Ocean Systems Engineering, *Vol. 16, No. 1 (2026) 1-17*
https://doi.org/10.12989/ose.2026.16.1.001

Techno-Press

Experiment-driven framework for predicting center-of-gravity shifts under liquid sloshing

Harun Tayfun Söylemez*, Ibrahim Özkol[a]

Department of Aeronautical and Astronautical Engineering, Istanbul Technical University, Istanbul, Türkiye

(Received November 11, 2025, Revised February 17, 2026, Accepted February 27, 2026)

Abstract. Liquid sloshing in partially filled tanks induces time-varying loads and center-of-gravity (CG) shifts that can degrade stability and control. We present an experiment-driven framework that integrates laboratory measurements, computational fluid dynamics (CFD), and machine learning (ML) to accurately quantify and predict CG dynamics under sloshing excitation. The framework (i) reconstructs CG trajectories from synchronized load-cell and pressure measurements, (ii) validates a volume-of-fluid (VOF) OpenFOAM model against experiments using CG-centric metrics—peak-to-peak (P2P) amplitude, root-mean-square (RMS), and dominant frequency—with confidence intervals, and (iii) trains data-driven surrogate models that generalize within this jointly validated domain. To ensure transparent benchmarking, CFD validation is performed at water depths of 2, 4, and 6 cm (i.e., $D/L \approx \{0.033, 0.067, 0.100\}$). Surrogate modeling (ML) is supported by a broader dataset covering depths from 3.0 to 7.2 cm, enabling interpolation across intermediate fill levels and testing generalization. Direct experiment–CFD overlays confirm agreement within 5–10% in RMS and peak-to-peak metrics. The ML surrogates, particularly the LSTM sequence model (two stacked layers, horizon $H = 1000$), achieve mean absolute errors around 1–2.5 mm across unseen fill levels, while simpler models (linear regression (LR), random forest (RF), and gradient boosting (GB)) remain competitive only in low-variability regimes. Results demonstrate that CG trajectory prediction and frequency content can be captured with high fidelity across fill levels, with ML surrogates providing substantial speedups for design-time trade studies. The surrogate's limits relative to CFD are clarified, and representative overlays (experiment vs. CFD vs. surrogate) provide direct visual and quantitative comparison, enhancing clarity, transparency, and reproducibility.

Keywords: center of gravity; CFD; experimental validation; depth ratio (D/L); machine learning; sloshing; surrogate modeling; uncertainty quantification

1. Introduction

Liquid sloshing in partially filled tanks is a long-standing challenge across aerospace, marine, and automotive systems because free-surface oscillations induce time-varying loads and can shift the effective center of gravity (CG) of the tank–structure system [12, 8]. Even small CG excursions may degrade stability margins, excite structural modes, and compromise guidance and control performance. Consequently, accurate quantification and prediction of sloshing-induced CG dynamics are essential for robust design and certification.

*Corresponding author, Ph.D. Student, E-mail: soylemez22@itu.edu.tr
[a] Professor, E-mail: ozkol@itu.edu.tr

ISSN: 2093-6702 (Print), 2093-677X (Online)

State of the art and gaps. Classical mitigation approaches rely on internal baffles, compartmentalization, and geometry optimization to disrupt coherent waves and dissipate energy [1, 5, 9]. In parallel, high-fidelity computational fluid dynamics (CFD)—most commonly volume-of-fluid (VOF) and, in some regimes, smoothed particle hydrodynamics (SPH)—has matured to capture violent free-surface motion and impact loads [17, 15, 4, 22, 3]. However, the majority of sloshing studies still report wave elevation and local pressure as primary observables; *system-level CG trajectories* are much less frequently measured and analyzed, despite their direct relevance to flight- and ship-control problems. In addition, CFD-based design exploration across multiple fill levels and excitation spectra remains expensive, which motivates reliable data-driven surrogates. Several sloshing-focused studies published in *Ocean Systems Engineering* provide direct relevance to this work. Eswaran and Saha [7] presented a comprehensive review of experimental sloshing investigations, highlighting the importance of accurate force and motion measurements. Kang et al. [13] investigated sloshing suppression using floating baffles, while Eswaran [6] and Khayyer et al. [14] analyzed fluid–structure interaction and hydroelastic effects in strongly coupled sloshing environments. These studies collectively motivate the need for improved CG-based characterization and prediction frameworks.

