all SLAC estimates (blue line) and the position error of the estimates that pass the two-stage FDE and that are used to update the Kalman filter (green circles).

error shown in Fig. 3. The CDF shows, in addition to the position errors of the Kalman filter and the unaided odometry, the error of the true SLAC inliers. Here the SLAC inliers are all estimates with an error below $2\,\mathrm{m}$ and are classified based on the ground truth. In the CDF the SLAC inliers achieve the most accurate result, which might be surprising on the first glance. However, from our experience the SLAC estimates are either very accurate or completely off the true position. Assuming the proposed outlier detection works reliably, fusing the SLAC estimates and the speed measurements will result in a CDF of the Kalman filter position error somewhere between the SLAC and odometer CDFs. The exact shape of the CDF depends on the availability of the SLAC estimates. During long periods without an accepted SLAC estimate the Kalman filter error will be dominated by the wheel encoder speed measurement errors. In contrast, when SLAC estimates are continuously available the Kalman filter error will closely follow the SLAC errors. This can be nicely seen from Fig. 4 that shows the position error of the Kalman filter, the unaided odometry, and the SLAC estimates over time. For the SLAC estimates in the bottom plot the blue line shows the error for each estimate and the green circles mark the estimates that have passed the FDE and that are used to update the Kalman filter position. When SLAC estimates are not available for some time, e.g. between $112\,\mathrm{s}$–$166\,\mathrm{s}$, the Kalman filter error starts to drift in a similar way as the unaided odometry error shown in yellow. In addition to the position error, the top plot of Fig. 4 shows six times the standard deviation of the position state estimated by the Kalman filter as a blue shaded area. This shows that the proposed positioning system not only provides accurate position estimates, but can also provide a conservative uncertainty estimate. Since a constant value for the standard deviation of the SLAC estimates is used in the Kalman filter update step, the performance of the FDE algorithm is crucial to maintain a conservative uncertainty estimate. Fortunately, the proposed two-stage FDE scheme detected $100\,\%$ of the SLAC outliers during the 19 evaluated runs. As it is typically the case for outlier detectors, there is a trade-off between the reliable detection of true outliers and the number of raised false alarms. In the evaluation, the FDE parameters were chosen to reliably detect all outliers and therefore the false alarm rate was increased. However, the false alarm rate, i.e. the percentage of all inliers falsely classified as outliers, was with $9.17\,\%$ small enough such that SLAC estimates were regularly provided to the Kalman filter to update its position state estimate.

V. CONCLUSION

This paper proposed a long-term stable meter-level accurate train localization system based on the measurements of a low-cost magnetometer and a wheel encoder. The proposed system uses distortions in the magnetic field along railway tracks that can be seen as magnetic "fingerprints" that allow to identify a specific positions on a track. With a map of magnetic fingerprints the absolute train position can be estimated by comparing the measurements of train-mounted magnetometer to the mapped fingerprints. For the comparison between the map and the magnetometer measurements the SLAC algorithm from our prior work [Siebler et al., 2025] was used. To improve the overall availability and accuracy of the estimated position, a Kalman filter was proposed that combines the SLAC estimates with speed measurements of a wheel encoder. Since the SLAC estimates suffer from large outliers, a FDE scheme was developed that reliably detect outliers before they are passed to the Kalman filter update step. The feasibility and the performance of the proposed localization system was evaluated on the basis of 19 runs with the ATL on a track network in the city of Berlin. The evaluation result showed that with the proposed system a sub-meter RMSE can be achieved due to the high accuracy of the SLAC estimates. Furthermore, the maximum position error was kept below $5\,\mathrm{m}$, which can be attributed to the reliable FDE scheme that was able to detect all SLAC outliers with errors beyond $2\,\mathrm{m}$.

ACKNOWLEDGEMENTS

Co-funded by
the European Union

Europe's Rail

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European union. Neither the European Union nor the granting authority can be held responsible for them. The project FP2-R2DATO is supported by the Europe's Rail Joint Undertaking and its members.

REFERENCES

[Bar-Shalom et al., 2002] Bar-Shalom, Y., Li, X.-R., and Kirubarajan, T. (2002). *Estimation with Applications to Tracking and Navigation: Theory, Algorithms and Software*. John Wiley & Sons, Inc.

[Canciani and Raquet, 2017] Canciani, A. and Raquet, J. (2017). Airborne Magnetic Anomaly Navigation. *IEEE Transactions on Aerospace and Electronic Systems*, 53(1):67–80.

[ERA, 2015] ERA (2015). SUBSET-041 "Performance Requirements for Interoperability" Issue 3.2.0. `https://www.era.europa.eu/system/files/2023-01/sos3_index014_-_subset-041_v320.pdf`.

[ERJU, 2026] ERJU (2026). R2dato - project structure. `https://rail-research.europa.eu/pages/fp2-r2dato/structure`.

[Frassl et al., 2013] Frassl, M., Angermann, M., Lichtenstern, M., Robertson, P., Julian, B. J., and Doniec, M. (2013). Magnetic Maps of Indoor Environments for Precise Localization of Legged and Non-legged Locomotion. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 913–920.

[Haverinen and Kemppainen, 2009] Haverinen, J. and Kemppainen, A. (2009). Global indoor self-localization based on the ambient magnetic field. *Robotics and Autonomous Systems*, 57(10):1028–1035.

[Kok and Schön, 2016] Kok, M. and Schön, T. B. (2016). Magnetometer Calibration Using Inertial Sensors. *IEEE Sensors Journal*, 16(14):5679–5689.

[Renaudin et al., 2010] Renaudin, V., Afzal, M. H., and Lachapelle, G. (2010). Complete Triaxis Magnetometer Calibration in the Magnetic Domain. *Hindawi Journal of Sensors*, 2010(1):967245.

[Shockley and Raquet, 2014] Shockley, J. A. and Raquet, J. F. (2014). Navigation of Ground Vehicles Using Magnetic Field Variations. *NAVIGATION*, 61(4):237–252.

[Siebler et al., 2023] Siebler, B., Lehner, A., Sand, S., and Hanebeck, U. D. (2023). Simultaneous Localization and Calibration (SLAC) Methods for a Train-Mounted Magnetometer. *NAVIGATION: Journal of the Institute of Navigation*, 70(1).

[Siebler et al., 2025] Siebler, B., Lehner, A., Sand, S., and Hanebeck, U. D. (2025). Snapshot Estimator for Magnetic Field-Based Train Localization with Uncalibrated Magnetometers. In *2025 33rd European Signal Processing Conference (EUSIPCO)*, pages 2142–2146.

[Siebler et al., 2021] Siebler, B., Sand, S., and Hanebeck, U. D. (2021). Localization with Magnetic Field Distortions and Simultaneous Magnetometer Calibration. *IEEE Sensors Journal*, 21(3):3388–3397.

[Solin et al., 2016] Solin, A., Särkkä, S., Kannala, J., and Rahtu, E. (2016). Terrain Navigation in the Magnetic Landscape: Particle Filtering for Indoor Positioning. In *2016 European Navigation Conference (ENC)*, pages 1–9.