Data Engineering and Applications

Research Paper

RWKV-TS for Underwater Robot Station Keeping from Ocean-Current-Disturbed Navigation Data

Ghassan Al-Marri[1, *] and Hadi Al-Suhaimi[1]

[1] College of Technological Innovation, Zayed University, Dubai, PO Box 4783, United Arab Emirates
*Corresponding author: ghassan.mar@zu.ac.ae

Abstract. In this study, a linear-time recurrent sequence model called RWKV-TS is used to map current-disturbed navigation histories to constrained thruster increments for horizontal-plane station-keeping. Doppler velocity, body acceleration, previous actuation, inertial position and orientation, and a causal current residual are the inputs. A channel-mixing gate differentiates between a short impulsive perturbation and a steadily increasing current bias, while a time-mixing block achieves a lengthy disturbance history without quadratic attention. A deterministic safety solution, which involves adding the learned increment to a low-gain stabilizing controller and using a saturation-aware allocator, will be applied if the sequence estimate is not provided. 63.4 hours of tank and coastal data from 18 flights in sheared, oscillatory, turbulent, and stable currents were analyzed. RWKV-TS is compared with PID, sliding-mode control, nonlinear model predictive control, LSTM-TS, and a causal Transformer using leave-one-current-profile-out testing. RWKV-TS has a root-mean-square position error of 0.23 m and a 95th-percentile error of 0.51 m with a medium current. The root-mean-square error was lowered by 18.0% when compared to the causal Transformer and by 27.1% when compared to the nonlinear model predictive control since the comparable values for the most turbulent profile were 0.41m and 0.88m. The recovery time following a 0.7 m/s lateral-current step was 6.8 s, and the average electrical energy was 11.6% less than that of sliding-mode control. Ablation experiments have demonstrated the effectiveness of current-residual conditioning, recurrent time mixing, and saturation feedback. With a measured inference delay of 2.7 ms on an embedded graphics module and a cautious fallback mode for out-of-distribution currents, the results show that RWKV-TS is a moderately effective data-driven augmentation for station-keeping.

Keywords: *Underwater Robot, Station Keeping, Ocean Current Disturbance, RWKV-TS, Navigation Time Series, Adaptive Control*

Received on 07 December2021, Accepted on 28 February 2022, Published on 07 March 2022

Introduction

Station-keeping capability is critical for hovering autonomous underwater robots performing inspection, docking, and underwater sampling tasks, where the vehicle is required to maintain a fixed horizontal position under complex marine hydrodynamic interference [1]. Continuous and time-varying ocean currents constitute the dominant disturbance source in practical operations. Even small instantaneous velocity errors induced by low-frequency steady currents can gradually accumulate into substantial positional drift, severely degrading mission reliability. Existing research has developed multiple disturbance-rejection control frameworks for underwater vehicles, including adaptive model predictive control built upon hydrodynamic characteristics, robust predictive control coping with model uncertainties and current interference, bionic-robot disturbance-suppression strategies, and disturbance-observer-enhanced sliding-mode and fuzzy adaptive controllers. Li et al. [1] built an adaptive model predictive control system for a spherical underwater robot in a current disturbance based on hydrodynamic characteristics. Station-keeping for an over-actuated hovering vehicle in the presence of current effects and model uncertainty has been introduced by Vu et al [2]. Robust predictive adjustment for external disturbances of a bionic underwater vehicle has been studied by Yang et al. [3]. Trajectory tracking for an underactuated autonomous underwater vehicle in ocean currents was studied by Wang [4]. Rong et al. [5] presented a fractional-order sliding-mode controller with a disturbance observer, and Zhou et al. [6] studied quantized-state feedback under environmental disturbance. Lakhekar et al. also introduced a fuzzy-adaptive S-surface controller and disturbance monitoring in reference [7]. Most of the above studies use calibrated dynamics, required observer structures or repeated online optimization, but taken as a whole, they have developed practical stability and disturbance-rejection mechanisms. Continuous low-frequency flow can cause position error accumulation even when the instantaneous velocity error is small, so station-keeping adds another problem. To prevent the learned compensator from overriding certified local stabilization, a realistic controller should retain the causal disturbance history, respect limited thruster authority, and account for delayed or erroneous navigation readings.

Navigation and actuator records automatically contain such history. Doppler velocity, inertial acceleration, prior thrust, accumulated position error, yaw response, and validity masks are displayed as the vehicle's motion and the actuators' reactions. The aforementioned correlations can be learned using recurrent networks, however conventional recurrent units are costly over extended periods of time; fixed-window Transformers need a big cache and may attenuate small but persistent interruption events. Directly learned rules can provide an operational risk since they may extrapolate beyond the training distribution, perform unpredictably in the absence of Doppler signals, or require impractical push. Thus, a successful station-keeping learner should use linear-time bounded memory, distinguish between slow current biases and short impulses, handle sensor validity and saturation explicitly, and behave as a residual correction function instead of a full-replacement controller. Provide a deterministic backup in cases where the command is impractical, the remaining size is too big, or the anticipated energy is uncertain. The length of the tail error, recovery time, command smoothness, energy consumption, dropout and delay robustness, safety rejection, embedded memory, latency, and long-term state stability are other performance indicators in addition to the mean position error.

RWKV-TS, a causal recurrent sequence model of station-keeping in ocean-current-disturbed navigation, is based on this work. Synchronous navigation, actuation, and environment recordings are used to generate body-frame tokens with location and velocity error, acceleration residual, previous command, yaw characteristics, and validity indicators. While RWKV-style time mixing preserves a constant-memory disturbance context, channel mixing distinguishes between slow-shift flow effects and fast-correction signals. Incorporate a bounded residual head into the conventional low-gain stabilising controller. A safety monitor will discard the learnt increment if the authority, energy, residual, or data-quality checks are unsuccessful. Mission-level splits stop relevant current episodes from leaking. Experiments are conducted to compare deterministic and learnt baselines under current speed, Doppler dropout, navigation delay, held-out disturbance profiles, energy consumption, recovery consistency, embedding latency, and safety events.

In the remainder of the paper, learned disturbance compensation and current-robust control are presented; the vehicle data, residual objective, and causal windows are defined; RWKV-TS and its saturation-aware training rule are introduced; accuracy, robustness, energy, ablation, navigation degradation, and deployment behaviour are evaluated; transfer limitations and future work on vertical currents, formal reachability, sensor ageing, and longer open-water trials are concluded. Poor navigation/performance in familiar currents and behaviour in the held-out profile will also be evaluated separately. As a result, a brief, correlated window won't seem more generalisable than it actually is. Report halting residuals, safety rejection, energy, and command changes. To find out if the statistical improvement is also practical, add the position error. A controller is deemed successful only when it can maintain the position more precisely without running the risk of actuator saturation, high power consumption, unreliable switching, or a longer stopping-distance tail.

Related Work

Current-Robust Underwater Control

Current-disturbed path-following has been addressed by the following control methods. Abdurahman et al. [8] have combined line-of-sight guidance with speed allocation, and Ding et al. [9] have used prescribed-performance sliding-mode tracking and a neural disturbance observer. Tang et al. [10] combined nonlinear predictive control with disturbance estimation for dynamic positioning, and Jiang et al. [11] developed vector-thruster control of articulated underwater robots under strong time-varying disturbances.

Station-keeping is not path-following; instead, a vehicle can have a small instantaneous velocity error but will accumulate a persistent low-frequency current of position error. Yan et al. [12] have achieved robust nonlinear predictive control for a biomimetic platform, and Nguyen et al. [13] have realized station-keeping without direct linear-velocity measurements. Based on the above results, disturbance rejection needs to consider both model uncertainty and the history of the position mismatch.

Observer-based Memory is one way; generally, its structure is selected from an assumed physical model. Liu and Du [14] used composite learning for unknown three-dimensional vehicle dynamics, and Oliveira et al. [15] combined a disturbance observer with predictive control for vehicle-manipulator systems. This study, on the other hand, learns a bounded residual from causal navigation history and keeps a deterministic nominal controller and explicit actuator limits.

Learned Disturbance Compensation

For image-based underwater manipulation, Gao and his associates have employed active disturbance rejection and predictive control [16]. An adaptive disturbance observer for the tracking path of underwater vehicles has been developed by Guerrero et al. [17]. Hou and associates investigated constrained predictive tracking for a biomimetic spherical robot [18]. Although the aforementioned experiments demonstrate that residual or observer information might enhance disturbance rejection, it is not guaranteed because a directly learnt policy may still surpass actuator limits or display unpredictable behavior outside of the training distribution. As a result, the current approach only learns a restricted residual command while maintaining deterministic allocation, monitoring, and fallback control.

Sequence learning is the other. The required control step can be derived from the recent current values using a causal model, and a typical feedback path is locally stable. Transformers' quadratic attention cost and long-range context may make them unnecessary for embedded station-keeping loops. A recurrently weighted key-value update without explicit pairwise attention is called RWKV. Because of this, there isn't a station-keeping-specific RWKV design that can employ navigation residuals, takes thruster limitations into account, and is evaluated under current profiles that aren't in the training data.