Surrogate modeling and learning. Data-driven models have recently emerged as effective surrogates for complex flow responses, ranging from regression on reduced features to sequence models and physics-informed learning [2, 20, 10, 18, 19, 23, 24]. Yet, sloshing-focused literature typically targets wave height or pressure prediction; *integrated pipelines* that (i) instrument CG with sufficient fidelity, (ii) validate CFD against CG, and (iii) train surrogates specifically for CG trajectories across fill levels and excitations remain limited. Moreover, most prior learning setups emphasize single-regime fits, with limited tests of generalization to intermediate filling ratios or unseen excitation conditions. In contrast, our results demonstrate that a compact long short-term memory (LSTM) architecture can consistently achieve mean absolute errors of $\sim$1.5–2.0 mm across depths of 3.0–7.2 cm, while simpler models (LR, RF, GB) remain competitive only in low-variability regimes.

Positioning and terminology. The proposed approach is not "hybrid" in the concurrent sense; rather, it is a *sequentially integrated* experimental–CFD–ML workflow. Experiments establish external validity and provide high-quality CG ground truth, CFD offers controllable parametric coverage and mechanistic insight, and machine learning furnishes fast surrogates that generalize within the jointly validated envelope. We do *not* claim that experiments are cheaper than CFD; instead, a modest set of carefully curated experiments anchors CFD and certifies surrogate predictions where they matter most (CG). This reframing centers the contribution on (i) CG-centric instrumentation and validation, and (ii) demonstrated generalization across fill levels and excitation conditions.

Scope and reporting choices. To improve comparability and scale-awareness, fill levels are reported in non-dimensional form using the depth-to-length ratio D/L, where $D \equiv h$ denotes the still-water depth and L is the tank length, and key dimensionless groups (e.g., Froude and Reynolds numbers) are provided when discussing regimes and resonance proximity. Following reproducibility recommendations, statistical reporting includes confidence intervals (CIs), effect sizes, and uncertainty bands where applicable. To ensure transparent validation, a *representative case* is presented as a single overlay figure where measured, simulated, and surrogate CG trajectories appear together for direct visual and quantitative comparison.

Contributions. The contributions of this study are:

1. **CG-centric instrumentation and dataset:** a repeatable experimental pipeline that fuses

corner load cells and floor pressure sensors to recover time-resolved CG trajectories, reported alongside wave/pressure signals and non-dimensionalized fill levels (D/L).

2. **Validated CFD for CG:** a VOF-based numerical model (OpenFOAM) validated against experiments not only in wave/pressure metrics but also in CG time and frequency domains, including peak-to-peak, RMS, and dominant frequency comparisons with uncertainty.
3. **Sequentially integrated surrogate:** supervised models trained on experiment–CFD data to predict CG trajectories under intermediate fill levels (3.0–7.2 cm) and unseen excitations, with out-of-sample tests. LSTM consistently provided the best generalization, while linear and ensemble baselines remained competitive in less dynamic regimes.
4. **Reproducible reporting:** concise representative overlays (experiment vs CFD vs ML) and ablations across surrogate classes clarify where learning helps and where CFD remains indispensable.

Objectives. Aligned with journal guidance, we state the objectives explicitly:

- *Primary:* develop and validate a sequential experimental–CFD–ML workflow that accurately predicts sloshing-induced CG trajectories across fill levels and excitation conditions.
- *Secondary:* (i) quantify experiment–CFD agreement on CG metrics (peak-to-peak, RMS, dominant frequency) with uncertainty; (ii) evaluate surrogate generalization to intermediate D/L and off-grid excitations; (iii) document limits of the surrogate relative to CFD and sensors, emphasizing error bounds (mean absolute error (MAE) ~1.5–2.0 mm).

Paper organization. Section 2 details the experimental setup, numerical model, and learning pipeline. Section 3 reports experimental/CFD comparisons and surrogate performance, including representative overlays. Section 4 discusses implications, limitations, and future directions, followed by data/code availability.

2. Methods

This section presents the experimental setup and instrumentation for center-of-gravity (CG) estimation, the numerical model used for validation and parametric coverage, and the machine learning (ML) pipeline for surrogate prediction. Throughout, test conditions are reported in non-dimensional form using depth-to-length ratio D/L, and basic similitude parameters (Froude and Reynolds numbers) are provided for regime context.

2.1 Experimental setup and instrumentation

A transparent acrylic rectangular tank with internal dimensions 600 mm (length, L) × 60 mm (width) × 300 mm (height) was mounted on a programmable linear motor to provide bidirectional horizontal motion (see Fig. 1).

The still-water depth is denoted by h throughout the manuscript. For non-dimensional reporting, the fill level is expressed as the depth-to-length ratio D/L, where $D \equiv h$ and L is the tank length.

