

Article

Target-Protected Multiscale GLRT-CUSUM Detection for Low-SNR Vector Magnetic Anomalies on Ocean Buoys

Yingdong Yang ^{1,2}, Peichuang Wang ¹, Keke Zhang ^{1,2,*}, Shixuan Liu ^{1,2}, Xiao Fu ^{1,2}, Bo Wang ¹, Qinglin Kong ², Zhijin Qiu ¹, Jiming Zhang ^{1,2} and Xianglong Yang ^{1,2}

¹ State Key Laboratory of Physical Oceanography, Institute of Oceanographic Instrumentation, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250316, China; yyd_ocean@qlu.edu.cn (Y.Y.); 10431240724@stu.qlu.edu.cn (P.W.); lxx@sdioi.com (S.L.); dawn_fu@163.com (X.F.); wangbo0532@qlu.edu.cn (B.W.); qzj@qlu.edu.cn (Z.Q.); jimingzhangwld@163.com (J.Z.); feiyehao6@sina.com (X.Y.)

² Laoshan Laboratory, Qingdao 266237, China; kongql66@163.com

* Correspondence: 18661475909@163.com

Abstract

Vector magnetic anomaly detection on ocean buoys is challenged by attitude changes, non-stationary backgrounds, and weak targets that adaptive filters may absorb. We propose an online target-protected framework combining attitude compensation, constant-false-alarm-rate normalized innovation squared (CFAR-NIS) threshold adaptation, target-dwell gating, whitening based on a target-free calibration covariance, and weighted multiscale generalized likelihood ratio test-cumulative sum (GLRT-CUSUM) fusion. The target-protection mechanism constrains covariance adaptation and the effective Kalman gain during suspected target intervals, while dual-threshold confirmation suppresses isolated background spikes. We evaluated the method using publicly available buoy attitude and triaxial magnetometer background data with injected ship magnetic signatures under semi-physical conditions. At SNR ≈ 2.7 , 200 paired Monte Carlo trials yielded a detection probability of 0.965, a background false-alarm rate of 0.045, and an average delay of 208.3 samples. Compared with standard and adaptive Kalman filtering, orthogonal-basis-function (OBF), minimum-entropy, and spectral-residual CUSUM baselines, the proposed method significantly improved detection probability and shortened delay. Its false-alarm rate was comparable to OBF and SR-CUSUM but higher than those of standard KF, adaptive KF, and minimum-entropy detection. Validation with real operational sea-trial data remains necessary.

Keywords: vector magnetic anomaly detection; ocean buoy; attitude compensation; target-protected adaptive filtering; online CFAR-NIS; multiscale GLRT-CUSUM

Academic Editor: José A. Orosa García

Received: 8 July 2026

Revised: 28 July 2026

Accepted: 3 August 2026

Published: 5 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Magnetic anomaly detection (MAD) is a passive, non-acoustic technique for underwater target monitoring that exploits magnetic field perturbations induced by ferromagnetic targets [1–4]. On ocean buoys, wave-induced attitude changes continuously alter the projection of the geomagnetic field onto the triaxial sensor coordinates and may mask weak target anomalies [5–7]. Platform motion, sensor drift, low-frequency disturbances, and triaxial correlated noise further make buoy-based vector magnetic anomaly detection

a sequential detection problem under low-SNR and non-stationary background conditions.

Existing approaches range from fixed-threshold and sliding-window energy detectors [8] to orthogonal-basis-function-based and noise-adaptive feature methods [9–12], adaptive state-space filtering [13–16], and learning-based or sensor-network approaches [17–21]. These methods provide useful signal extraction and interference suppression, but adaptive filters may absorb target-induced innovations into background statistics, whereas learning-based methods require representative labeled data and may suffer domain-shift effects across buoy sites and operating conditions. Statistical GLRT and CUSUM methods [22–26] are well suited to sequential detection in non-stationary backgrounds; however, conventional single-scale fixed-accumulation schemes do not simultaneously accommodate variations in target heading, passing distance, and response duration.

To address vector magnetic anomaly detection under real buoy background conditions, this paper proposes an interpretable statistical detection framework that integrates attitude compensation, online threshold-adaptive target-protected filtering, target-free calibration-based covariance-whitened GLRT, and dual-threshold weighted multiscale CUSUM. The main contribution of this work lies in the coupled design of the detection chain, which is specifically developed to mitigate buoy attitude coupling, threshold mismatch caused by background non-stationarity, and weak-target residual absorption by adaptive filtering.

The main contributions of this paper are summarized as follows. First, an attitude-compensated vector magnetic field observation model is established for ocean-buoy platforms, in which roll, pitch, and heading information is used to reduce residual contamination caused by geomagnetic projection variations. Second, an online CFAR-NIS target-protected adaptive filtering mechanism is developed. This mechanism estimates empirical NIS quantile thresholds in real time using an alarm-free background window and combines them with a continuous target-dwell counter to constrain covariance updating, delay the recovery of the gating weight, and reduce the effective Kalman gain during persistent suspected-target intervals. Third, a target-free calibration segment is used to estimate a calibration-based detection covariance, whiten the triaxial residuals, and construct the GLRT statistic. Fourth, a CFAR-calibrated normalized weighted multiscale evidence fusion strategy and a dual-threshold candidate-confirmation decision mechanism are proposed to improve low-SNR evidence accumulation while suppressing spike-induced false alarms.

Relative to adaptive-filtering-based magnetic tracking methods [13] and conventional GLRT/OBF detectors [22,23], the proposed method is distinguished by the coordinated use of online CFAR-NIS threshold adaptation and target-dwell memory to regulate both covariance updating and the effective Kalman gain. The standard KF and adaptive KF baselines use single-scale GLRT-CUSUM on their innovation sequences; OBF operates in an orthogonal-basis-function domain; and SR-CUSUM applies a causal spectral-residual transform to a scalar residual magnitude, followed by single-scale CUSUM accumulation [27,28]. These baselines do not jointly include target-protected filtering, whitening with a covariance estimated from a target-free calibration segment, multiscale GLRT-CUSUM fusion, and dual-threshold confirmation.

The overall workflow of the proposed ocean-buoy-based vector magnetic anomaly detection framework is illustrated in Figure 1.

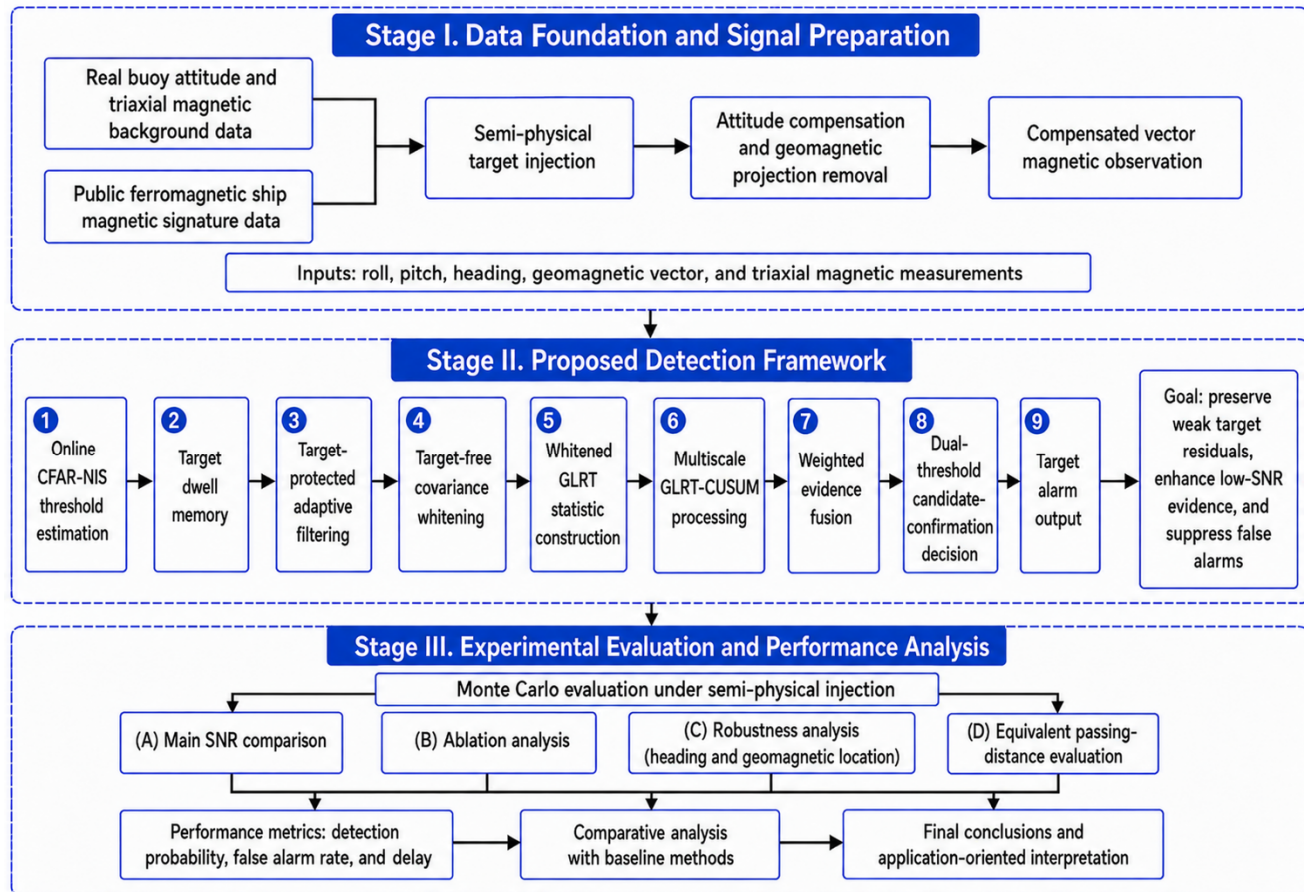


Figure 1. Overall framework of the proposed ocean-buoy-based vector magnetic anomaly detection method.

2. Materials and Methods

2.1. Attitude-Coupled Observation and Compensation Model

Let $\mathbf{y}_{\text{raw},k} \in \mathbb{R}^3$ denote the triaxial magnetic field observation in the sensor coordinate system at time k . This observation consists of the geomagnetic projection, local background, target magnetic anomaly, and measurement noise. Let roll, pitch, and heading be denoted by ϕ_k , θ_k , and ψ_k , respectively. Under the ZYX Euler-angle convention, the attitude rotation matrix from the navigation coordinate system to the sensor coordinate system is given by:

$$\mathbf{D}_k = \mathbf{R}_z(\psi_k)\mathbf{R}_y(\theta_k)\mathbf{R}_x(\phi_k) \quad (1)$$

Given the nominal geomagnetic vector $\mathbf{B}_e^n$ at the buoy location, the attitude-compensated triaxial residual signal is defined as:

$$\mathbf{y}_k = \mathbf{y}_{\text{raw},k} - \mathbf{D}_k \mathbf{B}_e^n \quad (2)$$

The compensated observation can be expressed as:

$$\mathbf{y}_k = \mathbf{b}_k + \mathbf{a}_k + \mathbf{v}_k \quad (3)$$

where $\mathbf{b}_k$ denotes slow drift, marine background, and residual platform disturbance; $\mathbf{a}_k$ is the target magnetic anomaly vector to be detected; and $\mathbf{v}_k$ is zero-mean measurement noise. The purpose of explicit attitude compensation is to remove the large-amplitude geomagnetic projection from the sensor observation, so that subsequent filtering and detection can focus on relatively small residual variations [5,6].

In the filtering model, the target-free background term $\mathbf{b}_k$ in (3) is modeled as $\mathbf{H}\mathbf{x}_k$. Therefore, the compensated observation used for recursive filtering can be written as $\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{a}_k + \mathbf{v}_k$. Under the target-free hypothesis, $\mathbf{a}_k = 0$, whereas under the target-present hypothesis, $\mathbf{a}_k$ may be nonzero. The target anomaly is not modeled as part of the slowly varying background state. Instead, the target-protection mechanism reduces the absorption of suspected target components into the background-state update, thereby helping retain target-related components in the innovation residual for subsequent GLRT-CUSUM detection.