The majority of learned controllers currently in use concentrate on path-following error along a predefined path, but station-keeping has a different failure mode: the commanded velocity is almost zero, and a slowly varying current can cause continuous position drift that is hard to detect with a single observation. Corresponding evidence is disseminated through channels and over time. Prior thrust records the vehicle's prior effort in a reaction; Doppler velocity is instantaneous drift; inertial acceleration is rapid disruption; and accumulated position error is ongoing loading. Consequently, a causal sequence model will be used to distinguish between externally induced motion and the effect of actuation.

Therefore, station-keeping is called causal residual sequence control in this context. It does not require a direct current meter or a complete hydrodynamic model to discover how the navigation residual changes following a known control operation. To preserve an explicit safety barrier, a limited scope is used: RWKV-TS only provides a finite compensation increment when command feasibility and navigation quality are satisfied, and the nominal controller, thruster allocation restrictions, and mission-abort rules continue to be deterministic.

Navigation Data and Problem Formulation

Vehicle, Sensors, and Current-Disturbed Dataset

The platform is a 38 kg hovering autonomous underwater vehicle with four horizontal thrusters, a pressure sensor, an inertial navigation unit, a Doppler velocity log, and an acoustic position reference used only for evaluation and periodic correction. The horizontal controller runs at 10 Hz. The dataset contains 18 missions and 63.4 h of valid operation: 21.7 h in a recirculating tank and 41.7 h at a sheltered coastal site. Current speed ranges from 0.05 to 0.82 m/s and includes steady, sinusoidal, sheared, cross-flow, and turbulent segments. The present work uses synchronized navigation, actuator, and environmental records without requiring direct force labels. Timestamps with Doppler lock loss longer than 0.8 s, acoustic position jumps greater than 1.5 m, or thruster-telemetry faults were excluded. Short Doppler gaps were retained with validity masks rather than interpolated, thereby preserving the causal relationship between sensor quality and control response.

$$x_t = [p_t, v_t, \psi_t, c_t] \quad \text{Eq. (1)}$$

The navigation vector x_t groups horizontal position p_t, body velocity v_t, yaw ψ_t, and contextual measurements c_t at the current sampling instant. This causal representation supplies the controller with motion, orientation, and operating-context information without using future observations.

Context includes acceleration, depth, current proxy, data masks, and previous thruster commands. The compact vector keeps all channels causal and allows the same preprocessing on the embedded computer. Position error is measured from the commanded station and expressed in a current-aligned local frame so that cross-current and along-current corrections are distinguishable. This alignment reduces unnecessary relearning when the vehicle heading changes.

$$e_t = p^* - p_t \quad \text{Eq. (2)}$$

The station-keeping error e_t is the difference between the commanded position p_* and the measured position p_t. Expressing this difference in the current-aligned local frame separates along-current and cross-current correction requirements. Figure 1 connects mission-scale data collection to model-ready causal samples: synchronized navigation and actuator streams are screened for sensor faults, transformed into residual-aware body-frame features, and segmented by current profile so that held-out conditions remain temporally independent.

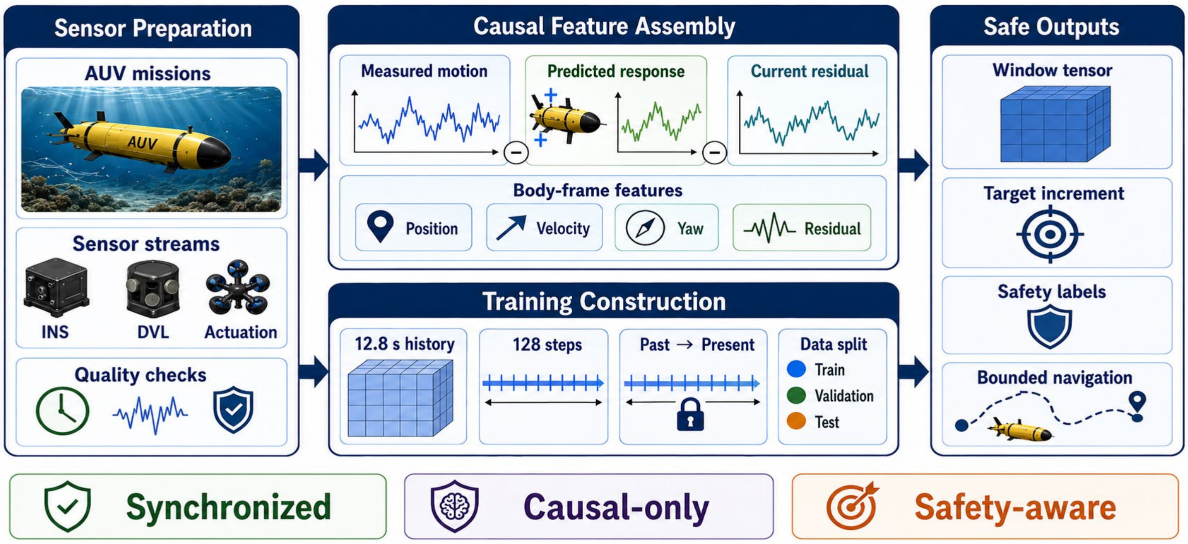

Figure 1. Overall architecture of the navigation-data pipeline and causal training-window construction.

Since sensor alignment is now a component of the control problem, it will no longer be handled independently. Get data from the Inertial Measurement Unit at 100 Hz. The output of the bottom lock runs at a thruster telemetry rate of 20 Hz and provides Doppler velocity data at 5–10 Hz. All channels have been resampled to the 10 Hz control clock with causal zero-order hold, and the age of each held measurement has been noted. The age channel won't be confused with an abrupt current drop because of delayed Doppler sampling. Calculate clock drift after obtaining hardware timestamps at the start and finish of each mission. Throw away the record if the inexplicable skew is greater than 25 ms. A training session is an incident that disrupts the entire system. Do not divide the window and the subsequent window into training and validation sets to prevent a small-leakage issue with a fluctuating current. Instead of smoothing using the inertial path, make a positioning correction in relation to the acoustic reference as an event. It is possible to differentiate between actual car motion and a navigation frame jump using a binary correction flag. Currently, only stratification evaluation for the profile label is carried out, and the controller does not acquire the profile ID during training or inference.

Disturbance Residual and Control Objective

A lower-risk residual formulation is adopted. A low-gain proportional-derivative controller produces the nominal command, while RWKV-TS predicts only the bounded increment required to compensate for persistent current effects. This arrangement preserves a deterministic stabilizing path and limits the learned model to disturbance correction.

The current residual is calculated as the mismatch between measured body acceleration and the acceleration predicted from commanded thrust by a coarse identified model. It is not treated as ground-truth current velocity. The residual is filtered at 2 Hz, normalized using training-mission statistics, and accompanied by a validity mask, providing disturbance-sensitive information without embedding a high-gain observer in the learned loop.

$$d_t = a_t - \hat{a}_t \quad \text{Eq. (3)}$$

The disturbance residual d_t compares the measured body acceleration a_t with the coarse-model prediction $\hat{a}_t$. Slowly varying residual bias indicates sustained current or model mismatch, whereas rapid peaks capture short disturbances and thruster interaction.

RWKV-TS Station-Keeping Model

Causal Tokenization and Time Mixing

Each training example contains 128 samples, corresponding to 12.8 s of history. Continuous channels are standardized, angular variables use sine-cosine encoding, and masks are concatenated to affected signals. A linear layer maps each time step to a 96-dimensional token. Six RWKV-TS blocks then apply recurrent time mixing and gated channel mixing. Unlike full attention, the recurrent state has constant size, so inference cost does not grow with mission duration. The time-mixing state accumulates keys and values with a learned exponential decay. Separate decay rates let some channels retain long current bias et al. emphasize recent impulses. Layer normalization precedes each mixing operation. Residual connections preserve the nominal navigation representation and make the model trainable without a large warm-up schedule. Token combines current navigation data and the previous command. Including prior actuation helps distinguish an external current change from vehicle motion caused by its own thrusters. Learned projections place both contributions in the same latent space.

$$r_t = W_x x_t + W_u u_{t-1} \quad \text{Eq. (4)}$$

The fused token r_t combines the current navigation representation x_t with the previously applied command u_{t-1}. The two learned projections place sensing and actuation history in the same causal latent space.

$$[k_t, q_t] = W r_t \quad \text{Eq. (5)}$$

The key vector k_t controls temporal weighting, whereas the value vector q_t carries the candidate disturbance response. These compact projections organize causal evidence without constructing a high-dimensional attention map.

$$\alpha_t = \sigma\left(k_t + \overline{k}_t\right) \quad \text{Eq. (6)}$$

The mixing coefficient α_t determines how strongly the current candidate response influences the recurrent state. Values near one emphasize new evidence, whereas smaller values preserve long-memory information.

$$h_t = \alpha_t q_t + (1 - \alpha_t) h_{t-1} \quad \text{Eq. (7)}$$

The updated state h_t combines the current candidate q_t with the retained state h_{t-1}. This channel-wise update represents both rapid current impulses and slowly varying bias without storing the full input sequence.