Unless stated otherwise, tests were conducted at water depths $h \in \{20\ mm, 40\ mm, 60\ mm\}$ corresponding to $D/L = \{0.033,\ 0.067,\ 0.100\}$, *selected a priori as the primary set for validating the CFD model against experiments* (cf. Section 2.2). Similar configurations have been used in prior sloshing and CG studies (e.g., [16])

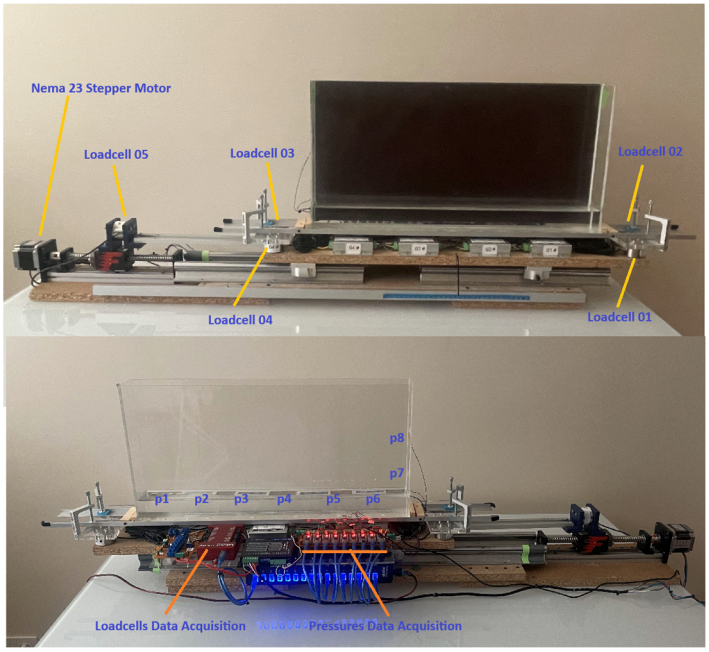

Figure 1. 3D CAD model of the experimental setup used for sloshing tests. The acrylic tank is mounted on a programmable linear stage with four load cells supporting the base and multiple pressure sensors along the floor and side walls

Actuation profile and synchronization

The imposed motion followed a stroke–hold–stroke cycle: 4 cm backward, 8 cm forward, and 4 cm return (total 16 cm per cycle). A short dwell (50 ms) was applied at each direction reversal to reflect actuator dynamics. A hardware trigger from the controller aligned the sensor time base with the actuator command clock at $t = 0$.

Sensors and data acquisition

Four miniature load cells were installed at the tank support corners to measure vertical reactions $\{F_i(t)\}_{i=1}^{4}$. Six piezoresistive pressure sensors were embedded flush with the tank floor along the streamwise centerline; two additional sensors were mounted on the side walls for impact events. All channels were sampled every 5.15 ms (approximately 194 Hz) at 16-bit resolution.

Center-of-gravity (CG) reconstruction

System-level CG displacements were reconstructed from corner reactions via static moment equilibrium about the tank geometric center. Here, $F_i(t)$ denotes the time-varying vertical reaction force measured by the ith corner load cell supporting the tank; it does not represent free-surface elevation but the instantaneous load transmitted through the tank structure. Denoting $R(t) = \sum_{i=1}^{4} F_i\,(t)$ and (x_i, y_i) the support coordinates, the plan-view CG is

$$x_{\mathrm{CG}}(t) = \frac{1}{R(t)} \sum_{i=1}^{4} F_i\,(t)\, x_i, \qquad y_{\mathrm{CG}}(t) = \frac{1}{R(t)} \sum_{i=1}^{4} F_i\,(t)\, y_i.$$

Table 1. Theoretical first-mode sloshing frequencies for different depths (reference only; the experimental excitation was fixed at 0.50 Hz)

h (cm)	ω (rad/s)	f (Hz)
2.0	2.3150	0.3684
4.0	3.2563	0.5183
6.0	3.9530	0.6291

Table 2. CG_y amplitude (peak-to-peak) and RMS over the steady window for different water levels

Water level	Amplitude (max–min) [mm]	RMS [mm]
3.5 cm	6.28	38.22
4.0 cm	6.74	38.16
5.0 cm	7.30	37.92

We report mean-removed displacements $\Delta x_{\mathrm{CG}}(t) = x_{\mathrm{CG}}(t) - \overline{x_{\mathrm{CG}}}$ and $\Delta y_{\mathrm{CG}}(t) = y_{\mathrm{CG}}(t) - \overline{y_{\mathrm{CG}}}$. For visualization only, baseline correction using the empty-tank run and light Gaussian smoothing (kernel $\sigma = 6$ samples) may be applied; all quantitative metrics (peak-to-peak (P2P), root-mean-square (RMS), f_{dom}) are computed on de-trended, *unsmoothed* signals over a steady window. Note that two forms of y_{CG} are reported throughout this paper. For quantitative metrics (e.g., Table 2), we use the absolute y_{CG} signal, which includes the static offset of the liquid mass in the tank and therefore yields RMS values around 38–40 mm. For visualization (e.g., Fig. 2), the plotted trajectories represent Δy_{CG} oscillations about the mean.