2.2. Online Threshold-Adaptive Target-Protected Kalman Filtering

The background state is described by a low-dimensional state-space model. Let $\mathbf{x}_k$ denote the latent, target-free background state after attitude compensation, and let $\mathbf{H}$ denote the observation matrix that maps the background state to the measured magnetic vector. Its evolution is modeled as a first-order state-transition process:

In the implementation, $\mathbf{x}_k \in \mathbb{R}^6$ is a latent two-block background state rather than a directly measured physical quantity. With $\mathbf{H} = [\mathbf{I}_3 \ \mathbf{I}_3]$, the modeled background contribution is approximately $\mathbf{b}_k \approx \mathbf{H}\mathbf{x}_k$, namely the sum of two latent three-dimensional components. The two components use different process-noise levels in $\mathbf{Q}_{k-1}$ to represent background variations at different time scales and should not be interpreted as two separate sensors.

$$\mathbf{x}_k = \mathbf{F}\mathbf{x}_{k-1} + \mathbf{w}_{k-1}, \mathbf{w}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{k-1}) \quad (4)$$

where $\mathbf{F}$ is the state-transition matrix, $\mathbf{Q}_{k-1}$ is the process-noise covariance, and $\mathbf{w}_{k-1}$ represents slow background variations and unmodeled disturbances. The one-step-ahead predicted state and prediction-error covariance are calculated by standard Kalman prediction:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}\mathbf{x}_{k-1|k-1} \quad (5)$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^T + \mathbf{Q}_{k-1} \quad (6)$$

For the main target-protected implementation, $\mathbf{F} = \mathbf{I}_6$ is used. Accordingly, the implemented prediction reduces to $\mathbf{x}_{k|k-1} = \mathbf{x}_{k-1|k-1}$ and $\mathbf{P}_{k|k-1} = \mathbf{P}_{k-1|k-1} + \mathbf{Q}_{k-1}$. This is the identity-transition specialization of Equations (5) and (6).

For the reported experiments, $\mathbf{Q} = \text{diag}(0.05\mathbf{I}_3, 0.02\mathbf{I}_3)$ is kept constant.

Thus, $\mathbf{P}_{k|k-1}$ is obtained by standard Kalman prediction-error propagation rather than by an independent empirical calculation at the current sample. To avoid temporal-order ambiguity between the recursive update of the measurement noise covariance and the calculation of the innovation statistic at the current time, the filtering innovation and its prior innovation covariance are defined as

$$\mathbf{v}_k = \mathbf{y}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1}, \quad (7)$$

$$\mathbf{S}_k = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R}_{k-1}. \quad (8)$$

where $\mathbf{R}_{k-1}$ is the measurement noise covariance updated at the previous time and is used as the prior covariance for the current innovation. To evaluate the statistical consistency of the filter in real time, the normalized innovation squared (NIS) is defined as

$$\eta_k = \mathbf{v}_k^T \mathbf{S}_k^{-1} \mathbf{v}_k \quad (9)$$

Under target-free conditions with a statistically consistent filtering model, η_k approximately follows a central chi-square distribution with m degrees of freedom. For the triaxial magnetic observations used in this paper, $m = 3$, namely

$$\eta_k \sim \chi_m^2, \quad m = 3 \quad (10)$$

The original fixed NIS threshold may be mismatched with the actual innovation distribution under a real non-stationary buoy background. Therefore, this paper replaces the fixed background-consistency threshold τ_b and target-protection threshold τ_t with an online threshold-adaptive form. Let $\mathcal{B}_k$ denote the alarm-free background NIS window before the current time, and let $\mathcal{G}_{k-1}$ denote the candidate alarms, confirmed alarms, and their protection buffers deter and mined up to time $k - 1$. Then,

$$\mathcal{B}_k = \{\eta_i \mid i \in [k - W_0, k - 1], i \notin \mathcal{G}_{k-1}\}, \quad \tau_b^{\text{raw}}(k) = \hat{Q}_{p_b}(\mathcal{B}_k), \quad \tau_t^{\text{raw}}(k) = \hat{Q}_{p_t}(\mathcal{B}_k). \quad (11)$$

where W_0 is the length of the online background window, $\hat{Q}_p(\cdot)$ denotes empirical quantile estimation, and p_b and p_t are the background-consistency quantile and target-protection quantile, respectively. To avoid threshold jumps caused by short-term spikes, exponential smoothing and a minimum-spacing constraint are further adopted:

$$\tau_b(k) = (1 - \alpha_\tau)\tau_b(k - 1) + \alpha_\tau\tau_b^{\text{raw}}(k), \quad (12)$$

$$\tau_t(k) = \max\{(1 - \alpha_\tau)\tau_t(k - 1) + \alpha_\tau\tau_t^{\text{raw}}(k), \tau_b(k) + \Delta_\tau\}. \quad (13)$$

The above online CFAR-NIS thresholds are updated only from the alarm-free background window. Target-passage samples, candidate-alarm samples, confirmed-alarm samples, and their protection buffers are all excluded from quantile estimation, thereby avoiding inflation of the protection threshold by target energy. When $|\mathcal{B}_k|$ is smaller than the minimum valid sample size, $\tau_b(k)$ and $\tau_t(k)$ retain their previous values or are initialized using the χ_3^2 quantile or empirical quantiles from a target-free calibration segment.

To further characterize the dwell behavior during continuous target passage, a continuous target-dwell counter D_k and a low-NIS recovery counter L_k are introduced. When $\eta_k > \tau_b(k)$, the suspected target-dwell strength increases, and the recovery counter is reset to zero. When $\eta_k \leq \tau_b(k)$, protection is fully released only after C_r consecutive low-NIS samples; otherwise, the dwell strength decays gradually. The update rule is

$$(D_k, L_k) = \begin{cases} (\min\{D_{k-1} + 1, D_{\max}\}, 0), & \eta_k > \tau_b(k), \\ (0, \min\{L_{k-1} + 1, C_r\}), & \eta_k \leq \tau_b(k), L_{k-1} + 1 \geq C_r, \\ (\max\{D_{k-1} - 1, 0\}, \min\{L_{k-1} + 1, C_r\}), & \eta_k \leq \tau_b(k), L_{k-1} + 1 < C_r. \end{cases} \quad (14)$$

Accordingly, the gating weight is no longer determined only by the current η_k but is jointly affected by the current NIS exceedance level and the continuous dwell strength. Define

$$r_k = \text{clip}\left(\frac{\eta_k - \tau_b(k)}{\tau_t(k) - \tau_b(k) + \varepsilon}, 0, 1\right), \quad s_k = \frac{D_k}{D_{\max}} \quad (15)$$

Then, the target-protection gating weight jointly driven by online threshold adaptation and continuous target-dwell memory is given by

$$\mu_k = \begin{cases} 1, & \eta_k \leq \tau_b(k), D_k = 0, \\ \max\{\mu_{\min}, \exp[-\beta_r r_k - \beta_c s_k]\}, & \eta_k \leq \tau_t(k), D_k > 0 \text{ or } \eta_k > \tau_b(k), \\ \mu_{\min}, & \eta_k > \tau_t(k). \end{cases} \quad (16)$$

The exponential mapping in Equation (11) is selected as a one-sided attenuation function rather than as a probabilistic link function. With $\beta_r \geq 0$, $\beta_c \geq 0$, and the normalized variables r_k and s_k defined in Equation (10), the term $\exp(-\beta_r r_k - \beta_c s_k)$ decreases smoothly and monotonically as the NIS exceedance and target-dwell memory increase. The lower bound $\mu_{\min}$ prevents the effective update weight from vanishing completely. Thus, the exponential form keeps the update weight close to one under weak

protection evidence and gradually reduces it under stronger NIS exceedance or longer suspected target dwell. A logistic mapping would require additional midpoint and slope parameters and would be less direct for this one-sided, lower-bounded attenuation purpose. The exponential form is therefore an engineering design choice based on monotonicity, smoothness, and simple implementation, rather than the only possible mapping.

Where $\mu_{\min}$ is the minimum update weight allowed during the target-protection stage, β_r controls the attenuation strength associated with the current NIS exceedance level, and β_c controls the suppression strength of continuous target-dwell memory on gating-weight recovery. Compared with a fixed NIS threshold, this mechanism can adaptively adjust the threshold when the background statistics drift slowly and can prevent the filter from prematurely resuming full updating due to a single low-NIS sample during continuous weak-target occurrence.

In the recursive update of the measurement noise covariance, the Sage–Husa forgetting factor is combined with the above gating weight to construct a composite adaptive updating factor, and the unconstrained covariance estimate is projected within bounded limits. To avoid notational ambiguity, the unconstrained estimate is given first, followed by the projected measurement noise covariance, effective Kalman gain, and state update:

$$\alpha_k = (1 - \lambda_k)\mu_k, \quad (17)$$

$$\tilde{\mathbf{R}}_k = (1 - \alpha_k)\mathbf{R}_{k-1} + \alpha_k(\mathbf{v}_k\mathbf{v}_k^T - \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T), \quad (18)$$

$$\mathbf{R}_k = \Pi_{[\mathbf{R}_{\min}, \mathbf{R}_{\max}]}(\tilde{\mathbf{R}}_k), \quad (19)$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}^T\mathbf{S}_k^{-1}, \quad (20)$$

$$\mathbf{K}_k^{\text{eff}} = \mu_k\mathbf{K}_k, \quad (21)$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k^{\text{eff}}\mathbf{v}_k. \quad (22)$$

Equation (22) describes the measurement-update step and should not be interpreted as the state-transition equation. The temporal evolution of the background state is determined by the state-transition and prediction model given in Equations (4)–(6). The effective gain $\mathbf{K}_k^{\text{eff}} = \mu_k\mathbf{K}_k$ controls the influence of the current innovation on the background-state update, particularly during suspected target intervals.

For completeness, the posterior covariance carried forward to the next prediction is given by Equation (23):

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k^{\text{eff}}\mathbf{H})\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{K}_k^{\text{eff}}\mathbf{H})^T + \mathbf{K}_k^{\text{eff}}\mathbf{R}_{k-1}(\mathbf{K}_k^{\text{eff}})^T \quad (23)$$

At sample k , $\mathbf{R}_{k-1}$ is used to calculate $\mathbf{S}_k$, the NIS, and $\mathbf{K}_k$. The covariance update obtained from the current innovation is then carried forward as $\mathbf{R}_k$ for the next sample. During suspected-target or dwell-protection intervals, the innovation-based covariance update is disabled, and the covariance is only weakly relaxed toward the baseline covariance.

Here, $\alpha_k = (1 - \lambda_k)\mu_k$ denotes the composite adaptive covariance-update factor. The forgetting factor λ_k controls the nominal adaptation rate, while the target-protection gating weight μ_k suppresses covariance and state updates during suspected target intervals. The unconstrained covariance estimate $\tilde{\mathbf{R}}_k$ is projected onto a bounded positive-definite set to avoid nonphysical covariance values caused by finite-sample fluctuations. The effective Kalman gain $\mathbf{K}_k^{\text{eff}}$ further reduces the influence of target-contaminated innovations on the background-state estimate.

Together, these constraints preserve target-related innovations while maintaining adaptive background tracking under stable or slowly varying conditions.

2.3. GLRT with Target-Free Calibration-Based Covariance Whitening

Recursive filtering uses the adaptive covariance to control background tracking and gain, whereas whitening and GLRT construction use a separate covariance estimated from target-free calibration data. This separation prevents transient or target-protection-related covariance adaptation from directly biasing the detection statistic.

During the detection stage, the unconstrained adaptive covariance within suspected target intervals is not used directly. Instead, the detection covariance $\mathbf{R}_{\text{det}}$ is estimated from a target-free calibration segment and then boundedly corrected. Let the detection covariance be

$$\mathbf{S}_{\text{det},k} = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T + \mathbf{R}_{\text{det}} \quad (24)$$

After applying Cholesky decomposition to $\mathbf{S}_{\text{det},k}$, i.e., $\mathbf{S}_{\text{det},k} = \mathbf{C}_k \mathbf{C}_k^T$, the whitened innovation is obtained as

$$\mathbf{z}_k = \mathbf{C}_k^{-1}\mathbf{v}_k \quad (25)$$

When the target direction and amplitude are unknown, the target can be modeled as an unknown mean shift in the whitened residual space. The single-sample GLRT energy statistic is

$$T_k = \|\mathbf{z}_k\|_2^2 \quad (26)$$

The GLRT increment actually used for sequential accumulation is

$$\ell_k = 0.5T_k - \gamma \quad (27)$$

where γ is the drift compensation parameter. In the main configuration, $\gamma = 1.5$ is used to obtain a lower false-alarm level under real buoy background conditions.

2.4. Dual-Threshold Weighted Multiscale GLRT-CUSUM Fusion

To accommodate target magnetic anomaly responses with different durations, this paper performs windowed accumulation of the single-step GLRT increment ℓ_k over multiple time scales. Let the scale set be $\mathcal{Q} = \{S, M, L\}$, where S , M , and L denote the short, medium, and long scales, respectively, with corresponding window lengths $W_S = 3$, $W_M = 10$, and $W_L = 30$.