Figure 2 clarifies how the learned sequence pathway and conventional controller cooperate: recurrent memory extracts persistent current effects, the bounded residual head proposes incremental corrections, and safety checks can suppress them while the nominal stabilizing path continues to command the allocator.

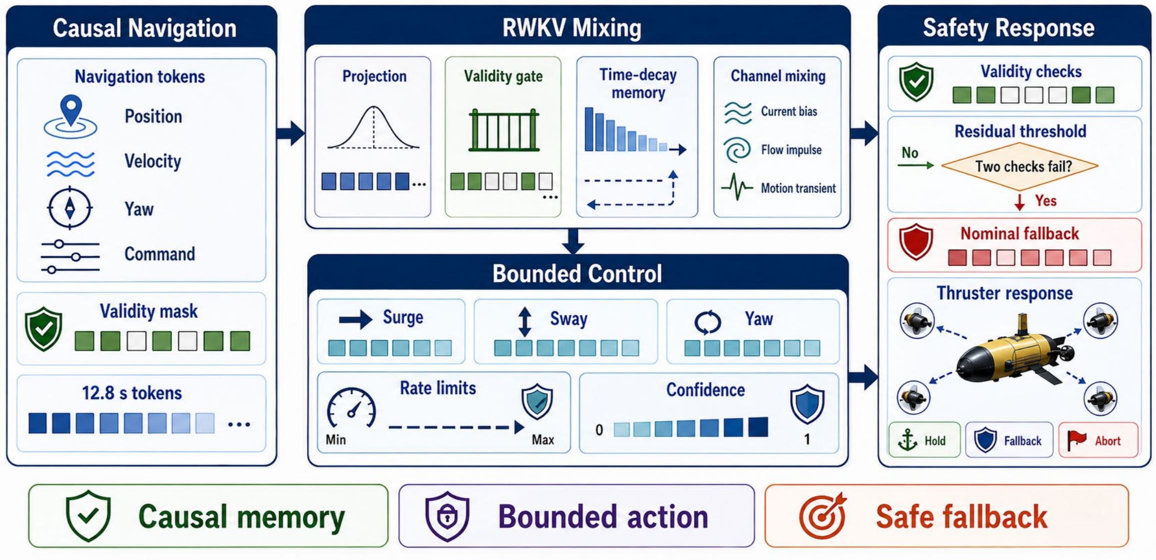

Figure 2. Overall architecture of the RWKV-TS residual control and deterministic safety framework.

The recurrent summary was reset to zero at vehicle startup and then updated continually. As a warm-up, the first two seconds of every offline mini-batch were not included in the loss. Only the forward-pass state was retained, and the gradients of the recurrent state were truncated after 256 samples in order to limit training memory during a lengthy flight. The position inaccuracy rose by 9.6% while maintaining the nominal control window's continuity and resetting the state at each window. The decay parameters were restricted by making them sigmoidal in shape. As a result, all recurrent channels are stable and no single long-term memory unit has developed an unlimited bias. While slower channels detected Doppler residuals and sustained sway demands, faster-learned channels reacted to acceleration impulses and yaw-rate variations. This division was not manually placed in frequency bands; rather, it emerged during training. Performance is largely unaffected by slight variations in window length because the pertinent history is carried in the recurrent state outside of the explicit input window.

Channel Mixing and Safe Control Head

Time mixing captures persistence, whereas channel mixing resolves interactions among pose error, velocity, current residual, and prior thrust. A gated feed-forward unit expands token width by a factor of three and returns to 96 dimensions. Dropout is disabled during control inference. The output head predicts surge, sway, and yaw command increments, not absolute motor signals. A conventional proportional-derivative command remains active in parallel. The learned increment is limited to 45% of available horizontal authority, and a control-barrier check rejects increments that would increase a one-step Lyapunov estimate beyond a calibrated margin. Rejected increments decay smoothly to zero over 0.5 s, avoiding an abrupt change in thrust.

$$g_t = \sigma\left(W_g h_t\right) \quad \text{Eq. (8)}$$

The channel gate g_t is obtained from the recurrent state h_t through a sigmoid projection. Its bounded components select latent features relevant to the current decision and suppress channels associated with degraded measurements.

$$y_t = W_o(g_t h_t) \quad \text{Eq. (9)}$$

The output vector y_t maps the gated recurrent features g_t and h_t through the output projection. Its three normalized components represent the surge, sway, and yaw command increments before authority rescaling and safety checks.

The safety path's three-phase exam. Test of navigation: Doppler lock, acoustic correction magnitude, and measurement age. Comparing the anticipated one-step error with the error anticipated from a nominal controller allows for dynamic testing. The authority test establishes if the given motor commands are short-lived and continuous. Only after passing all three exams is a learnt increment allowed. Although station keeping prioritizes predictable conduct over sporadic aggressive error reduction, this conjunction is conservative. The

https://doi.org/10.64972/dea.2022.v1i1.4127d:84-99

rejection state has a 0.3s entry and a 0.5s departure persistence to avoid chatter at the acceptance border. Instead of being instantly set to zero following rejection, the learned increment is decreased by a steady decline. Since the nominal controller is always determined using the current condition, a warm start is not necessary. Logs facilitate post-mission diagnostics without expanding the control loop's computation by storing the failed test, anticipated increment, assigned thrust, and recurrent-state norm.

Objective and Saturation-Aware Update

Training minimizes multi-step position error, velocity error, command variation, and a penalty for exceeding available control authority. Losses are evaluated over a 20-step rollout, which exposes drift that a one-step objective would miss. Teacher forcing is reduced from one to zero over the first 30 epochs so the recurrent state learns to operate on its own previous commands.

The allocator maps body commands to four thrusters with a fixed pseudoinverse and rescales all channels when any motor would saturate. The applied command, not the unconstrained prediction, is returned to the next input token. This feedback is essential because current rejection depends on the actuation actually delivered.

$$L = L_p + \lambda_v L_v + \lambda_u L_u \quad \text{Eq. (10)}$$

The total objective L combines the position term L_p, velocity term L_v, and control-variation term L_u using the corresponding weights. This balance prioritizes station accuracy while discouraging rapid command reversals and excessive actuation.

$$u_t = \mathrm{sat}(u_{t-1} + \Delta u_t) \quad \text{Eq. (11)}$$

The applied command u_t is the saturated sum of the previous command u_{t-1} and the learned increment Δu_t. Feeding the feasible applied command back to the next input prevents recurrent-state drift when a thruster reaches its authority limit.

Experimental Results

Experimental Protocol

Temporally adjacent sections could not be covered by data division since the missions were split based on the whole current-profile families rather than at random by time. The held-out test data was not utilised to create targets, define hyperparameters, or select checkpoints since one profile family was retained for testing in each fold, one mission from the remaining families was used to aid in validation, and all other missions were used for training. With AdamW, a weight decay of 0.01, gradient clipping set to 0.8, and 96 windows per batch, cosine learning rate decay was applied across 140 epochs. Incorporate navigation bias, actuation delay, Doppler dropout, and current-residual scaling into the data. There isn't a single trustworthy sensor anymore because of all the major navigation issues mentioned above. Recovery Time, Mean Electrical Energy, Maximum Error, 95th Percentile Error, Horizontal Position RMSE, and Command are the six assessment indices of Total Variation. Choose one that satisfies the safety requirements and has a comparatively low weighted validation score. Consequently, just a somewhat significant tail error will marginally raise the mean position error; neither the energy consumption nor the actuator operation will increase.

Establish two limited phases for the control objectives. The measured next-step velocity residual gave a bounded adjustment under the same authority limit as RWKV-TS after a nonlinear predictor produced commands that followed the calibrated vehicle model. The technique avoided risky policy-gradient exploration in hardware without treating the predictor as a flawless expert because both procedures were performed offline and omitted all held-out test segments. Each held-out profile was optimized using five randomly chosen seeds. The selected checkpoint should have a comparatively low weighted score and satisfy the validation safety standards. There was no improvement following eighteen epochs of early stopping. The recurrent accumulators were still full-precision, but mixed-precision training decreased memory utilization. At each epoch, note the saturation frequency, gradient statistics, and validation rollout duration. Additionally, checkpoints with recurrent-state norms were eliminated even if their short-horizon errors were greater than the upper bound.

All controller settings, safety thresholds, baseline definitions, and deployment conditions were fixed prior to the evaluation of the held-out current profiles. LSTM-TS, nonlinear model predictive control, disturbance-observer sliding-mode control, proportional-integral-derivative control, and the six-layer causal Transformer all employed the same sensor filtering, yaw representation, data partitioning, control frequency, thruster allocation, and authority constraints. Only hyperparameters from the validation runs were utilized to provide an upper bound for comparison; preprocessing and actuator constraints were not changed. Initially, a number of arrangements are screened for disturbance rejection. Ultimately, the most disturbed-observer sliding-mode type was selected among the aforementioned. Any variations in performance may be traced to the controller rather than inconsistent actuation limits or preprocessing because all methods used the same sensor filters, thruster allocators, control frequencies, and authority limitations. RWKV-TS is a conservative switch that takes command viability, navigation quality, predicted energy gain, and current-residual size into account. If none of the acceptance conditions are satisfied, the nominal controller is still employed while the learning increment is progressively reduced. RWKV-TS only regulates surge, sway, and yaw; as depth is within the vehicle's permitted depth controller range, improvements in horizontal station-keeping won't have an impact on the independently stabilized vertical loop.

The finished RWKV-TS network comprises 1.84 million parameters and is roughly 7.6MB in half-precision. One training fold on a single 24GB graphics card took 2.3 hours. Preprocessing, recurrent inference, and safety checks took 2.7 ms per update on an NVIDIA Jetson-class module; CPU-only inference took 18.4 ms; both were well inside the 100 ms control interval and had enough computing headroom for communication, navigation, and thruster allocation.

Over the held-out current profiles, RWKV-TS exhibited a horizontal-position RMSE of 0.31 m, a 95th-percentile error of 0.69 m, and a maximum error of 1.18 m. The causal Transformer was 0.37, 0.83, and 1.46 m for the same indicators, while LSTM-TS reached 0.40, 0.91, and 1.57 m. The separation remained constant for shearing, oscillatory, turbulent, and stable flows; the Transformer's RMSE values were 0.22, 0.28, 0.42, and 0.50m, while the RWKV-TSs were 0.18, 0.23, 0.34, and 0.41m. The Transformer (8.1 s) was faster than sliding-mode control (7.5 s), nonlinear predictive control (9.4 s), and PID (15.2 s). The proposed controller re-entered the 0.25 m/s holding radius in 6.8 s when the current increased from 0.22 to 0.70 m/s. As a result, Figure 3(a) has a shorter, less oscillatory return route and a smaller final error; Figure 3(b) has a lower median and a narrower upper tail; and Figure 3(c) indicates that the transient has entered the steady-error region after roughly 6–8 seconds. Due to a decrease in both path length reduction and dispersion, the disturbance memory has increased overall accuracy and predictability close to the inspection target or docking structure. Dai et al. [19] have shown that the predictive compensation is limited; Hasan and Abbas [20] have reported that adaptive disturbance rejection can reduce the persistent underwater tracking error, and in this case, comparison is used to observe the trend rather than to verify the measured values. As a near-constant residual can correct for steady flow, the error ordering is physically reasonable. Due to oscillations and turbulence, it is necessary to determine whether a controller signal is a short-term impulse or a slow-changing bias. Therefore, carrying the history of cause disturbances is more advantageous than using only the most recent sample, as shown by the expanded gap under turbulent flow. A shorter return path has a smaller length of the large position error and avoids frequent reversals that could excite the vehicle body. The Trajectory Panel can provide information not contained in a scalar RMSE. To limit the upper tail of the controller near the infrastructure, a distribution box is also needed; otherwise, only the mean would be increased. The recovery panel shows that the car switches to the normal stabilisation path when it is near the station and has added the learned residual during the transient period. Thus, after the disturbance has passed, this high corrective command will be less likely to continue indefinitely. From an engineering standpoint, a wide loop or an extended error tail may be unacceptable even if the average accuracy is relatively high; instead, a combined reduction in path length, error dispersion and settling time that is suitable for inspection and docking can be achieved. This result is further supported by the held-out-profile design: the observed order cannot be explained by nearby, highly correlated windows in both training and testing because whole current families were excluded from training.

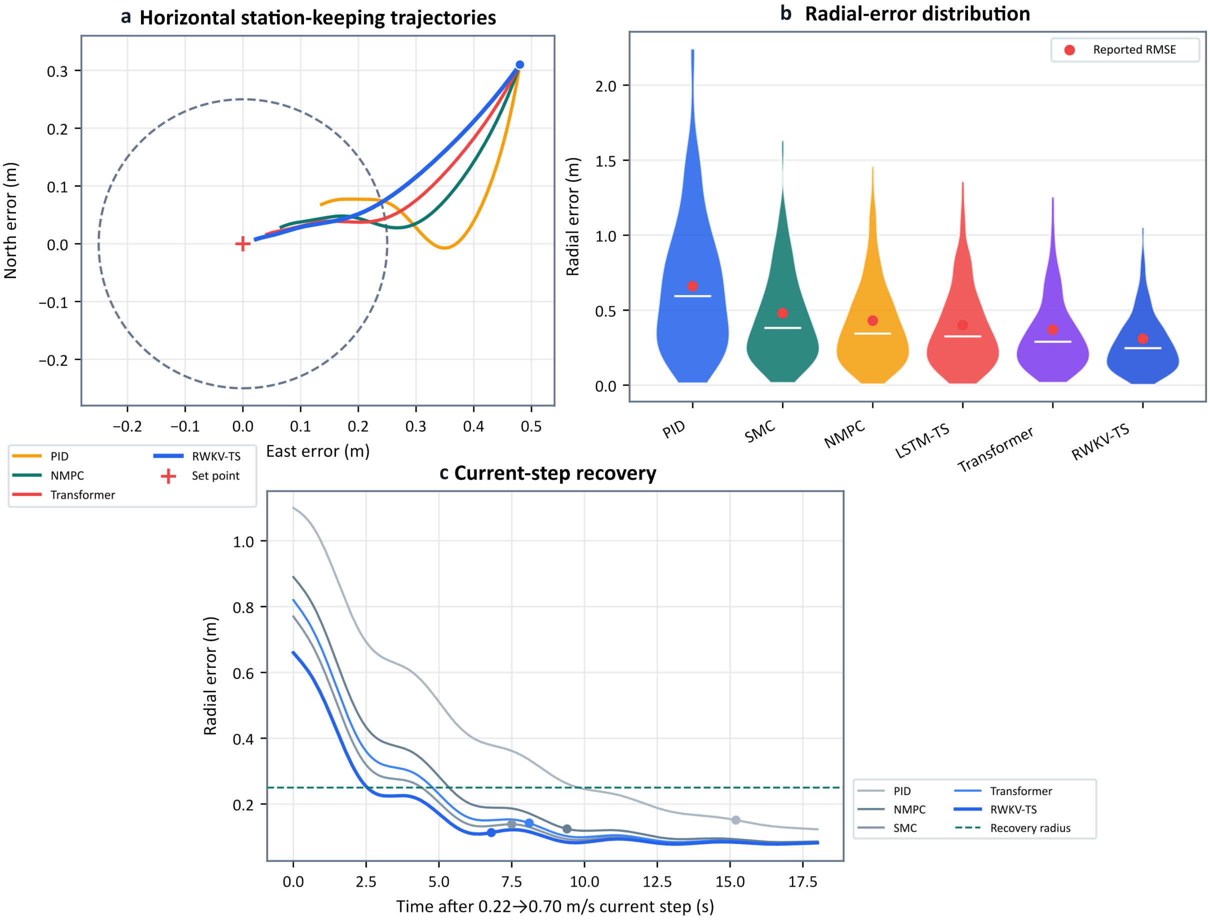

Figure 3. Station-keeping accuracy: (a) horizontal trajectories; (b) radial-error distribution; (c) current-step recovery.