For the q -th scale, the windowed GLRT increment is first constructed as

$$\bar{\ell}_{q,k} = \frac{1}{W_q} \sum_{i=0}^{W_q-1} \ell_{k-i}, \quad q \in \mathcal{Q} \quad (28)$$

Based on this windowed increment, the CUSUM statistic at the q -th scale is defined as

$$G_{q,k} = \max(0, G_{q,k-1} + \bar{\ell}_{q,k}), \quad q \in \mathcal{Q} \quad (29)$$

where the different window lengths W_q are used to match response characteristics under different target passing speeds, heading conditions, and magnetic anomaly durations. The short-scale branch is more sensitive to local abrupt changes, the medium-scale branch captures typical target responses, and the long-scale branch enhances accumulated evidence for weak persistent anomalies.

Since CUSUM statistics at different scales and the direct GLRT increment have different numerical ranges and background distributions, a target-free calibration segment

$\mathcal{H}_0$ is used to perform robust non-negative normalization for different evidence statistics. Let

$$S_D(k) = \ell_k, \quad S_q(k) = G_{q,k}, \quad q \in \mathcal{Q} = \{S, M, L\} \quad (30)$$

Then, the normalized evidence is defined as

$$Z_j(k) = \max \left[0, \frac{S_j(k) - \text{median}(S_j^{\mathcal{H}_0})}{1.4826 \text{MAD}(S_j^{\mathcal{H}_0}) + \varepsilon} \right], \quad j \in \{D, S, M, L\} \quad (31)$$

where D denotes the direct single-step GLRT increment evidence and S , M , and L denote the short-, medium-, and long-scale CUSUM evidence, respectively. Therefore, $Z_D(k)$ is the normalized evidence of the direct GLRT increment, while $Z_S(k)$, $Z_M(k)$, and $Z_L(k)$ are the normalized evidence of the short-, medium-, and long-scale CUSUM statistics, respectively. $\text{MAD}(\cdot)$ denotes the median absolute deviation, and ε is a very small positive constant used to ensure numerical stability. This treatment reduces the influence of magnitude differences among different evidence statistics on the fusion result and improves robustness against abnormal background spikes.

The factor 1.4826 is the Gaussian consistency factor for the median absolute deviation. It is defined by $c_{\text{MAD}} = \frac{1}{\Phi^{-1}(0.75)} \approx 1.4826$, where $\Phi^{-1}(\cdot)$ denotes the inverse cumulative distribution function of the standard normal distribution. The fourth decimal place is not essential; replacing 1.4826 with 1.483 changes the scale factor by only approximately 0.027% and is not expected to materially affect the results, provided that the detection thresholds are recalibrated consistently.

To prevent the long-scale branch from accumulating continuously on its own under low-frequency background disturbances, a long-scale participation gate is introduced:

$$I_L(k) = \begin{cases} 1, & Z_M(k) > 0.8 \text{ or } Z_D(k) > 1.5, \\ 0, & Z_M(k) \leq 0.8 \text{ and } Z_D(k) \leq 1.5. \end{cases} \quad (32)$$

The final fusion statistic is defined as

$$F_k = 0.10Z_D(k) + 0.80Z_M(k) + 0.10Z_L(k)I_L(k) \quad (33)$$

where $Z_D(k)$, $Z_M(k)$, and $Z_L(k)$ denote the robustly normalized results of the direct GLRT evidence, medium-scale CUSUM evidence, and long-scale CUSUM evidence, respectively. The long-scale evidence participates in the fusion only when supported by the medium-scale or direct GLRT evidence. In the final low-false-alarm configuration, the short-scale branch is mainly used for ablation experiments and multiscale response analysis and is not directly included in the final weighted fusion.

Then, two CFAR thresholds, a low threshold and a high threshold, are calibrated for F_k using a target-free calibration segment:

$$h_{\text{low}} = Q_{99.55\%}(F^{\mathcal{H}_0}), \quad h_{\text{high}} = Q_{99.85\%}(F^{\mathcal{H}_0}) \quad (34)$$

where $Q_p(\cdot)$ denotes the p -th quantile of the empirical distribution. A candidate event is generated when $F_k > h_{\text{low}}$. It is further confirmed if

$$F_k > h_{\text{high}} \quad (35)$$

or if the medium-scale evidence satisfies the following condition within the subsequent $N_c = 8$ sampling points:

$$\max_{0 \leq r \leq N_c} Z_M(k+r) > \eta_c, \quad \eta_c = 1.10, \quad N_c = 8 \quad (36)$$

The candidate event is then confirmed as a target alarm. Otherwise, it is regarded as an isolated background spike or short-term disturbance.

This dual-threshold candidate-confirmation mechanism avoids false-alarm accumulation caused by simple hard-OR multiscale voting. The low threshold h_{low} is used to retain weak target candidate responses, whereas the high threshold h_{high} and the medium-scale confirmation condition are used to suppress false alarms caused by background spikes, thereby achieving a better trade-off among detection probability, false-alarm rate, and detection delay under real buoy background conditions.

For readability, the principal notation is summarized by processing stage in Table 1. Symbols not listed here are defined when they are first introduced.

Table 1. Notation summary for the proposed detection chain.

Processing Stage	Principal Symbols	Role in the Detection Chain
Observation and state transition	$y_{\text{raw},k}, y_k, b_k, x_k, a_k, v_k, H, F, Q_{k-1}, w_{k-1}$	Raw and compensated observations; slow background; latent target-free state; target anomaly; measurement noise; observation matrix; state-transition matrix; process-noise covariance; and process disturbance.
Kalman prediction and update	$\hat{x}_{k k-1}, P_{k k-1}, v_k, S_k, R_{k-1}, \eta_k, K_k, K_k^{\text{eff}}, \hat{x}_{k k}, P_{k k}, R_k$	Predicted state and covariance; innovation and prior innovation covariance; previous measurement covariance; NIS; nominal and effective Kalman gains; posterior state and covariance; and adaptive filter covariance.
Online CFAR-NIS target protection	$B_k, G_{k-1}, \tau_{b,k}, \tau_{t,k}, W_0, p_b, p_t, D_k, L_k, C_r, r_k, S_k, \mu_k, \mu_{\min}, \beta_r, \beta_c$	Alarm-free NIS window and prior alarm/protection record; adaptive background and target-protection thresholds; quantile/window settings; target-dwell and low-NIS recovery counters; normalized exceedance and dwell terms; gating weight, lower bound, and attenuation coefficients.
Covariance adaptation	λ_k, α_k	Forgetting factor and composite covariance-update factor. The latter combines nominal adaptation with the target-protection gating weight.
Target-free whitening and GLRT	$R_{\text{det}}, S_{\text{det},k}, C_k, z_k, T_k, \ell_k, \gamma$	Target-free calibration-based detection covariance; detection covariance matrix; Cholesky factor; whitened innovation; GLRT energy; sequential GLRT increment; and drift-compensation parameter. The Cholesky factor uses a separate symbol so that the recovery-counter notation remains unambiguous.
Multiscale evidence accumulation	$Q, q, W_q, G_{q,k}, Z_{D,k}, Z_{S,k}, Z_{M,k}, Z_{L,k}$	The scale-set notation is distinct from the process-noise covariance in the state-transition model. The scale index identifies the short, medium, and long branches; the branch window length determines the accumulation horizon; and the four normalized evidence terms represent direct, short-, medium-, and long-scale evidence.
Fusion, confirmation, and indexing	$F_k, h_{\text{low}}, h_{\text{high}}, \eta_c, N_c, k, q$	Fused detection statistic; low and high confirmation thresholds; medium-scale confirmation threshold; confirmation-window length; sample index; and scale-branch index. The fused statistic is distinct from the state-transition matrix.

The subscript k denotes the sample index, whereas q denotes the selected scale branch. The short-scale branch is retained for ablation and response analysis but is not directly included in the final weighted fusion.

2.5. Overall Detection Procedure

Algorithm 1 summarizes the causal online sequence from attitude compensation and target protection to residual whitening, multiscale evidence accumulation, and alarm confirmation.

Algorithm 1. Proposed target-protected multiscale GLRT-CUSUM detection procedure

Input: raw triaxial magnetic data, roll, pitch, heading, geomagnetic vector

Output: target alarm sequence

1. Perform attitude compensation using Equation (2).
2. Compute innovation and NIS using Equations (7)–(10).
3. Update online CFAR-NIS thresholds using Equations (11)–(13).
4. Update dwell counter and gating weight using Equations (14)–(16).
5. Perform target-protected Kalman filtering using Equation (17)–(23).
6. Whiten residuals and compute the GLRT statistic using Equations (24)–(27).
7. Accumulate multiscale CUSUM evidence using Equations (28)–(31).
8. Fuse normalized evidence and confirm alarms using Equations (32)–(36).

All steps use the current sample and previously available quantities; threshold and dwell updates precede filtering, and only the protected residual is passed to the downstream detection stages.

3. Experimental Setup

3.1. Data Sources and Semi-Physical Target Injection

The target magnetic anomaly source was obtained from a publicly available ferromagnetic ship magnetic signature reproduction dataset [29,30]. In this study, the KD0 track direction in DataS2_V1.mat was selected for the experiments. At the V1 geomagnetic location, the geomagnetic field magnitude is approximately 50.0 μT , and the peak amplitude of the dominant unscaled ship magnetic component is approximately 420.7 nT. To prevent the target signal from occupying the entire time axis, only the central ± 80 m interval of the spatial magnetic profile was retained and converted into a time series using a constant ship speed of 5.0 m/s. The sampling interval was set to 0.05 s; therefore, the detection delay in sampling points can be converted into physical time by multiplying it by 0.05.

The background data were obtained from the public buoy dataset of the U.S. Department of Energy Wind Data Hub/Atmosphere to Electrons Data Archive and Portal, which contains real buoy roll, pitch, heading, and triaxial magnetometer background records. A total of 96 IMU files were used, yielding approximately 1,727,749 valid samples covering about 48 h of records. To avoid overlap between parameter tuning and performance testing, the second half of the records was used as the test background. The attitude-compensated real buoy background was normalized so that the standard deviation of the calibration segment was 20 nT. Six low to medium-low SNR conditions were considered, namely 2.3, 2.4, 2.5, 2.6, 2.7, and 2.8. For each SNR condition, 200 Monte Carlo trials were conducted.

3.2. Evaluation Metrics and Threshold Calibration

The event-level detection probability was defined as the probability that at least one alarm occurred within the target-passage interval. The background false-alarm rate was calculated using samples outside the target interval and its protection buffer. All detection thresholds and online NIS thresholds were calibrated only from a target-free calibration segment or an alarm-free background window. Target-passage samples, candidate-alarm samples, confirmed-alarm samples, and their protection buffers were excluded from threshold calibration to avoid target information leakage. The comparison methods

included standard KF single-scale CUSUM, adaptive KF single-scale CUSUM [13,14], OBF, minimum entropy, SR-CUSUM [27], and the proposed method. To ensure a fair comparison, all methods used the same target data, real buoy background data pool, SNR scaling rule, Monte Carlo random seed division, target injection interval, and event-level evaluation metrics. The thresholds of all methods were calibrated from a target-free calibration segment or an independent background segment using empirical quantiles, without using any target-interval information. Standard KF, adaptive KF, OBF, minimum entropy, SR-CUSUM, and the proposed method were all applied to the same attitude-compensated triaxial residual signals. Standard KF and adaptive KF used single-scale GLRT-CUSUM detection, OBF and minimum entropy used the same calibration segment to determine their detection thresholds, and SR-CUSUM used the same target-free calibration rule for spectral-residual saliency accumulation. The proposed method further introduced online CFAR-NIS target protection and dual-threshold weighted multiscale fusion under the same low-false-alarm constraint.

For statistical validation, the trial-level detection probability, background false-alarm rate, and detection delay were retained for all 200 paired Monte Carlo trials at each SNR condition. The 95% confidence intervals were estimated by nonparametric bootstrap resampling with 10,000 replicates. Because all methods were evaluated on the same trial realizations, two-sided paired sign-flip permutation tests with 10,000 replicates were used to compare the proposed method with each baseline. The resulting p -values were adjusted using the Holm procedure separately for each metric and SNR condition, and a corrected significance level of $p < 0.05$ was adopted.

For reproducibility, the final reference implementation was executed with Python 3.10.20 and NumPy 2.2.6 on Windows 10. The main SNR-sweep trials used a deterministic seed generated as 100 times the Monte Carlo trial index plus the SNR-condition index, while fixed-SNR supplementary trials used 100 times the trial index plus 99. The attitude-compensation, parameter-sensitivity, and threshold-sweep supplementary trials used seeds starting at 10,000 with increments of 100. The bootstrap and paired sign-flip permutation analyses used the base seed 20,260,726, with 10,000 replicates each and deterministic offsets by SNR condition, metric, and method. Each primary and supplementary condition used 200 Monte Carlo trials. The final parameter configuration is summarized in Table 2 and retained in the experiment script. The scripts, processed background segments, injected target data, configuration settings, and result tables are available from the corresponding author upon reasonable request; no public GitHub or Zenodo repository was available at the time of this revision.

Table 2. Main experimental and detection parameters.

Item	Setting
Target data	DataS2_V1, KD0 track direction
Background data	Public buoy attitude data and triaxial magnetometer background from DOE Wind Data Hub/A2E
Test segment	Second 50% of the 48 h record
Ship speed	5.0 m/s
Sampling interval	0.05 s
Noise/background standard deviation	20 nT after normalization
SNR condition	2.3, 2.4, 2.5, 2.6, 2.7, and 2.8
Monte Carlo	200 trials for each SNR
Proposed windows	3/10/30 sampling points
Dual thresholds	low = 99.55%, high = 99.85%
Fusion weights	0.10/0.80/0.10
Heading robustness experiment	KD0, PD0, and SD0, fixed SNR ≈ 2.7
Geomagnetic-location robustness experiment	V1-V6, fixed SNR ≈ 2.7
Online NIS background window	$W_0 = 1000$ sampling points; candidate/alarm protection-zone samples excluded from updating
NIS quantile thresholds	$p_b = 99.0\%$, $p_t = 99.7\%$
Threshold smoothing and spacing	$\alpha\tau = 0.02$, $\Delta\tau = 1.0$

Dwell-protection parameters	$D_{\max} = 4, C_r = 4, \mu_{\min} = 0.05$
Equivalent passing-distance experiment	60–120 m, with 80 m as the reference distance
Reference condition	80 m corresponding to $\text{SNR} \approx 2.70$
Distance attenuation model	$B(r) \propto 1/r^3$
Reliable-distance criterion	$P_d \geq 0.90$ and $P_{fa} \leq 0.05$
SR-CUSUM baseline parameters	SSR window = 64; CUSUM window = 20; threshold percentile = 99.5%

3.3. Baseline Methods

To improve the reproducibility of the comparison experiments, both OBF and minimum entropy were applied to the attitude-compensated triaxial residuals. For OBF, discrete orthogonal polynomials of orders 0–3 were used as background basis functions, with the window length set to $N_{\text{OBF}} = 30$. For each axis, the background component was estimated within the sliding window by least squares, and the triaxial reconstructed residual energy

$$T_{\text{OBF}}(k) = r_x^2(k) + r_y^2(k) + r_z^2(k) \quad (37)$$

was used as the detection statistic. For the minimum-entropy method, the triaxial magnitude was first calculated as

$$a(k) = \sqrt{x^2(k) + y^2(k) + z^2(k)} \quad (38)$$

Then, within a sliding window of $N_{\text{ME}} = 30$, $B = 20$ histogram bins were used to estimate the probability $p_i(k)$, and the window entropy was calculated as

$$H(k) = - \sum_{i=1}^B p_i(k) \log[p_i(k) + \varepsilon] \quad (39)$$

The detection statistic was defined as

$$T_{\text{ME}}(k) = H_0 - H(k) \quad (40)$$

where H_0 denotes the average entropy in the target-free calibration segment and ε is a small positive constant used to avoid logarithmic singularity. The detection thresholds of both methods were kept consistent with those in the main experiment and were calibrated using empirical quantiles from the target-free calibration segment, without tuning on target-interval samples.

In addition, an SR-CUSUM baseline was included to represent a lightweight spectral-residual time-series anomaly detector [27,28]. The SR-CUSUM baseline was adapted from the spectral-residual saliency idea originally proposed by Hou and Zhang [27] and its later extension to time-series anomaly detection by Ren et al. [28]. The attitude-compensated triaxial residuals were first converted into a scalar magnitude sequence after robust axis-wise normalization. A causal sliding-window spectral-residual transform was then applied to obtain the saliency score. The saliency sequence was robustly normalized using the target-free calibration segment and accumulated by a single-scale CUSUM detector. The SR-CUSUM detection threshold was calibrated using the same target-free empirical-quantile rule as the other baseline methods, without tuning on target-interval samples.

3.4. Robustness and Equivalent Passing-Distance Experimental Design

In addition, to evaluate the adaptability of the proposed algorithm to variations in target heading and geomagnetic environment, $\text{SNR} \approx 2.7$ was selected as a representative operating point. Under 200 Monte Carlo repetitions, two supplementary experiments were conducted. In the first experiment, DataS2 and the V1 geomagnetic location were fixed, and the detection performance was compared under three heading conditions: keel-direction longitudinal passage (KD0), port-direction transverse passage (PD0), and star-board-direction transverse passage (SD0). In the second experiment, DataS2 and the KD0

track direction were fixed, and the detection performance was evaluated at six geomagnetic locations, namely V1–V6: North Sea 55°N/2°E (V1), Arctic 90°N (V2), Southeast Asia 15°N (V3), near the equator 10°N (V4), southern Australia 25°S (V5), and near Antarctica 63°S (V6). SNR $\approx$ 2.7 was selected because it lies in the upper part of the low-SNR test range and provides moderate detection difficulty. This setting reduces the influence of random background fluctuations under extremely low-SNR conditions while preserving performance differences among methods in terms of detection probability, false-alarm rate, and detection delay. The supplementary experiments used the same real buoy background segmentation, threshold calibration rules, and event-level evaluation metrics as those in the main experiment.

To provide an application-oriented interpretation of the low-SNR detection results, an equivalent passing-distance experiment was further designed. Since the far-field magnetic anomaly amplitude of a ferromagnetic target approximately follows a cubic distance attenuation relationship, i.e., $B(r) \propto 1/r^3$, the 80 m case was selected as the reference distance corresponding to SNR $\approx$ 2.70. Based on this reference condition, the target magnetic signature was scaled to simulate equivalent passing distances of 60, 70, 80, 90, 100, 110, and 120 m. For each equivalent distance, the same real buoy background pool, target injection strategy, threshold calibration rule, and Monte Carlo evaluation procedure were used. The evaluation metrics included event-level detection probability, background false-alarm rate, and average detection delay. It should be noted that this experiment represents an equivalent passing-distance analysis under the semi-physical injection framework, rather than a direct sea-trial measurement of detection range.

The main experimental settings and detection parameters are summarized in Table 2.

4. Results

4.1. Residual Separability and Typical Detection Process

The H_0 and H_1 mean values reported in Table 3 represent the drift-compensated single-step GLRT increments after subtraction of the drift parameter γ , rather than the original non-negative GLRT energy statistic. Consequently, negative mean values may occur. As shown in Table 3, target-protected adaptive filtering combined with target-free calibration-based covariance whitening increases the H_1 – H_0 mean difference of the GLRT increment from -0.033 for the standard KF residual to 0.377 for the proposed residual. Although this improvement in separability cannot completely eliminate the overlap between the H_0 and H_1 distributions, it provides a more discriminative residual basis for subsequent sequential accumulation and dual-threshold fusion.

Table 3. Separability of the drift-compensated single-step GLRT increments under H_0/H_1 at SNR $\approx$ 2.7.

Method	H_0 Mean	H_1 Mean	Separability
Standard KF	0.152	0.120	-0.033
Proposed residual	-0.625	-0.248	0.377

Figure 2 presents a representative single-run detection process at SNR $\approx$ 2.7. The attitude-compensated triaxial residuals exhibit local anomalous responses within the target interval; however, these responses overlap with real buoy background disturbances, making it difficult for single-step statistics to generate stable detection decisions. The proposed method preserves suspected target residuals through online CFAR-NIS target-protected filtering, accumulates short-term and persistent anomaly evidence in parallel using three-scale CUSUM, and outputs alarms through normalized weighted fusion and a dual-threshold confirmation mechanism. In this representative trial, standard KF, adaptive KF,

OBF, and minimum entropy fail to provide an effective target alarm within the target interval. SR-CUSUM responds to the target, but it also produces a higher background false-alarm rate. By contrast, the proposed method generates an effective target alarm while maintaining a lower background false-alarm level in this trial. This result illustrates that the proposed target-protected multiscale GLRT-CUSUM framework can enhance weak magnetic anomaly detectability while improving false-alarm suppression under real buoy background conditions.



Figure 2. Example of a typical single-run detection process at $\text{SNR} \approx 2.7$ with six detection methods.

4.2. Main Performance Comparison Under Different SNR Levels

Table 4 shows that, under the real DOE/A2E buoy background, the compared detection methods exhibit clear differences in the trade-off between detection probability and background false-alarm rate. The standard KF and adaptive KF methods show relatively low detection probabilities across all SNR conditions, indicating that single-scale residual accumulation has limited capability to preserve weak magnetic anomaly responses under low-SNR and non-stationary buoy background conditions. The OBF method achieves a certain detection capability under low-false-alarm constraints, and SR-CUSUM further provides a competitive lightweight spectral-residual anomaly-detection baseline. Compared with standard KF, adaptive KF, and minimum entropy, SR-CUSUM generally achieves higher detection probabilities, indicating that spectral-residual saliency is effective in enhancing weak anomaly responses. However, its false-alarm rate remains close to the upper low-false-alarm constraint, and its detection delay at $\text{SNR} \approx 2.7$ is still longer than that of the proposed method. Although SR-CUSUM obtains a slightly higher detection probability at the lowest SNR of 2.3, the proposed method provides a better overall balance between detection probability, false-alarm suppression, and detection latency over the full SNR range. At $\text{SNR} \approx 2.7$, the proposed method achieves $P_d = 0.96$, $P_{fa} = 0.05$, and an average delay of 208.3 sampling points.

Table 4. Detection results under different SNR conditions (P_d/P_{fa}).

SNR	Standard KF	Adaptive KF	OBF	Minimum Entropy	SR-CUSUM	Proposed Method
2.3	0.72/0.03	0.74/0.03	0.89/0.05	0.73/0.02	0.90/0.05	0.89/0.04
2.4	0.67/0.03	0.71/0.03	0.89/0.05	0.74/0.02	0.87/0.05	0.90/0.05
2.5	0.70/0.03	0.71/0.03	0.84/0.05	0.70/0.02	0.89/0.05	0.91/0.05
2.6	0.73/0.02	0.76/0.02	0.88/0.04	0.71/0.02	0.91/0.04	0.97/0.04
2.7	0.74/0.03	0.81/0.03	0.88/0.05	0.74/0.02	0.89/0.05	0.96/0.05
2.8	0.71/0.03	0.79/0.03	0.90/0.05	0.67/0.02	0.92/0.04	0.96/0.04

It should be noted that the detection probabilities of some baseline methods are not strictly monotonic with increasing SNR. This is mainly because each SNR condition was evaluated using finite Monte Carlo trials with randomly selected real buoy background segments. The real DOE/A2E background contains non-stationary fluctuations, impulsive disturbances, and local correlation variations; therefore, adjacent SNR levels may exhibit small statistical fluctuations. Nevertheless, the overall results indicate that the proposed method consistently achieves higher detection probability under comparable low-false-alarm constraints.

Figure 3 further compares the event-level detection probability and background false-alarm rate of different methods over the tested SNR range.

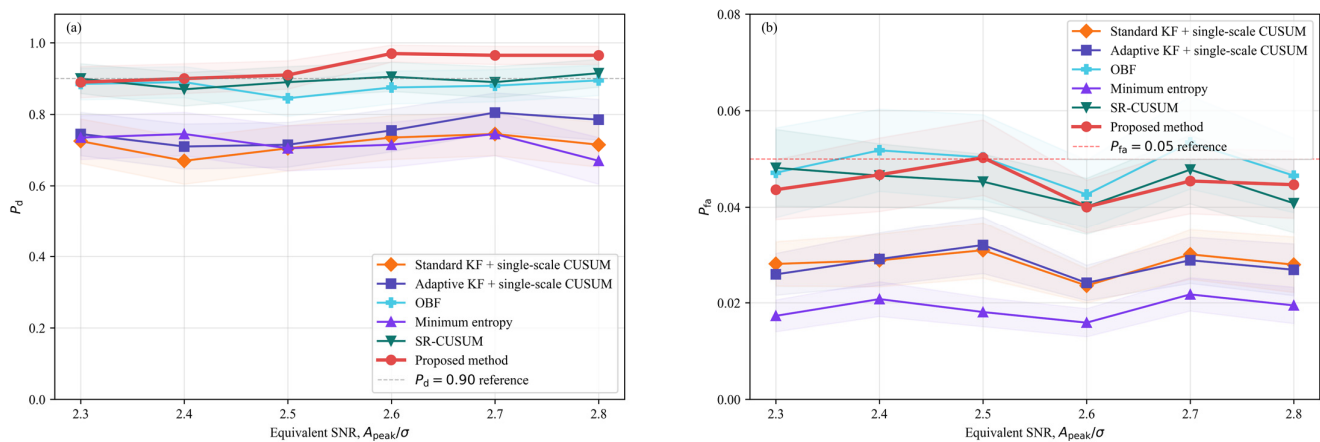


Figure 3. Event-level detection probability P_d and background false-alarm rate P_{fa} versus equivalent SNR for the six detection methods. (a) Event-level detection probability; (b) background false-alarm rate. The shaded regions denote 95% confidence intervals, and the dashed lines indicate the reference levels $P_d = 0.90$ and $P_{fa} = 0.05$.