Accuracy, Robustness, and Energy

Because the recommended controller uses the highest horizontal-duty bands less frequently, Figure 4(b) does not rule out continuous high-duty actuation as the exclusive cause of this benefit. All controllers have an increasing position error for the current-speed sweep, as seen in Figure 4(a), but RWKV-TS has the lowest error at 0.8 m/s and the shallowest rise. Because it is the left most of them and has a shorter upper tail, as seen in Figure 4(c), it recovers from episodes more rapidly and steadily. The orientation test is interpreted similarly; for heads at a 90-degree angle, the RMSE ranges from 0.29 to 0.33 meters. The RMSE is momentarily increased to 0.46 m by a 4°/s rotating current before falling below 0.30 m after 9.1 seconds of the rotation ceasing. The energy data demonstrate that RWKV-TS's reasonably selective compensation mechanism, which uses 46.8 Wh per hour of episode—11.6% less than sliding-mode control—limits reversals, power demand, and thruster stress. Cut the overall command variation by 34.7%. Huang et al. [21] have also linked learned disturbance adaptation with resilience to unknown disturbances; Song et al. [22] have connected parameter adaptation to improved active-disturbance rejection; and Nerkar et al. [23] have highlighted the necessity of observer robustness, and the lower duty burden in the present experiments also suggests that the gain was obtained via a smoother bounded correction. The relatively shallow error slope in the current-speed panel shows that the bounded residual uses the available thrust more sparingly at higher speeds. If the improved accuracy had been achieved solely by a higher feedback gain, the duty-cycle panel would be expected to show greater occupancy in the highest-duty bands; instead, the shift away from these bands indicates that the disturbance history reduces redundant corrective action. The recovery-time panel is a typical case of the above distinction. A narrow recovery-time distribution means that operators know roughly how long the vehicle will be out of service after a current event, but a long tail is unpredictable, and an inspection may need to be aborted. The small spread of the four headings also shows that current alignment and body-frame actuation are relatively insensitive to orientation. The temporary degradation during continuous current rotation indicates a remaining problem: the disturbance frame changes more rapidly than the slow recurrent channels can fully adapt, but the subsequent return below 0.30 m shows that the state does not have a harmful obsolete bias. Energy and command-variation results are relatively stable mechanisms. A smaller total variation indicates fewer direction changes and a lower-amplitude oscillating thrust; thus, the energy loss is relatively small. The combination is especially suitable for battery-limited vehicles; thus, the improved efficiency will extend the service life and reduce thermal and mechanical stress on the thrusters.

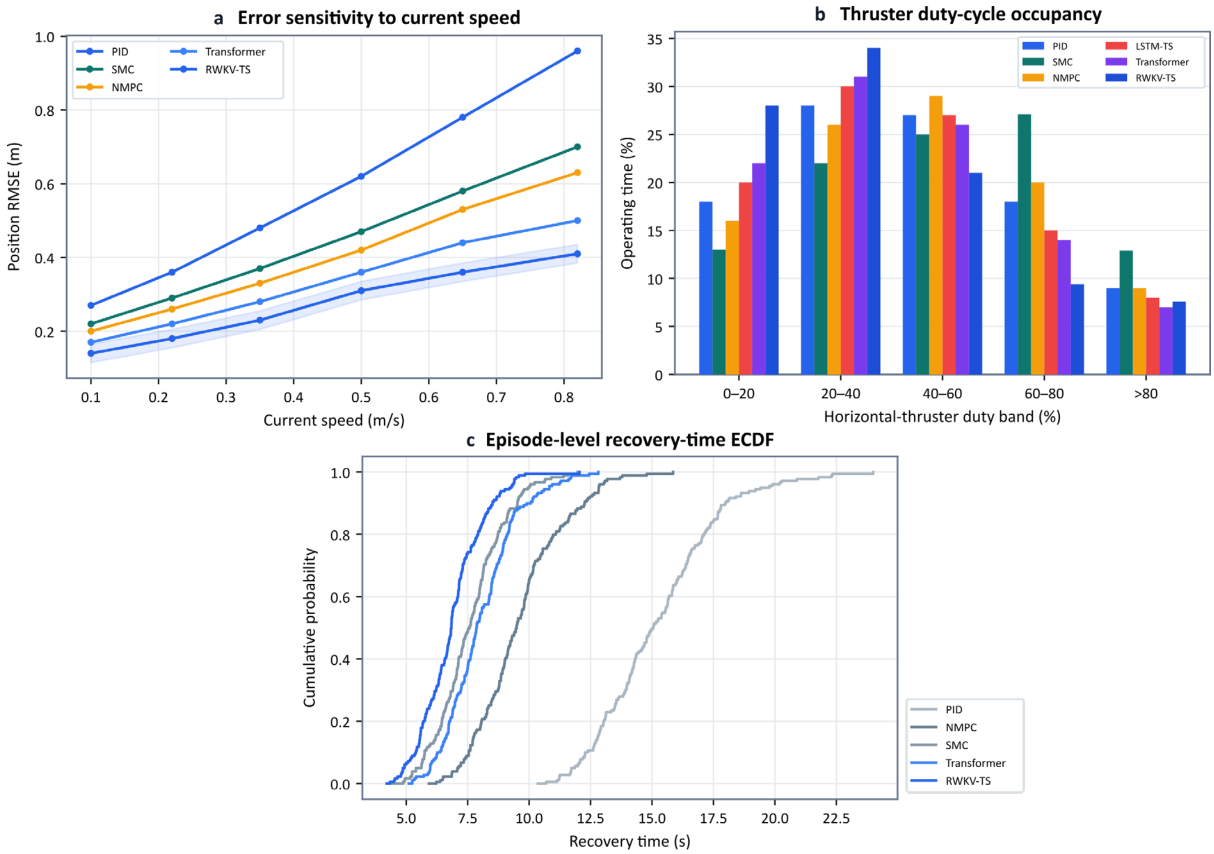

Figure 4. Current robustness: (a) error versus current speed; (b) thruster duty cycle; (c) recovery-time distribution.

Saturation feedback, recurrent memory, disturbance representation, and deterministic monitoring are all independently successful, according to ablation tests. The RMSE of the complete model, as seen in Figure 5(a), is 0.31m; without the current-residual pathway, it is 0.39m; when recurrent time mixing is replaced by a feed-forward window, it is 0.43m; and without saturation feedback, it is 0.48m. As seen in Figure 5(b), the episode energy is 46.8 Wh, which is nearly 52 Wh. Wh for the entire model, and this is due to fewer repetitive surge and sway adjustments rather than an increase in control effort. As Figure 5(c) illustrates, nominal mistakes are still somewhat uncommon; however, authority-limit events rise with the delay, and even if the nominal error has barely changed, there are nine violations when the safety monitor is withdrawn. The aforementioned modules demonstrate how independent monitors reduce off-nominal behavior, preserve saturation data to maintain state-command consistency, and enhance performance using recurrent memory. Liang et al. [24] have shown that learned disturbance information can strengthen underwater sliding-mode control; Zhu et al. [25] have proposed representative choices to improve tracking robustness; and Wan et al. [26] have demonstrated the advantages of an adaptive disturbance-rejection structure. Currently, the aforementioned phenomena operate differently: recurrent memory improves efficiency, saturation feedback preserves practical actuation, and deterministic rejection lowers the likelihood of faulty observations. The aforementioned explanations are likewise in turmoil. The network loses a compact indicator of the unmodeled acceleration after eliminating the current residual, so the recurrent state can only identify the disturbance using location, velocity, and command history. To maintain measurements and eliminate the persistent state needed to connect a slow bias with earlier actuation, a feed-forward window can be utilised in place of time mixing. The vehicle's real command should be returned to the sequence model if there is a significant penalty without saturation feedback; otherwise, the hidden state would change as if an unreasonable thrust had been applied. As a result, both an isolated efficiency

https://doi.org/10.64972/dea.2022.v1i1.4127d:84-99

outcome and increased state-command consistency contribute to the energy attribution displayed in the energy attribution panel. The safety-event panel displays an alternative use for the safety monitor. The majority of the learnt increments are already achievable; eliminating them won't alter the average time, and a delayed observation causes the perceived and actual vehicle conditions to differ. If the rise resulting from that mismatch is not accepted, the taught path cannot improve a previous repair. Nominal accuracy and off-nominal protection should be distinguished in certification; if it limits behavior under sensor deterioration, a never-triggered component would be needed. The findings thus support the multi-level interpretation: recurrent memory carries out performance recall, saturation feedback guarantees physical consistency, and deterministic monitoring provides independent deployment protection.