As shown in Table 5, at $\text{SNR} \approx 2.7$, the compared detection methods exhibit clear differences in detection probability, background false-alarm rate, and detection delay. The standard KF single-scale and adaptive KF single-scale methods show relatively low detection probabilities, indicating that single-scale residual accumulation has limited ability to preserve weak magnetic anomaly responses in real non-stationary buoy backgrounds. The OBF and SR-CUSUM baselines achieve higher detection probabilities than the standard KF and adaptive KF methods; however, their average detection delays remain longer than that of the proposed method. The minimum-entropy method yields the lowest false-alarm rate, but its detection probability is only 0.745, with an average delay of 339.3 sampling points, indicating overly conservative detection behavior. In contrast, the proposed method achieves a high detection probability of 0.965, controls the background false-alarm rate at 0.045, and reduces the average detection delay to 208.3 sampling points. Overall, the proposed method provides a better balance among detection probability, false-alarm rate, and detection delay, demonstrating its effectiveness for weak magnetic anomaly detection under real buoy background conditions.

Table 5. Comprehensive performance comparison at $\text{SNR} \approx 2.7$. Mean values are reported over 200 paired Monte Carlo trials; brackets denote 95% nonparametric bootstrap confidence intervals.

Method	P_d	P_{fa}	Delay/Samples
Standard KF single-scale	0.745 [0.685, 0.805]	0.030 [0.0250, 0.0354]	338.5 [304.8, 372.4]
Adaptive KF single-scale	0.805 [0.750, 0.860]	0.029 [0.0243, 0.0338]	319.2 [288.9, 350.0]
OBF	0.880 [0.835, 0.925]	0.053 [0.0440, 0.0635]	255.6 [225.5, 285.6]
Minimum entropy	0.745 [0.685, 0.805]	0.022 [0.0185, 0.0252]	339.3 [306.3, 372.7]
SR-CUSUM	0.890 [0.845, 0.930]	0.048 [0.0409, 0.0550]	256.5 [228.5, 285.1]
Proposed method	0.965 [0.935, 0.990]	0.045 [0.0388, 0.0525]	208.3 [188.2, 228.8]

Paired sign-flip permutation tests showed that the proposed method achieved a significantly higher event-level detection probability than the standard KF single-scale, adaptive KF single-scale, OBF, minimum-entropy, and SR-CUSUM baselines (Holm-adjusted p -values ≤ 0.0022). Its detection delay was also significantly shorter than those of all five baselines (Holm-adjusted p -values ≤ 0.0064). In contrast, its background false-alarm rate was significantly higher than those of standard KF, adaptive KF, and minimum entropy (Holm-adjusted p -values < 0.001) but did not differ significantly from OBF ($p =$

0.0868) or SR-CUSUM ($p = 0.6052$). Thus, the statistically supported advantages are higher detection probability and shorter delay, while the false-alarm rate represents a trade-off: it is comparable to OBF and SR-CUSUM but higher than the more conservative KF and minimum-entropy baselines.

4.3. Ablation Study

4.3.1. Multiscale Fusion Module

Table 6 shows that the single short-scale and medium-scale branches maintain relatively low false-alarm rates, but they also lead to longer detection delays. Combining the medium- and long-scale branches shortens the detection delay and improves the detection probability; however, it increases the false-alarm rate to $P_{fa} = 0.080$. In contrast, the proposed dual-threshold weighted fusion maintains a high detection probability of $P_d = 0.975$, controls the false-alarm rate at $P_{fa} = 0.049$, and reduces the average detection delay to 213.8 sampling points. These results indicate that normalized multiscale evidence fusion, together with the candidate-confirmation mechanism, achieves a favorable balance between sensitivity to target responses of different durations and false-alarm suppression. The corresponding ablation results are also illustrated in Figure 4.

Table 6. Ablation study of the multiscale fusion module at $\text{SNR} \approx 2.7$.

Variant	P_d	P_{fa}	Delay/Samples
AKF + short scale	0.955	0.041	224.8
AKF + medium scale	0.965	0.043	217.1
AKF + long scale	0.965	0.068	217.6
AKF + short + medium	0.965	0.053	213.8
AKF + medium + long	0.980	0.080	196.4
Proposed dual-threshold weighted fusion	0.975	0.049	213.8

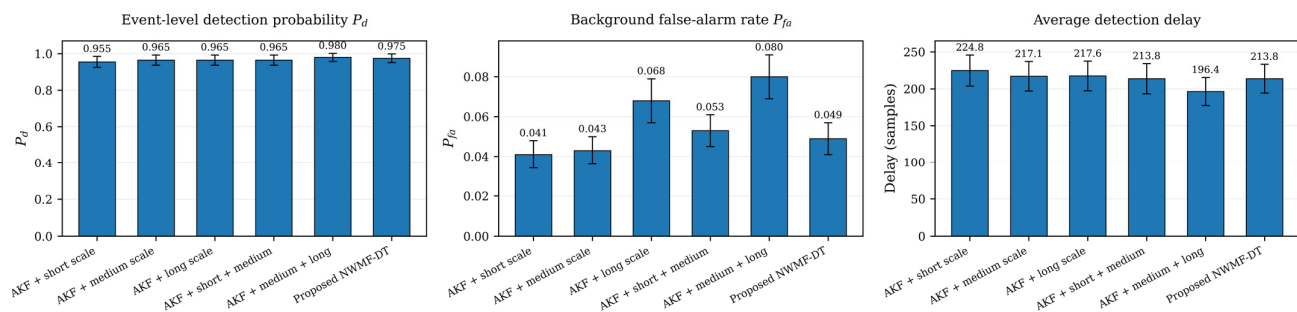


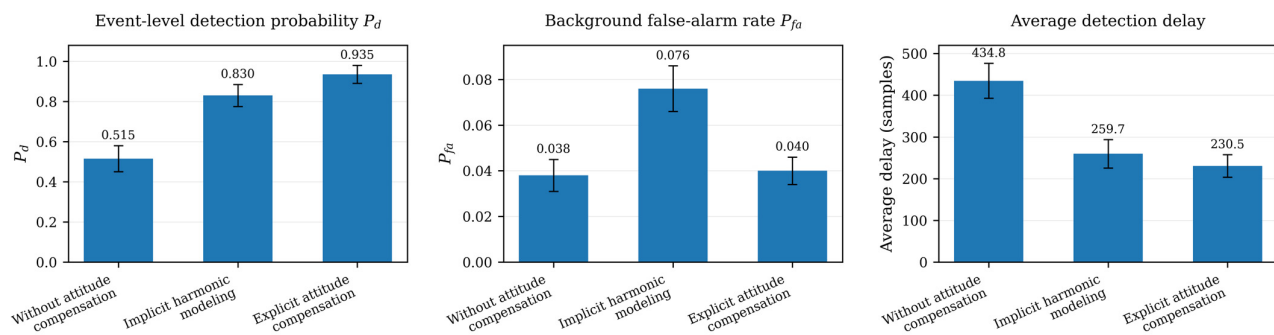
Figure 4. Ablation results of the multiscale fusion module at $\text{SNR} \approx 2.7$.

4.3.2. Independent Contribution of Attitude Compensation

Table 7 shows that, without attitude compensation, geomagnetic projection variations directly contaminate the triaxial observations, resulting in a low detection probability of $P_d = 0.515$ and a long average detection delay of 434.8 sampling points. Modeling attitude motion as an implicit harmonic background improves the detection probability to $P_d = 0.830$ but increases the false-alarm rate to $P_{fa} = 0.076$, indicating that this approach may mix platform-motion components with target residuals. In contrast, explicit attitude compensation achieves a detection probability of $P_d = 0.935$ while reducing the false-alarm rate to $P_{fa} = 0.040$. These results demonstrate that physical rotation correction is an important front-end module for suppressing attitude-induced geomagnetic projection interference and controlling false alarms under real buoy background conditions. The corresponding attitude-processing ablation results are also illustrated in Figure 5.

Table 7. Attitude compensation ablation experiment at SNR ≈ 2.7 .

Attitude Processing	P_d	P_{fa}	Delay/Samples
Without attitude compensation	0.515	0.038	434.8
Implicit harmonic modeling	0.830	0.076	259.7
Explicit attitude compensation	0.935	0.040	230.5

**Figure 5.** Detection performance comparison under different attitude processing modes at SNR ≈ 2.7 . Error bars denote 95% confidence intervals of the mean over 200 Monte Carlo trials.

4.4. Parameter Sensitivity and Detection–False Alarm Trade-Off

As summarized in Table 8, the drift parameter γ mainly controls the trade-off between false-alarm rate and detection delay. A smaller γ reduces the detection delay but significantly increases the false-alarm rate, whereas a larger γ suppresses false alarms at the cost of reduced detection probability and slower response. When the online NIS quantiles vary within a reasonable range, the performance changes only slightly, indicating that the online threshold adaptation and target-dwell-protection mechanism are robust to moderate quantile variations. For the window-length configuration, longer windows help shorten the detection delay and improve the detection probability, but they may also increase false alarms caused by low-frequency background accumulation. Therefore, $\gamma = 1.5$ and the 3/10/30 sampling-point window combination are retained as the main low-false-alarm configuration in this study.

Table 8. Parameter sensitivity at SNR ≈ 2.7 .

Parameter	Parameter Variation	Result Summary
Drift parameter γ	$\gamma = 1.0 \rightarrow 2.0$	P_d : 0.970 $\rightarrow$ 0.925; P_{fa} : 0.092 $\rightarrow$ 0.039; Delay: 173.3 $\rightarrow$ 232.3.
Online NIS quantiles p_b / p_t	$p_b = 97.5\% \rightarrow 99.2\%, p_t = 99.5\% \rightarrow 99.8\%$	P_d, P_{fa} , and delay change slightly
Window lengths $W_s/W_M/W_L$	3/10/30 $\rightarrow$ 10/30/90	$P_d \approx 0.935 - 0.965$; $P_{fa} \approx 0.038 - 0.077$; Delay $\approx 197.8 - 229.2$.

To make the parameter-setting procedure explicit, the final values were selected according to the functional role of each parameter rather than by re-tuning each test case. The statistical thresholds were calibrated from target-free background data: the NIS quantiles were set to $p_b = 99.0\%$ and $p_t = 99.7\%$, and the dual detection thresholds were set at the 99.55% and 99.85% quantiles. The temporal and fusion parameters were selected using the target-free calibration results and the sensitivity analyses reported in the manuscript. Specifically, the 3/10/30-sample windows and 0.10/0.80/0.10 fusion weights were retained to balance short-term response, persistent evidence accumulation, and false-alarm control; $\alpha_t = 0.02$ and $\Delta_t = 1.0$ were used to avoid abrupt threshold changes; and $D_{\max} = 4$, $C_r = 4$, and $\mu_{\min} = 0.05$ were used to limit rapid recovery of the

protection gate during suspected-target intervals. These values were applied consistently to the reported test cases. If the sea state, sensor noise, or buoy platform changes, the calibration-dependent thresholds and covariance-related quantities should be recalibrated using representative target-free data, whereas the processing structure and multiscale fusion logic can be retained. Thus, the method is transferable at the algorithmic level, but the numerical parameter values are not claimed to be universally optimal.

As shown in Figure 6, at a fixed SNR of approximately 2.7, different methods exhibit distinct detection–false alarm operating characteristics as the detection threshold varies. The standard KF single-scale and adaptive KF single-scale methods are generally located in the lower detection-probability region, indicating that single-scale residual accumulation has limited ability to preserve weak target responses in real non-stationary buoy backgrounds. The minimum-entropy method maintains a relatively low false-alarm rate, but its detection probability remains limited, showing conservative detection behavior. The OBF and SR-CUSUM baselines achieve higher detection probabilities than the KF-based and entropy-based baselines; however, their operating points tend to move toward higher false-alarm rates when higher detection probabilities are obtained. In contrast, the proposed method achieves a high detection probability under a low false-alarm rate, providing a more favorable operating point in the detection–false alarm trade-off. These results indicate that online target-protected filtering, target-free calibration-based covariance whitening, and dual-threshold multiscale fusion can effectively improve weak-target detectability while maintaining false-alarm suppression under real buoy background conditions.

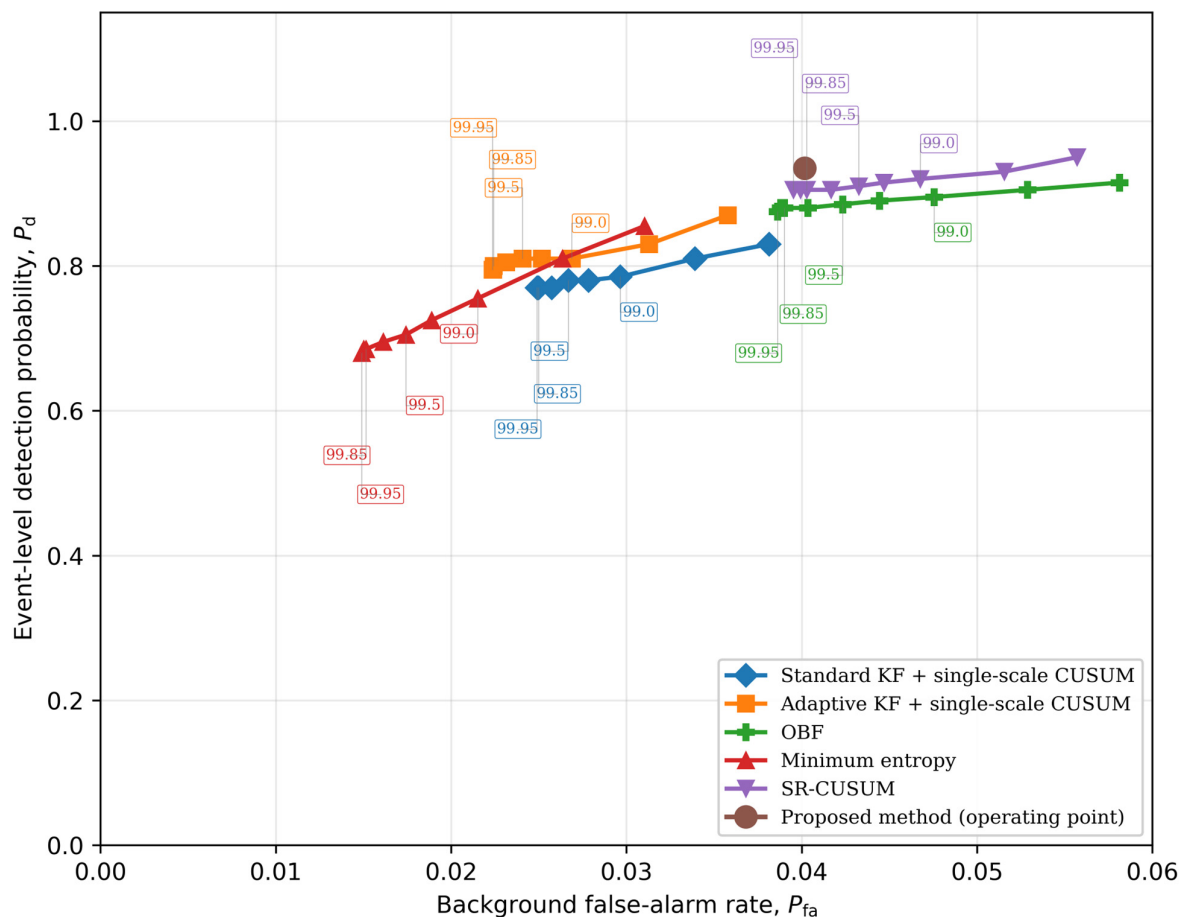


Figure 6. Detection–false alarm trade-off curves of different methods at SNR $\approx$ 2.7.

4.5. Robustness to Heading and Geomagnetic Location

To further examine the adaptability of the proposed method to target-heading variations, comparative experiments were conducted under fixed DataS2, V1 geomagnetic location, and $\text{SNR} \approx 2.7$. Three target-heading conditions, namely KD0, PD0, and SD0, were selected, and 200 Monte Carlo trials were performed for each condition. The corresponding results are summarized in Table 9 and Figure 7.

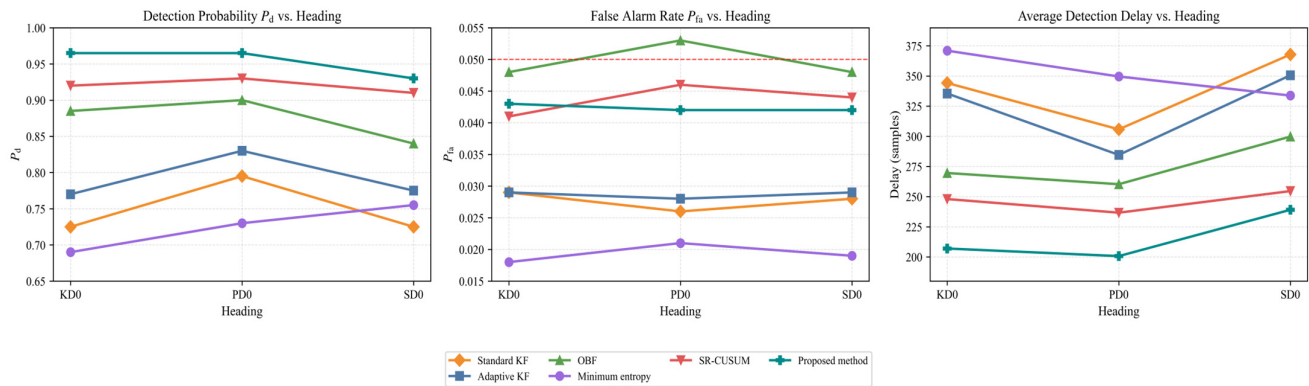


Figure 7. Performance comparison of the six detection methods under different target headings at DataS2, V1, and $\text{SNR} \approx 2.7$. The dashed horizontal line in the false-alarm-rate panel denotes the $P_{fa} = 0.05$ reference level.

Table 9. Robustness comparison of different methods under different target heading conditions.

Heading	Method	P_d	P_{fa}	Delay/Samples
KD0	Standard KF	0.725	0.029	344.2
	Adaptive KF	0.770	0.029	335.5
	OBF	0.885	0.048	269.6
	Minimum entropy	0.690	0.018	371.0
	SR-CUSUM	0.920	0.041	248.0
	Proposed method	0.965	0.043	207.0
PD0	Standard KF	0.795	0.026	305.8
	Adaptive KF	0.830	0.028	284.6
	OBF	0.900	0.053	260.3
	Minimum entropy	0.730	0.021	349.6
	SR-CUSUM	0.930	0.046	236.7
	Proposed method	0.965	0.042	200.7
SD0	Standard KF	0.725	0.028	367.8
	Adaptive KF	0.775	0.029	350.6
	OBF	0.840	0.048	299.8
	Minimum entropy	0.755	0.019	333.8
	SR-CUSUM	0.910	0.044	254.6
	Proposed method	0.930	0.042	239.1

To evaluate the adaptability of the proposed method to target-heading variations, comparative experiments were conducted using the DataS2_V1 dataset under a fixed equivalent SNR of approximately 2.7. Three target-heading conditions, namely KD0, PD0, and SD0, were considered, and 200 Monte Carlo trials were performed for each case. As shown in Table 9, the proposed method achieved event-level detection probabilities of 0.965, 0.965, and 0.930 under KD0, PD0, and SD0, respectively. The corresponding background false-alarm rates were 0.043, 0.042, and 0.042, while the average detection delays

were 207.0, 200.7, and 239.1 samples. Compared with the standard KF and adaptive KF single-scale methods, the proposed method consistently improves detection probability and reduces detection delay under all heading conditions. Compared with OBF, it provides higher detection probability and faster response while maintaining a comparable false-alarm level. Compared with the minimum-entropy method, the proposed method has a slightly higher false-alarm rate in some cases, but its improvements in detection probability and detection delay are more significant. Compared with SR-CUSUM, the proposed method achieves higher detection probabilities under all heading conditions and shorter detection delays while maintaining a similar background false-alarm rate. Although the detection probability decreases slightly under the SD0 heading, the proposed method still maintains the best overall balance among detection probability, false-alarm control, and response speed. These results demonstrate that the online CFAR-NIS target-protected filtering and dual-threshold weighted multiscale GLRT-CUSUM mechanism has good robustness to target-heading variations.

To further evaluate the adaptability of the proposed algorithm to different geomagnetic intensities and inclination conditions, DataS2, the KD0 track direction, and $\text{SNR} \approx 2.7$ were fixed, and the same Monte Carlo evaluation procedure was conducted at six geomagnetic locations: V1–V6. The corresponding results are summarized in Table 10.

Table 10. Robustness experiment of the proposed method under different geomagnetic locations.

Location	Corresponding Region	Geomagnetic Intensity/ μT	P_d	P_{fa}	Delay/Samples
V1	North Sea 55°N/2°E	50.0	0.965	0.042	207.0
V2	Arctic 90°N	52.5	0.955	0.042	208.8
V3	Southeast Asia 15°N	37.0	0.985	0.040	173.5
V4	Near the equator 10°N	30.0	0.985	0.042	158.9
V5	Southern Australia 25°S	55.0	0.980	0.038	174.0
V6	Near Antarctica 63°S	67.0	0.975	0.041	183.0

Table 10 shows that, across the six geomagnetic locations V1–V6, the detection probability of the proposed method remains within the range of 0.955–0.985, while the average detection delay ranges from 158.9 to 208.8 sampling points. The background false-alarm rate remains below 0.042 under all V1–V6 conditions. Overall, the proposed method maintains high detection probability and low false-alarm levels under different geomagnetic intensities and inclination conditions. These results indicate that explicit attitude compensation effectively reduces the influence of geomagnetic projection variations on triaxial observations, while target-free calibration-based covariance whitening improves robustness to changes in background correlation at different geomagnetic locations. Moreover, the online CFAR-NIS target-protected filtering and dual-threshold weighted multiscale GLRT-CUSUM fusion mechanism maintains a favorable detection–false alarm trade-off across different geomagnetic environments.

4.6. Equivalent Passing-Distance Evaluation

To further provide an application-oriented interpretation of the low-SNR detection capability of the proposed method, an equivalent passing-distance experiment was conducted. Based on the magnetic field attenuation relationship $B(r) \propto 1/r^3$, the 80 m case was used as the reference condition, corresponding to an equivalent $\text{SNR} \approx 2.70$. The target magnetic amplitude was then scaled to simulate equivalent passing distances ranging from 60 m to 120 m. Detection probability, background false-alarm rate, and average detection delay were evaluated under the same Monte Carlo configuration.

As shown in Figure 8, the detection probabilities of all methods generally decrease as the equivalent passing distance increases, whereas the average detection delays increase accordingly. Table 11 further summarizes the equivalent passing-distance results of the proposed method. Under the constraint of $P_d \geq 0.90$ and $P_{fa} \leq 0.05$, the maximum reliable equivalent passing distance of the proposed method is approximately 80.0 m. At this distance, the corresponding equivalent SNR is approximately 2.70, the target peak amplitude is approximately 54.00 nT, the detection probability is 0.955, the background false-alarm rate is 0.040, and the average detection delay is 220.7 sampling points, corresponding to 11.04 s. When the distance exceeds 80 m, the detection probability decreases to 0.845 at 90 m and 0.710 at 120 m, while the false-alarm rate remains stable at approximately 0.040–0.041. These results indicate that the proposed method maintains stable low-false-alarm behavior under distance-dependent signal attenuation and that the reliable equivalent-distance boundary under the adopted semi-physical injection setting is approximately 80 m.

Table 11. Equivalent passing-distance results of the proposed method.

Distance (m)	Eq. SNR	Target Peak Amplitude/nT	P_d	P_{fa}	Delay/Samples
60.0	6.400	128.00	1.000	0.039	146.1
70.0	4.030	80.61	1.000	0.040	169.1
80.0	2.700	54.00	0.955	0.040	220.7
90.0	1.896	37.93	0.845	0.040	273.5
100.0	1.382	27.65	0.780	0.041	311.9
110.0	1.039	20.77	0.740	0.041	328.3
120.0	0.800	16.00	0.710	0.041	338.4

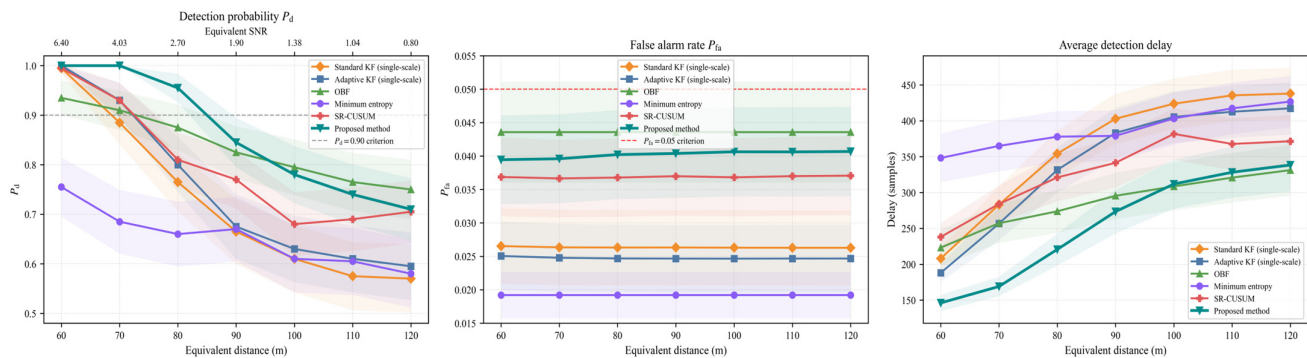


Figure 8. Equivalent passing-distance evaluation under the magnetic field attenuation assumption $B(r) \propto r^{-3}$. The 80 m reference distance corresponds to $SNR \approx 2.70$, and the shaded bands denote 95% confidence intervals.