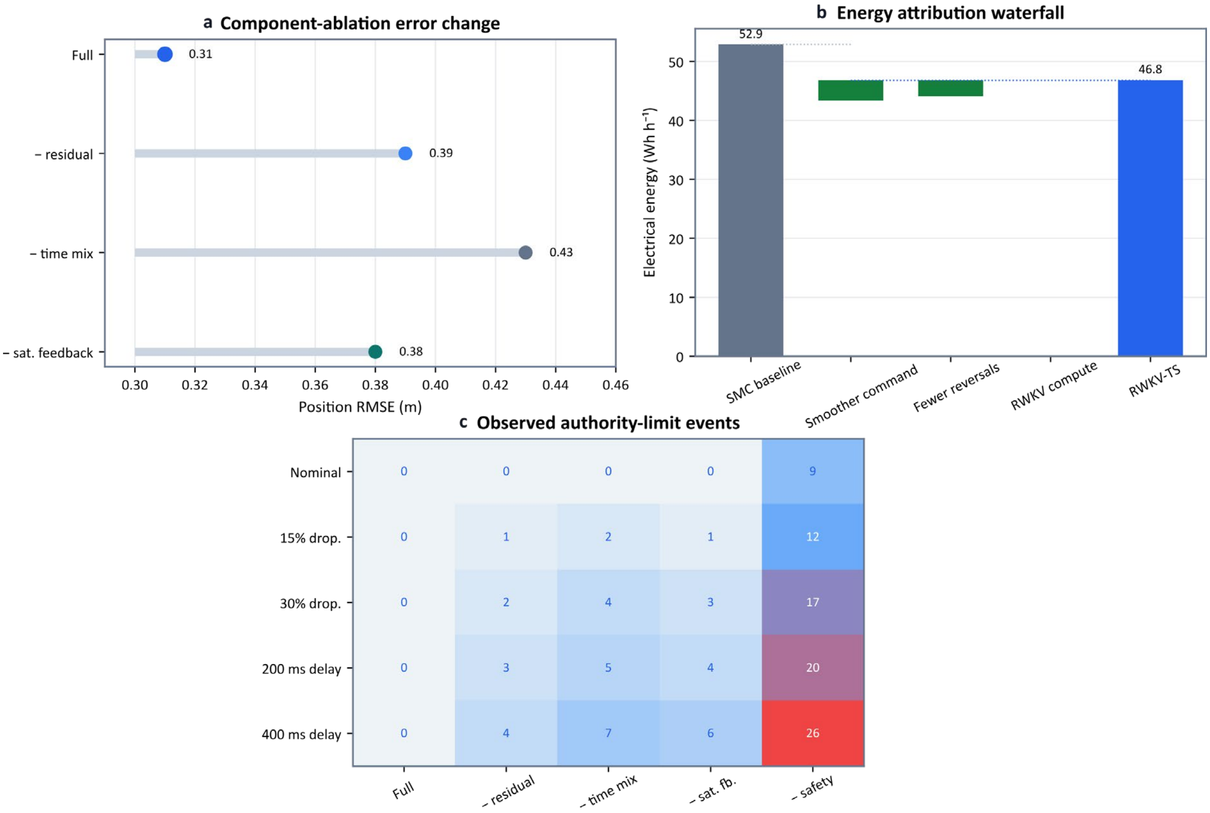

Figure 5. Ablation and energy attribution: (a) position-error change; (b) energy change; (c) safety-limit events.

Degradation of navigation is a gradual, rather than an abrupt, decline in performance. RWKV-TS performs better than Transformer (0.61m) and LSTM-TS (0.68m) for all Doppler dropout levels from 0% to 30%, as seen in Figure 6(a), with the lowest error being 0.49m at 30%. Figure 6(b) illustrates another kind of risk; a respectable measurement may still be made even after a 400 ms delay, with an error of about 0.62 m. Figure 6(c) illustrates the intended safety reaction; the rejected learning increments are 12.6% prior to the safety radius being exceeded. Based on a width-depth comparison, the 96-dimensional six-block model is within the practical operating range; a smaller four-block design has a latency of 1.6ms but an error of 0.36m; an 8-block model with a larger size would need roughly four times the latency and memory for a 0.01m improvement. The chosen threshold produces an RMSE of 0.31 m and does not show the observed violation; eight authority-margin exceedances occur while accepting all learning increments lowers the nominal error to 0.30 m. Consequently, the actuation risk resulting from delayed or absent navigation is somewhat mitigated by the minor accuracy cost. Szymak and Kot [27] have pointed out that stable trajectory control in the operating environment always requires continuous actuation, and Li and Zhang [28] have introduced a disturbance-rejection device for difficult underwater operations. The new navigation-degradation results expand this reasoning by demonstrating that selective fallback trades a slight loss in accuracy for constrained authority under absent or delayed measurements. Dropout and delay studies are two instances of information loss. By utilizing the prior motion and actuation data, the recurring state can cover a brief duration, and random dropout prevents erratic velocity readings. Because a legitimate measurement might be made after the car has already left, it is riskier. Although

https://doi.org/10.64972/dea.2022.v1i1.4127d:84-99

the observations are retained, the fixed delay causes them to become outdated. Consequently, the safety-coverage panel's increased rejection rate is probably the intended result rather than a sign that the learned controller has failed. It is evident that the monitor transfers control to the nominal controller when the measurement relevance declines; a steady rise in inaccuracy suggests that this transfer does not produce a sudden decline. A Width-Depth Study offers an additional deployment trade-off. Even tho a tiny network is quick, it could not be accurate enough to reach the intended operating point, and a huge network needs a large increase in resources for a slight improvement. For allocation, communication, and navigation, the chosen six-block configuration in the 100 ms control cycle is still sufficient. The same is true for threshold selection: all increases will result in a tiny nominal error but allow authority overruns; a high threshold will prevent necessary adjustments and increase the error. A certain level of safety can be purchased with a comparatively modest nominal penalty, and the selected value falls between the two above. According to the aforementioned research findings, position-error optimization alone should not be employed to construct a monitor in real-world operation; instead, explicit authority and navigation-quality standards should be applied.

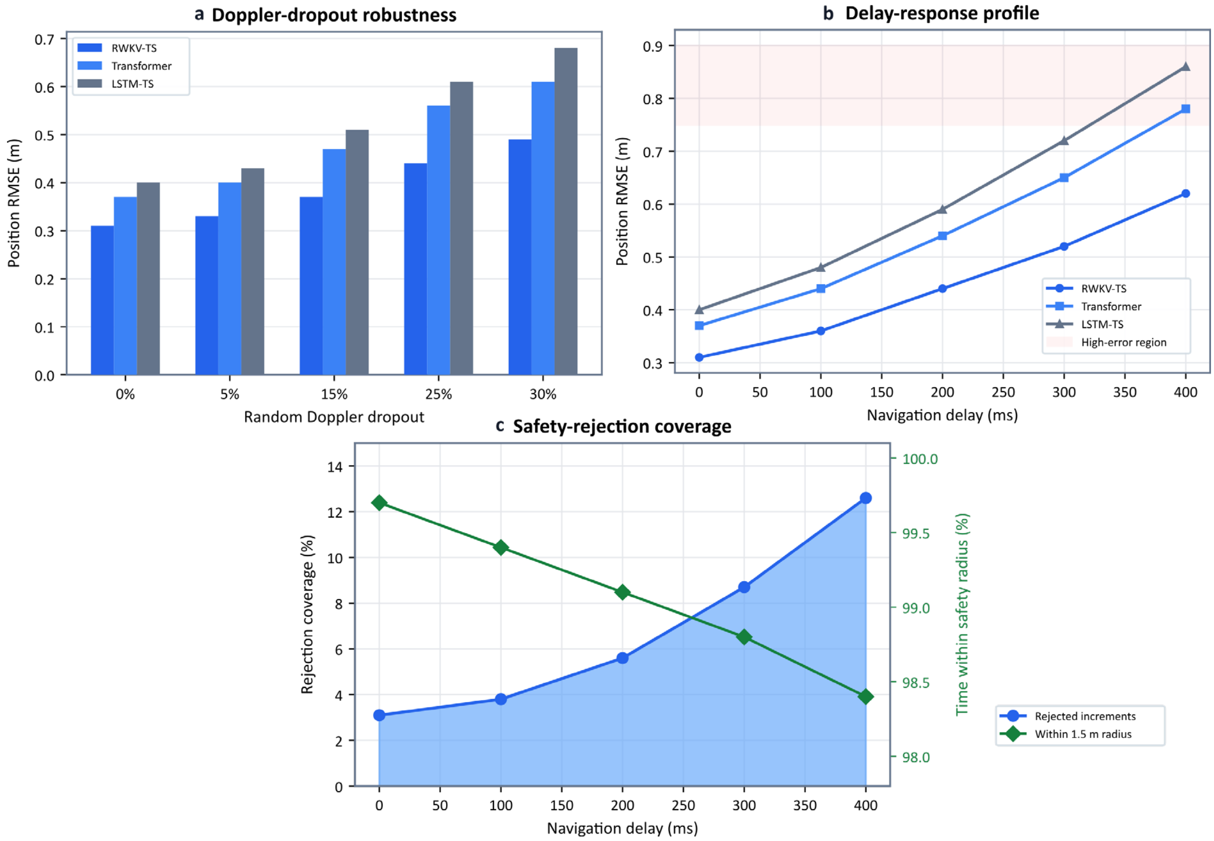

Figure 6. Navigation robustness: (a) Doppler dropout; (b) navigation delay; (c) safety-rejection coverage.

Embedded Deployment and Failure Analysis

More over 96 ms of every 100 ms control cycle remained for state estimation, communication, and allocation because embedded execution was less than 3.4 ms at the 99th percentile and 122 MB were utilised for runtime buffers. In contrast to RWKV-TS, which had a fixed-size recurrent state, the Causal Transformer required 8.9 ms and 311 MB because of the expanding rolling cache with retained context. It was not necessary to periodically reset the collected disturbance data in order to achieve minimal latency and memory use because the above state remained numerically stable throughout the whole 4.2-hour operation. Since the controller is only one component of the on-board loop, the latency margin is quite tiny. Even if its own prediction is correct, a method that consumes the majority of the cycle may cause a delay in state estimation or allocation. Since RWKV-TS is a fixed-size state, neither memory nor computation will grow as the mission goes on. Another issue is the 4.2 h state-stability observation: even if a brief offline window seems stable, recurrent models may accrue numerical drift during extended operation. Continuous-state deployment without frequent hard resets was possible because there was no such increase during the stated operation. Compared to the cached Transformer, constant-memory recurrence offers a reduced latency and a smaller runtime footprint, freeing up more embedded memory for mission supervision and observation.