Although the 80 m equivalent-distance case corresponds to an equivalent SNR of approximately 2.70, its numerical results are not identical to those of the main SNR experiment. This difference occurs because the equivalent passing-distance experiment was conducted as an independent Monte Carlo evaluation based on distance-dependent amplitude scaling, whereas the main SNR experiment was organized using predefined SNR levels. Variations in randomly selected real-background segments, target injection positions, and finite Monte Carlo sampling may lead to slight differences in P_d , P_{fa} , and average delay. Therefore, the 80 m result should be interpreted as a consistent reference condition under the equivalent-distance evaluation framework, rather than a duplicate of the main $SNR = 2.7$ experiment.

5. Discussion

5.1. Mechanism Interpretation

The performance improvement of the proposed method mainly comes from the separation between background adaptation and target evidence preservation. The key methodological difference from prior adaptive-filtering-based magnetic tracking approaches [13] is not the use of an adaptive Kalman filter alone but the way in which the adaptive update is constrained during suspected-target intervals. In conventional adaptive filtering, anomalous innovations may be gradually incorporated into the background estimate when the target response is weak, temporally extended, or intermittently below a fixed consistency threshold. In contrast, the proposed online CFAR-NIS mechanism estimates the background-consistency and target-protection thresholds from a rolling alarm-free background window and couples them with a continuous target-dwell counter. This dwell-based memory delays the recovery of the gating weight until a sustained sequence of low-NIS samples is observed. More importantly, the gating weight is applied jointly to the covariance update and the effective Kalman gain, so that both the adaptive noise statistics and the state update are constrained during suspected target passages. As a result, weak magnetic anomaly residuals are less likely to be absorbed into the adaptive background estimate and can be preserved for subsequent GLRT-CUSUM evidence accumulation. The subsequent target-free covariance whitening and normalized multiscale GLRT-CUSUM fusion further transform the protected residuals into comparable detection evidence across different response durations.

The proposed detector also differs from conventional GLRT-based magnetic anomaly detectors [22] in its treatment of the detection covariance. Instead of directly using the online adaptive covariance for detection, the adaptive covariance is used only for recursive background filtering and gain control, whereas a separate target-free calibration covariance is used for residual whitening and GLRT statistic construction. This separation reduces the risk that the detection statistic is biased by the covariance adaptation process itself. After whitening, the normalized multiscale GLRT-CUSUM fusion and the dual-threshold candidate-confirmation rule further improve robustness against short impulsive disturbances and low-frequency background fluctuations.

5.2. Practical Applicability

The supplementary experiments on target heading and geomagnetic location further demonstrate that the proposed method can maintain high detection probability under different target track directions and geomagnetic backgrounds. However, the false-alarm rate still exhibits slight fluctuations across different geomagnetic locations, indicating that online threshold calibration should be combined with local target-free background data in practical deployments. The online CFAR-NIS threshold estimation introduced in this study updates $\tau_b(k)$ and $\tau_t(k)$ using an alarm-free background window, thereby reducing the mismatch between fixed χ^2 -based thresholds and the actual NIS distribution under locally stationary or slowly drifting backgrounds. In addition, the target-dwell counter delays the recovery of the gating weight, which helps alleviate residual absorption during continuous target passage.

From the perspective of computational complexity, the proposed detector has a low-dimensional structure compatible with future online embedded implementation because all matrix operations are performed in low-dimensional spaces. Here, n_x and m denote the background-state and observation dimensions, respectively; $m = 3$ for the triaxial observation. The dominant per-sample cost comes from the target-protected adaptive Kalman filtering step, whereas whitening and GLRT construction only require a 3×3 Cholesky decomposition, a triangular solve, and a squared-norm calculation. The

multiscale CUSUM branches and dual-threshold fusion are scalar recursive operations. Since the sampling interval is 0.05 s, the detector only needs to complete one iteration within 50 ms. Table 12 summarizes the per-sample computational complexity of the main online processing steps.

Table 12. Per-sample computational complexity of the proposed online detector.

Processing Step	Main Operation	Complexity
Attitude compensation	3×3 rotation/subtraction	$O(m^2)$
Innovation and NIS	Covariance and quadratic form	$O(n_x^2 m + n_x m^2 + m^3)$
Protected Kalman update	Gain, covariance, state update	$O(n_x^3 + n_x^2 m + n_x m^2 + m^3)$
CFAR-NIS threshold update	Sliding-window quantile	$O(W_0 \log W_0)$; lower if incremental
Dwell memory and gating	Scalar comparison/exponential gate	$O(1)$
Whitened GLRT statistic	3×3 Cholesky solve and norm	$O(m^3)$
CUSUM accumulation and fusion	Scalar recursions and threshold tests	$O(Q)$
Overall online detector	Dominated by low-dimensional Kalman filtering	$O(n_x^3 + n_x^2 m + n_x m^2 + m^3)$

To assess the online processing burden, the proposed method was benchmarked using the final Python implementation on a Lenovo laptop equipped with a 13th Gen Intel Core i7-13650HX CPU and 24 GB DDR5-4800 memory. The GPU was not used. Across 20 repetitions of a 3000-sample sequence, the mean, median, and 95th-percentile batch-processing times were 297.8, 297.3, and 303.2 ms, respectively. These values correspond to equivalent mean and 95th-percentile processing times of 99.3 and 101.1 μ s per sample, with a throughput of approximately 10,091 samples/s. At the nominal sampling rate of 20 Hz, the measured throughput is approximately 505 times the required input rate. Because the benchmark processes the sequence as a batch, the batch duration is not directly compared with the 50 ms sampling interval; the throughput comparison is used to assess the tested CPU reference implementation. The 0.401 MB tracemalloc value represents Python-level allocations within the timed function rather than the complete process memory footprint and is therefore not interpreted as the total RAM requirement. Hardware-level latency and memory consumption on a target embedded processor were not measured and remain to be verified in future work.

To complement the runtime benchmark, the streaming working-memory requirement can be expressed at the algorithm level. Let W_0 denote the length of the online NIS background window, $|Q|$ denote the cardinality of the multiscale CUSUM branch set, n_x denote the background-state dimension, and m denote the observation dimension. The branch-set cardinality $|Q|$ is distinct from the process-noise covariance $\mathbf{Q}_{k-1}$ used in the state-space model. Excluding program code and stored input/output records, the algorithm-level working-memory complexity is $O(W_0 + |Q| + n_x^2 + m^2)$. These terms represent the rolling window, branch states, state-space covariance matrices, and triaxial covariance/whitening matrices. For the reported configuration, $W_0 = 1000$, $|Q| = 3$, $n_x = 6$, and $m = 3$. This is an algorithm-level estimate, not an embedded-device measurement; target RAM, latency, and power remain to be verified with a dedicated C/CMSIS-DSP implementation.

5.3. Limitations and Future Work

The results of this study should be interpreted as semi-physical validation based on real buoy backgrounds and injected publicly available ship magnetic signatures, rather than as complete sea-trial validation. The real background data were obtained from

DOE/A2E buoy attitude and triaxial magnetometer records; therefore, the background disturbance, attitude variation, and triaxial correlation are supported by measured data. However, the target signal was generated by injecting publicly available numerical ship magnetic signatures [29,30]. Therefore, the conclusions demonstrate the advantage of the proposed method as a representative training-free statistical baseline under real buoy background conditions, but they do not directly verify the detection range in actual maritime operations or the robustness of the method under all sea states.

Failure cases and numerical stability should also be considered. The detector is expected to become less reliable when the target amplitude is attenuated below the calibrated detection range or when background disturbances are not sufficiently represented by the target-free calibration data. In the equivalent passing-distance experiment (Table 11), the detection probability decreased from 0.955 at 80 m ($\text{SNR} \approx 2.70$) to 0.845 at 90 m ($\text{SNR} \approx 1.896$) and 0.710 at 120 m ($\text{SNR} \approx 0.800$), illustrating the degradation of performance under the tested low-SNR conditions. Short isolated spikes are intended to be rejected by the dual-threshold confirmation rule, whereas weak persistent targets may produce delayed or missed alarms if their accumulated evidence does not reach the confirmation threshold. Given the implemented bounds on $0 \leq \mu_{\min} \leq 1$, $\beta_r \geq 0$ and $\beta_c \geq 0$ and the projection of R_k onto prescribed covariance bounds, the recursive update remains numerically bounded in the tested implementation. These constraints help limit numerical excursions but do not constitute a formal convergence proof for arbitrary non-stationary sea states. Further stability and long-term false-alarm behavior should therefore be evaluated using real sea-trial data.

An additional limitation is that the present experiments do not include a supervised deep learning detector. Although the semi-physical construction provides target/background labels for evaluation, it does not yet provide the diversity and volume of real labeled sea-trial data needed for a fair training-based comparison. Accordingly, the present results should not be interpreted as evidence that the proposed statistical detector is universally superior to AI-based methods. Future work will establish matched training, validation, and test protocols using real operational data and compare the proposed method with representative learning-based detectors under the same false-alarm and latency constraints.

Future work will focus on validating the proposed framework using sea-trial data with real target passages, extending the evaluation to different ship types and sea states, and examining the influence of buoy deployment depth, sensor installation error, and multi-source environmental disturbances on long-term false-alarm stability. Hardware-level timing, memory use, and power consumption on a target embedded processor were not measured in this study and remain future work.

6. Conclusions

This paper proposed an online threshold-adaptive target-protected filtering method combined with dual-threshold weighted multiscale GLRT-CUSUM for ocean-buoy-based vector magnetic anomaly detection. The proposed method first explicitly compensates for geomagnetic projection variations using an attitude rotation matrix. Then, an online CFAR-NIS target-protection mechanism is introduced to reduce the absorption of suspected target innovations by the adaptive filter. In this mechanism, the NIS background-consistency threshold and target-protection threshold are estimated online from an alarm-free background window using empirical quantiles, while a target-dwell counter delays the recovery of the gating weight during continuous suspected-target intervals. Subsequently, a target-free calibration segment is used to estimate a calibration-based detection covariance and construct a whitened GLRT statistic. Finally, alarms are generated through

CFAR-normalized weighted multiscale evidence fusion and a dual-threshold candidate-confirmation rule.

Semi-physical experiments based on real DOE/A2E buoy backgrounds and injected publicly available ferromagnetic ship magnetic signatures show that, at $\text{SNR} \approx 2.7$ with 200 Monte Carlo trials, the proposed method achieves $P_d = 0.965$, $P_{fa} = 0.045$, and an average delay of 208.3 sampling points. Compared with standard KF, adaptive KF, OBF, minimum-entropy, and SR-CUSUM baselines, the proposed method provides a better overall balance between weak-target detection probability, false-alarm suppression, and response delay. The results indicate that protecting suspected target residuals before sequential GLRT-CUSUM evidence accumulation is critical for low-SNR buoy-based magnetic anomaly detection.

Supplementary experiments further show that the proposed method maintains high detection probability under the three heading conditions KD0, PD0, and SD0, and its detection performance remains generally stable across different geomagnetic locations: V1–V6. These results indicate that the proposed method has good robustness to target heading and geomagnetic-location variations under the adopted semi-physical experimental framework. Future work will introduce real target-passage sea-trial data, multiple headings and ship types, different buoy deployment sites, and cross-scenario learning-based baselines. In addition, the stability of online threshold estimation and target-dwell-protection mechanisms will be further verified under long-term buoy deployment conditions.

Nevertheless, the present validation is semi-physical, and the numerical thresholds and covariance-related parameters should be recalibrated using representative target-free data when the sea state, sensor noise, or buoy platform changes. Validation with real operational sea-trial data remains necessary to assess the generalization and long-term performance of the proposed method.