3.1% of learnt increments were rejected during routine coastal deployment because of Doppler reacquisition, sudden acoustic corrections, or currents that exceeded the training limit. The deterministic nominal controller had an RMSE of 0.73m during this period; the decay-and-reentry logic did not execute a discontinuous switch, and learnt augmentation resumed within 1.2s when the navigation quality improved. Of the 38 events with radial error more than 1.0 m, 14 followed acoustic position jumps, 11 combined growing cross-current with yaw saturation, 8 happened during Doppler-quality fluctuation close to the seabed, and 5 followed reversals that were faster than those in the training set. Only three incidents went over 1.5 meters, and they all recovered on their own without operator assistance. Event distribution distinguishes sensor discontinuities from truly out-of-distribution flow and demonstrates why the larger fallback error is a reasonable trade-off: while the learnt residual is unclear, the nominal path loses some accuracy but maintains predictable, authority-limited behaviour. Consequently, the biggest mistakes did not exhibit persistent recurrent-state instability and were mostly associated with distinguishable sensor or disturbance regimes. The first is that training was limited to a certain vehicle and sensor combination. The relationship between the prior command and actual acceleration is altered by a changing thruster configuration. Consequently, a short local dataset and a conservative learned-command limit should be the first transfer. Leak detection, depth protection, collision avoidance, and mission-level abort logic cannot be implemented in this manner. Additionally, the reasons for rejection can be interpreted operationally. Out-of-distribution residuals are produced by currents beyond the training range, and Doppler reacquisition and acoustic adjustments may cause discontinuities that resemble actual acceleration. It is always safer to stop the learnt rise than to continue from an unstable level. Despite having a higher error during rejection, the nominal controller's response is still predictable and falls within the certified range. In order to prevent a second transient during the measurement quality recovery, smoothly resume within 1.2 seconds. Acoustic leaps and Doppler fluctuations are primarily sensor-induced mistakes; combined yaw saturation, cross-current increase, and quick reversal are typical of the latter. The event decomposition further separates sensor-induced errors from disturbance-induced errors. When the learnt contribution is not accessible, recovery without operator involvement demonstrates that the backup method is still operational.

RWKV-TS has the lowest energy consumption and memory usage among the learnt alternatives, and Figure 7(a) demonstrates that it has a nearly 99% station-keeping success rate. The entire model obtains roughly 12 more successful holds, the small-memory variant approaches 6, and the Transformer is even lower during challenging disturbance periods, as Figure 7(b) illustrates. Stopping-distance residuals are near zero in Figure 7(c), and the majority of observations fall within the designated range of -0.5 m to 0.5 m, with only a tiny tail outside of this range. When taken as a whole, the panels demonstrate that more success is not attained by increasing energy, including embedded memory, or methodically stopping overshoot; rather, the remaining tail highlights sudden navigation corrections and quickly rotating currents as the emphasis for additional verification. The aforementioned metrics are aggregated into the three deployment panels. The trade-off panel states that energy and memory are necessary for success, thus a controller that uses excessive power or embedded resources to maintain position might not be appropriate for a long-term mission. In line with the held-out-profile evaluation, the cumulative-gain panel demonstrates that the advantage builds up across challenging episodes rather than happening in a single advantageous trial. The residual-distribution panel then assesses if a troublesome stopping-distance tail is produced by the same controller. Allowing systematic overshoot will not improve the success, as indicated by residuals near 0 and the majority of observations in the indicated range. The limited remaining tail suggests that only a small number of episodes—mostly quickly rotating currents and abrupt navigation corrections—need to be addressed in future validation. When combined, the panels can be deployed as a bounded augmentation instead of an unrestricted replacement controller. When the input is reliable, the learnt pathway enhances disturbance compensation; otherwise, the allocator, nominal controller, and monitor maintain control based on the uncertain inferred residual. Relatively little change of the mission-level protection is required because the division of duties is in charge of adapting the procedure to various vehicles.

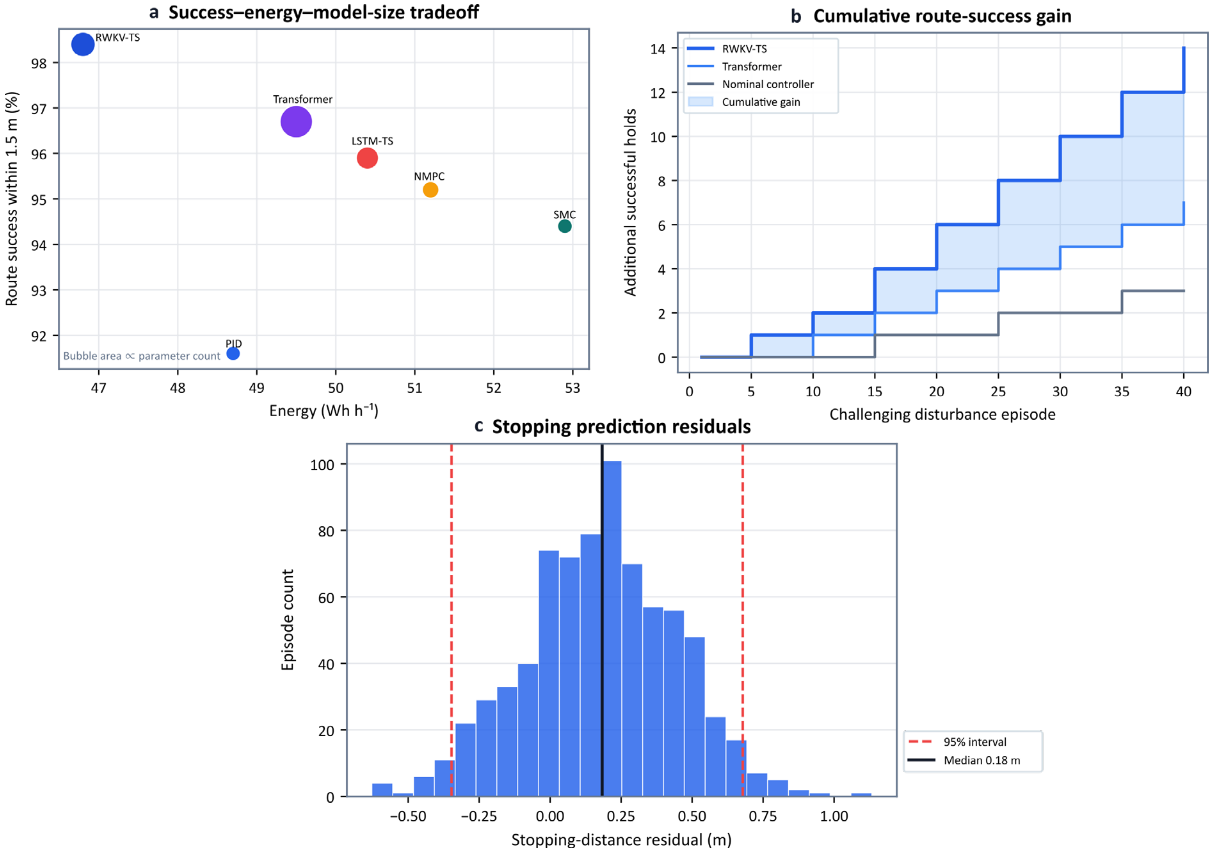

Figure 7. Deployment and safety attribution: (a) success–energy–size tradeoff; (b) cumulative route-success gain; (c) stopping-distance residual distribution.

Conclusion

Based on navigation records impacted by ocean currents, RWKV-TS is suggested as a causal recurrent sequence controller for underwater robot station-keeping. Navigation error, Doppler velocity, acceleration residual, prior actuation, and validity masks are all included in the model's constant-memory state. The traditional stabilising controller is supplemented with a bounded residual head, and deterministic monitoring and saturation feedback are employed for a workable backup mode. According to the test results, this partition improves embedded efficiency, reduces energy consumption, increases disturbance adjustment and recovery consistency, and does not completely transfer control to the learnt portion.

There are still a few restrictions. It cannot be immediately transferred to different platforms because the learnt relationship between the preceding thrust and the observed acceleration depends on the vehicle, sensors, thrusters, etc. Vertical flow, severe biofouling, seasonal hydrodynamic fluctuations, extended sensor ageing, and all quickly rotating disturbances are not properly represented by the training data, which only includes a small collection of horizontal current profiles. There is currently no explicit reachability guarantee for all operational settings, and the safety proof is empirical and dependent on calibrated authority and navigation-quality standards.