Author Contributions: Conceptualization, Y.Y., P.W. and K.Z.; methodology, P.W., Y.Y. and K.Z.; software, P.W.; validation, P.W., Y.Y., K.Z., J.Z. and S.L.; formal analysis, P.W., Y.Y. and K.Z.; investigation, Y.Y., P.W., J.Z. and K.Z.; resources, K.Z., Q.K., B.W., Z.Q. and X.Y.; data curation, P.W., K.Z., S.L., X.F. and J.Z.; writing—original draft preparation, P.W., Y.Y. and K.Z.; writing—review and editing, K.Z., Y.Y., P.W., J.Z., Q.K., Z.Q. and X.Y.; visualization, P.W. and K.Z.; supervision, K.Z., Q.K., Z.Q. and X.Y.; project administration, K.Z., Q.K. and X.Y.; funding acquisition, K.Z. and Q.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Major Innovation Project of Science, Education and Industry Integration Pilot Project of Qilu University of Technology (Shandong Academy of Sciences) (Grant No. 2025ZDYS01); the Laoshan Laboratory (Grant No. LSKJ202502400); the Major Scientific Research Project for the Construction of State Key Laboratory at Qilu University of Technology (Shandong Academy of Sciences) (Grant No. 2025ZDZG01); the Key Research and Development Program of Shandong Province, China (Grant Nos. 2025CXPT184 and 2025CXGC010711); and the Taishan Industrial Program “Marine Observation and Detection Buoy Equipment R&D and Industrialization Team” (Grant No. 2024).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The real ocean-buoy background data used in this study were obtained from the publicly available U.S. Department of Energy Wind Data Hub/Atmosphere to Electrons (A2E) Data Archive and Portal. These data include buoy attitude records and triaxial magnetometer background measurements used for attitude compensation, background construction, and semi-physical injection experiments. The ferromagnetic ship magnetic signature data used for target injection were obtained from the public dataset “Magnetic signature reproduction of ferromagnetic

ships at arbitrary geographical position, direction and depth using a multi-dipole model—source and verification dataset with description”, provided by Gdańsk University of Technology, DOI: 10.34808/5zp7-3z19. No access restrictions apply to the above public source datasets. The processed background segments, semi-physical target-injection data, parameter settings, numerical results, and Python scripts used for preprocessing, threshold calibration, Monte Carlo evaluation, and figure generation are available from the corresponding author upon reasonable request.

Acknowledgments: The authors would like to thank the providers of the public DOE/A2E buoy dataset and the public ferromagnetic ship magnetic signature dataset.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MAD	Magnetic anomaly detection
GLRT	Generalized likelihood ratio test
CUSUM	Cumulative sum
GLRT-CUSUM	Generalized likelihood ratio test-cumulative sum
CFAR	Constant false-alarm rate
NIS	Normalized innovation squared
CFAR-NIS	Constant-false-alarm-rate normalized innovation squared
KF	Kalman filter
AKF	Adaptive Kalman filter
OBF	Orthogonal basis function
SNR	Signal-to-noise ratio
IMU	Inertial measurement unit
DOE	U.S. Department of Energy
A2E	Atmosphere to Electrons
DOE/A2E	U.S. Department of Energy Atmosphere to Electrons program
MC	Monte Carlo
KD0	Keel-direction track condition
PD0	Port-direction track condition
SD0	Starboard-direction track condition
SR-CUSUM	Spectral residual-cumulative sum

References

1. Wang, C.; Cui, Y.; Song, X.; Zhang, Y.; Yuan, Z. A novel underwater target detection method based on low-frequency magnetic signal generated by portable transmitter. *IEEE Sens. J.* **2023**, *23*, 8459–8465. <https://doi.org/10.1109/JSEN.2023.3253708>.
2. Li, H.; Luo, J.; Zhang, J.; Li, J.; Zhang, Y.; Zhang, W.; Zhang, M. Determinants of maximum magnetic anomaly detection distance. *Sensors* **2024**, *24*, 4028. <https://doi.org/10.3390/s24124028>.
3. Dai, F.; Peng, D.; Chen, Z.; Liu, B. Magnetic anomaly detection method based on principal component analysis and cascaded stochastic resonance. *Rev. Sci. Instrum.* **2026**, *97*, 015201. <https://doi.org/10.1063/5.0274445>.
4. Zhao, Y.; Zhang, J.H.; Li, J.H.; Liu, S.; Miao, P.X.; Shi, Y.C.; Zhao, E.M. A brief review of magnetic anomaly detection. *Meas. Sci. Technol.* **2021**, *32*, 042002. <https://doi.org/10.1088/1361-6501/abd055>.
5. Wang, H.; Dong, X.; Yang, Y. Influence of marine environmental factors on characteristics of composite magnetic field of underwater vehicles. *J. Mar. Sci. Eng.* **2025**, *13*, 1850. <https://doi.org/10.3390/jmse13101850>.
6. Oehler, J.-F.; Schifano, V.; Marquis, G.; Reiller, H.; Lucas, S.; Bougeault, C. Processing of unmanned underwater vehicle vector magnetometer data. *Geophys. Prospect.* **2024**, *72*, 2902–2918. <https://doi.org/10.1111/1365-2478.13556>.
7. Tong, Y.; Huang, X.; Chen, Y.; Li, W.; Zha, F. A compensation method for the geomagnetic measurement error of an underwater ship-borne magnetometer based on constrained total least squares. *Sensors* **2024**, *24*, 3478. <https://doi.org/10.3390/s24113478>.
8. Sun, H.; Wang, Z.; Lan, T.; Zhang, S.; Qiu, J. Adaptive cascade detection of weak magnetic anomalies based on marine predators algorithm-stochastic resonance. *AIP Adv.* **2023**, *13*, 025147. <https://doi.org/10.1063/9.0000593>.

9. Wang, Z.; Zheng, E.; Liu, J.; Guo, T. Adaptive Orthogonal Basis Function Detection Method for Unknown Magnetic Target Motion State. *Appl. Sci.* **2024**, *14*, 902. <https://doi.org/10.3390/app14020902>.
10. Qin, Y.; Li, K.; Yao, C.; Wang, X.; Ouyang, J.; Yang, X. Magnetic anomaly detection using full magnetic gradient orthonormal basis function. *IEEE Sens. J.* **2020**, *20*, 12928–12940. <https://doi.org/10.1109/JSEN.2020.3003680>.
11. Tzannetou, L.; Bakoglou, E.; Priftis, P.; Angelopoulos, S.; Ktena, A.; Hristoforou, E. Magnetic Anomaly Detection of Moving Objects. *Phys. B Condens. Matter* **2024**, *676*, 415659. <https://doi.org/10.1016/j.physb.2023.415659>.
12. Li, W.; Zhou, Z.; Li, H.; Tang, J.; Xu, M. Optimization of Magnetic Anomaly Detection Under $1/f^\alpha$ Noise Based on Karhunen–Loève Expansion and Frequency Characteristic Function. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 9516813. <https://doi.org/10.1109/TIM.2025.3557098>.
13. Sun, M.; Hodgskin-Brown, R.; Davies, M.E.; Proudler, I.K.; Hopgood, J.R. Adaptive kernel Kalman filter for magnetic anomaly detection-based metallic target tracking. In *Proceedings of the 2023 Sensor Signal Processing for Defence Conference (SSPD), Edinburgh, UK, 12–13 September 2023*; IEEE: Piscataway, NJ, USA, 2023. <https://doi.org/10.1109/SSPD57945.2023.10256968>.
14. Jiang, S.; Chen, Y.; Liu, Q. Advancements in buoy wave data processing through the application of the Sage-Husa adaptive Kalman filtering algorithm. *Sensors* **2023**, *23*, 7298. <https://doi.org/10.3390/s23167298>.
15. Lv, B.; Qin, Y.; Dai, H.; Han, J. A differential signal and extended-Kalman-filter-based anti-interference magnetic tracking method for surgical scenes. *IEEE Trans. Ind. Electron.* **2024**, *71*, 15140–15150. <https://doi.org/10.1109/TIE.2024.3366214>.
16. Yan, Z.; Yu, H.; Leng, Z.; Yao, Y.; Zhao, X.; Wang, Z. Comprehensive Interference Estimation and Correction Methods Based on Unscented Kalman Filter for Magnetic Anomaly Detection. *Measurement* **2025**, *248*, 116858. <https://doi.org/10.1016/j.measurement.2025.116858>.
17. Chen, Z.; Lou, Y.; He, P.; Xu, P.; Zhang, X. Magnetic anomaly detection based on Attention-Bi-LSTM network. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 2520011. <https://doi.org/10.1109/TIM.2024.3403210>.
18. Meng, Y.; Zhang, X.; Xu, Z.; Li, H.; Zhang, J.; Li, J.; Zhang, Y.; Zhang, W. Underwater Target Detection Based on Magnetic Gradient Tensor. In *Proceedings of the OCEANS 2024—Singapore, Singapore, 15–18 April 2024*; IEEE: Piscataway, NJ, USA, 2024; pp. 1–6. <https://doi.org/10.1109/OCEANS51537.2024.10706271>.
19. Gastaldo, P.; Zunino, R.; Bellesi, A.; Carbone, A.; Gemma, M.; Ragusa, E. A two-dimensional (2-D) sensor network architecture with artificial intelligence models for the detection of magnetic anomalies. *Sensors* **2026**, *26*, 764. <https://doi.org/10.3390/s26030764>.
20. Wang, H.; Jiang, C.; Feng, Z. Research on the combined detection of magnetic anomaly and shaft-rate magnetic field signals. *IET Radar Sonar Navig.* **2025**, *19*, e70091. <https://doi.org/10.1049/rsn2.70091>.
21. Dong, X.; Wang, H.; Yang, Y.; Chen, Y. Spatiotemporal characteristics of submarine magnetic anomaly under typical motion scenarios. *J. Mar. Sci. Appl.* **2025**. <https://doi.org/10.1007/s11804-025-00761-x>.
22. Pepe, P.; Zozor, S.; Rouve, L.-L.; Coulomb, J.-L.; Serviere, C.; Muley, J. Generalization of GLRT-Based Magnetic Anomaly Detection. In *Proceedings of the 2015 23rd European Signal Processing Conference (EUSIPCO), Nice, France, 31 August–4 September 2015*; IEEE: Piscataway, NJ, USA, 2015; pp. 1930–1934. <https://doi.org/10.1109/EUSIPCO.2015.7362720>.
23. Kay, S.M. *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*; Prentice Hall: Upper Saddle River, NJ, USA, 1998; ISBN 978-0-13-504135-2.
24. Xu, C.; Chen, J.; Fang, J.; Xiao, K.; Xu, Y.; Chen, J.; Wei, C. Real-time anomaly detection for submarine cables based on multi-variate CUSUM. In *Proceedings of the 2025 23rd International Conference on Optical Communications and Networks (ICOON), Zhangjiajie, China, 28–31 July 2025*; IEEE: Piscataway, NJ, USA, 2025; pp. 1–3. <https://doi.org/10.1109/ICOON67308.2025.11145409>.
25. Page, E.S. Continuous Inspection Schemes. *Biometrika* **1954**, *41*, 100–115. <https://doi.org/10.1093/biomet/41.1-2.100>.
26. Basseville, M.; Nikiforov, I.V. *Detection of Abrupt Changes: Theory and Application*; Prentice Hall: Englewood Cliffs, NJ, USA, 1993; ISBN 978-0-13-126780-8.
27. Hou, X.; Zhang, L. Saliency detection: A spectral residual approach. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Minneapolis, MN, USA, 17–22 June 2007*; IEEE: Piscataway, NJ, USA, 2007; pp. 1–8. <https://doi.org/10.1109/CVPR.2007.383267>.
28. Ren, H.; Xu, B.; Wang, Y.; Yi, C.; Huang, C.; Kou, X.; Xing, T.; Yang, M.; Tong, J.; Zhang, Q. Time-Series Anomaly Detection Service at Microsoft. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019*; IEEE: Piscataway, NJ, USA, 2019; pp. 3009–3017. <https://doi.org/10.1145/3292500.3330680>.

29. Wołoszyn, M.; Tarnawski, J. Magnetic signature reproduction of ferromagnetic ships at arbitrary geographical position, direction and depth using a multi-dipole model. *Sci. Rep.* **2023**, *13*, 14601. <https://doi.org/10.1038/s41598-023-41702-4>.
30. Wołoszyn, M.; Tarnawski, J. Magnetic signature reproduction of ferromagnetic ships at arbitrary geographical position, direction and depth using a multi-dipole model-source and verification dataset with description. *Gdansk University of Technology* 2023, Dataset. Available online: <https://mostwiedzy.pl/en/open-research-data/magnetic-signature-reproduction-of-ferromagnetic-ships-at-arbitrary-geographical-position-direction-,102011752152696-0> (accessed on 2 August 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.