Thus, the next effort will investigate adaptive residual-authority limitations, sensor-uncertainty-aware recurrent updates, vertical-current estimation, and conservative cross-vehicle adaptation using limited local datasets. The learned increment and deterministic fallback logic will be combined with formal reachability analysis, and long-term open-water experiments will assess state stability in the face of seasonal fluctuations, biofouling, and slow sensor deterioration. The purpose of these expansions is to determine whether the bounded residual architecture can be used safely outside of the vehicle and current operating conditions under consideration.

Author Contributions

Ghassan Al-Marri contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision. Hadi Al-Suhaimi contributes to conceptualization, validation. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

[1] Li, A., Guo, S., Liu, M., & Yin, H. (2022). Hydrodynamic characteristic-based adaptive model predictive control for the spherical underwater robot under ocean current disturbance. *Machines*, 10(9), 798. https://doi.org/10.3390/machines10090798

[2] Vu, M. T., Le Thanh, H. N. N., Huynh, T. T., Thang, Q., Duc, T., Hoang, Q. D., & Le, T. H. (2021). Station-keeping control of a hovering over-actuated autonomous underwater vehicle under ocean current effects and model uncertainties in horizontal plane. IEEE Access, 9, 6855-6867. https://doi.org/10.1109/access.2020.3048706

[3] Yang, H., Yan, Z., Zhang, W., Gong, Q., Zhang, Y., & Zhao, L. (2022). Trajectory tracking with external disturbance of bionic underwater robot based on CPG and robust model predictive control. Ocean Engineering, 263, 112215. https://doi.org/10.1016/j.oceaneng.2022.112215

[4] Wang, L. (2019). Trajectory tracking control of underactuated autonomous underwater vehicle in the presence of ocean currents. International Journal of Performability Engineering, 15(6), 1716-1723. https://doi.org/10.23940/ijpe.19.06.p23.17161723

[5] Rong, S., Wang, H., Li, H., Sun, W., Gu, Q., & Lei, J. (2022). Performance-guaranteed fractional-order sliding mode control for underactuated autonomous underwater vehicle trajectory tracking with a disturbance observer. Ocean Engineering, 263, 112330. https://doi.org/10.1016/j.oceaneng.2022.112330

[6] Zhou, B., Su, Y., Huang, B., Wang, W., & Zhang, E. (2022). Trajectory tracking control for autonomous underwater vehicles under quantized state feedback and ocean disturbances. Ocean Engineering, 256, 111500. https://doi.org/10.1016/j.oceaneng.2022.111500

[7] Lakhekar, G. V., Waghmare, L. M., & Roy, R. G. (2019). Disturbance observer-based fuzzy adapted s-surface controller for spatial trajectory tracking of autonomous underwater vehicle. IEEE Transactions on Intelligent Vehicles, 4(4), 622-636. https://doi.org/10.1109/tiv.2019.2938082

[8] Abdurahman, B., Savvaris, A., & Tsourdos, A. (2019). Switching LOS guidance with speed allocation and vertical course control for path-following of unmanned underwater vehicles under ocean current disturbances. Ocean Engineering, 182, 412-426. https://doi.org/10.1016/j.oceaneng.2019.04.021

[9] Ding, Z., Wang, H., Sun, Y., & Qin, H. (2022). Adaptive prescribed performance second-order sliding mode tracking control of autonomous underwater vehicle using neural network-based disturbance observer. Ocean Engineering, 260, 111939. https://doi.org/10.1016/j.oceaneng.2022.111939

[10] Tang, L., Wang, L., Wang, Y., & Zhang, Y. (2022). An enhanced trajectory tracking control of the dynamic positioning ship based on nonlinear model predictive control and disturbance observer. Ocean Engineering, 265, 112482. https://doi.org/10.1016/j.oceaneng.2022.112482

[11] Jiang, X., Yang, F., & Shi, S. (2022). Design and full-link trajectory tracking control of underwater snake robot with vector thrusters under strong time-varying disturbances. Ocean Engineering, 266, 113012. https://doi.org/10.1016/j.oceaneng.2022.113012

[12] Yan, Z., Yang, H., Zhang, W., Gong, Q., Zhang, Y., & Zhao, L. (2022). Robust nonlinear model predictive control of a bionic underwater robot with external disturbances. Ocean Engineering, 253, 111310. https://doi.org/10.1016/j.oceaneng.2022.111310

[13] Nguyen, L. H., Hua, M. D., Allibert, G., & Hamel, T. (2021). A homography-based dynamic control approach applied to station keeping of autonomous underwater vehicles without linear velocity measurements. IEEE Transactions on Control Systems Technology, 29(5), 2065-2078. https://doi.org/10.1109/tcst.2020.3029564

[14] Liu, J., & Du, J. (2021). Composite learning tracking control for underactuated autonomous underwater vehicle with unknown dynamics and disturbances in three-dimension space. Applied Ocean Research, 112, 102686. https://doi.org/10.1016/j.apor.2021.102686

[15] Oliveira, É. L., Orsino, R. M. M., & Donha, D. C. (2021). Disturbance-observer-based model predictive control of underwater vehicle manipulator systems. IFAC-PapersOnLine, 54(16), 348-355. https://doi.org/10.1016/j.ifacol.2021.10.115

[16] Gao, J., Liang, X., Chen, Y., Zhang, L., & Jia, S. (2021). Hierarchical image-based visual serving of underwater vehicle manipulator systems based on model predictive control and active disturbance rejection control. Ocean Engineering, 229, 108814. https://doi.org/10.1016/j.oceaneng.2021.108814

[17] Guerrero, J., Torres, J., Creuze, V., & Chemori, A. (2020). Adaptive disturbance observer for trajectory tracking control of underwater vehicles. Ocean Engineering, 200, 107080. https://doi.org/10.1016/j.oceaneng.2020.107080

[18] Hou, X., Guo, S., Shi, L., Xing, H., Yin, H., Li, Z., Zhou, M., & Xia, D. (2020). Improved model predictive-based underwater trajectory tracking control for the biomimetic spherical robot under constraints. Applied Sciences, 10(22), 8106. https://doi.org/10.3390/app10228106

[19] Dai, Y., Gao, H., Yu, S., & Ju, Z. (2022). A fast tube model predictive control scheme based on sliding mode control for underwater vehicle-manipulator system. Ocean Engineering, 254, 111259. https://doi.org/10.1016/j.oceaneng.2022.111259

[20] Hasan, M. W., & Abbas, N. H. (2022). Disturbance rejection for underwater robotic vehicle based on adaptive fuzzy with nonlinear PID controller. ISA Transactions, 130, 360-376. https://doi.org/10.1016/j.isatra.2022.03.020

[21] Huang, F., Xu, J., Yin, L., Wu, D., Cui, Y., Yan, Z., & Chen, T. (2022). A general motion control architecture for an autonomous underwater vehicle with actuator faults and unknown disturbances through deep reinforcement learning. Ocean Engineering, 263, 112424. https://doi.org/10.1016/j.oceaneng.2022.112424

[22] Song, W., Chen, Z., Sun, M., & Sun, Q. (2022). Reinforcement learning based parameter optimization of active disturbance rejection control for autonomous underwater vehicle. Journal of Systems Engineering and Electronics, 33(1), 170-179. https://doi.org/10.23919/jsee.2022.000017

[23] Nerkar, S. S., Londhe, P. S., & Patre, B. M. (2022). Design of super twisting disturbance observer-based control for autonomous underwater vehicle. International Journal of Dynamics and Control, 10(1), 306-322. https://doi.org/10.1007/s40435-021-00797-1

[24] Liang, J., Huang, W., Zhou, F., Liang, J., Lin, G., Xiao, E., Li, H., & Zhang, X. (2022). Double-loop PID-type neural network sliding mode control of an uncertain autonomous underwater vehicle model based on a nonlinear high-order observer with unknown disturbance. Mathematics, 10(18), 3332. https://doi.org/10.3390/math10183332

[25] Zhu, C., Jun, L., Huang, B., Su, Y., & Zheng, Y. (2022). Trajectory tracking control for autonomous underwater vehicle based on rotation matrix attitude representation. Ocean Engineering, 252, 111206. https://doi.org/10.1016/j.oceaneng.2022.111206

[26] Wan, J., Liu, H., Yuan, J., Shen, Y., Zhang, H., Wang, H., & Zheng, Y. (2021). Motion control of autonomous underwater vehicle based on fractional calculus active disturbance rejection. Journal of Marine Science and Engineering, 9(11), 1306. https://doi.org/10.3390/jmse9111306

[27] Szymak, P., & Kot, R. (2022). Trajectory tracking control of autonomous underwater vehicle called PAST. Pomiary Automatyka Robotyka, 26(3), 17-22. https://doi.org/10.14313/par_245/17

[28] Li, S., & Zhang, X. (2022). The welding tracking technology of an underwater welding robot based on sliding mode active disturbance rejection control. Assembly Automation, 42(6), 891-900. https://doi.org/10.1108/aa-07-2022-0171