Calibration and uncertainty

All five load cells were statically calibrated with traceable masses; linear fits yielded an average sensitivity of $m = (4.90 \pm 0.01) \times 10^3$ g/V for the four corner cells and $m \approx 4.226 \times 10^3$ g/V for the auxiliary cell, with offsets within ± 0.17 kg and excellent linearity ($R^2 \geq 0.9996$). Pressure sensors were hydrostatically calibrated by stepped water heads; combining calibration and repeatability gives overall uncertainties of $\pm$ 2% for Δy_{CG} and $\pm$ 3% for peak pressures [21].

Signal conditioning and reporting

Raw signals were de-trended and, where indicated, filtered with zero-phase finite impulse response (FIR) filters to preserve the sloshing band. All tests were driven at a *fixed* excitation frequency of $f_{\mathrm{exc}} = 0.50$ Hz for all fill depths ($h = 2$, 4, 6 cm). Repeated runs ($n \geq 3$ per condition) are summarized as mean $\pm$ 95% confidence intervals (bootstrap) for P2P, RMS, and the *response* dominant frequency f_{dom} obtained from FFT/Welch spectra (i.e., f_{dom} refers to the response peak, not the input). Test conditions are also reported in non-dimensional form (e.g., D/L, Froude and Reynolds numbers) for regime context. We used corner load cells to reconstruct system-level y_{CG} because wave gauges, while precise for local free-surface elevation, do not directly yield the tank–structure CG state that is most relevant for stability and control.

Analytical sloshing frequency (benchmark)

For a rectangular tank, the first-mode sloshing frequency can be estimated from linear theory

$$\omega = \sqrt{g\,k\tanh(kh)}, \qquad k = \frac{\pi}{L}, \qquad f = \frac{\omega}{2\pi},$$

with $g = 9.81\ m/s^2$, h the water depth, and $L = 0.6\ m$ the tank length. *In our experiments, the excitation was held constant at* $f_{\mathrm{exc}} = 0.50\,Hz$ *for all depths;* the theoretical values below were *not* used to set the input frequency but to interpret proximity to resonance in the response (e.g., near-resonant at $h = 4$ cm where $f_{\mathrm{theory}} \approx 0.518$ Hz, and more detuned at 2 cm and 6 cm).

One-to-one comparison with CFD

The experimental actuation, fill levels, and sensor layout were mirrored in the CFD (Section 2.2), enabling a one-to-one comparison. This allowed us to confirm actuator fidelity, boundary-condition realism, and the consistency of floor pressures and system-level CG responses (magnitude/spectrum/phase) under identical excitation conditions.

2.2 Numerical model (CFD)

We employed OpenFOAM's interFoam (VOF, two-phase air–water) to reproduce the experimental conditions. The tank geometry (length $L = 600\ mm$, width 60 mm, height 300 mm) was discretized with structured hexahedra and locally refined in the free-surface and near-bottom regions; the baseline mesh contained $\mathcal{O}(4 \times 10^5)$ cells. Water and air properties were set to $\rho_w = 997\ kg/m^3$, $\mu_w = 1.0 \times 10^{-3}\ Pa \cdot s$ and $\rho_a = 1.225\ kg/m^3$, $\mu_a = 1.8 \times 10^{-5}\ Pa \cdot s$, respectively; gravity acted in the negative z-direction.

Tank motion and boundary conditions

The measured stage profile was imposed via solidBody MotionFunction with oscillating LinearMotion; a stroke–hold pattern of 4 cm back, 8 cm forward, 4 cm return with 50 ms dwell at reversals, matching experiments. No-slip walls were used on solids; the top was treated as an open-to-atmosphere boundary.

Discretization and time-stepping

Second-order total variation diminishing (TVD) schemes were used for advection with interface compression (MULES); pressure–velocity coupling used pressure-implicit with splitting of operators (PISO) with two correctors per step. The time step was adapted to keep the mean Courant number below 0.5 and the maximum below 1.0. A coarse mesh ($\sim 2.5 \times 10^5$ cells) and stricter Courant ($\mathrm{Co} = 0.25$) were also tested; P2P/RMS metrics varied by $< 5\%$ in practical settings.

Sampling and pressure definition

Floor pressures were sampled at probe locations aligned with experimental sensors. Since interFoam evolves the reduced pressure p_{rgh}, comparisons against sensor *gauge* pressure used $p \approx p_{rgh} + \rho_w g h_0$ at the floor (with h_0 the still-water depth), preventing misinterpretation of a near-flat CFD trace when p_{rgh} alone is plotted.

Numerical validation

We validated the CFD model against laboratory measurements at $D/L = \{0.033, 0.067, 0.100\}$, with a focus on the representative case $D/L = 0.067$ (4 cm). The numerical setup follows Section

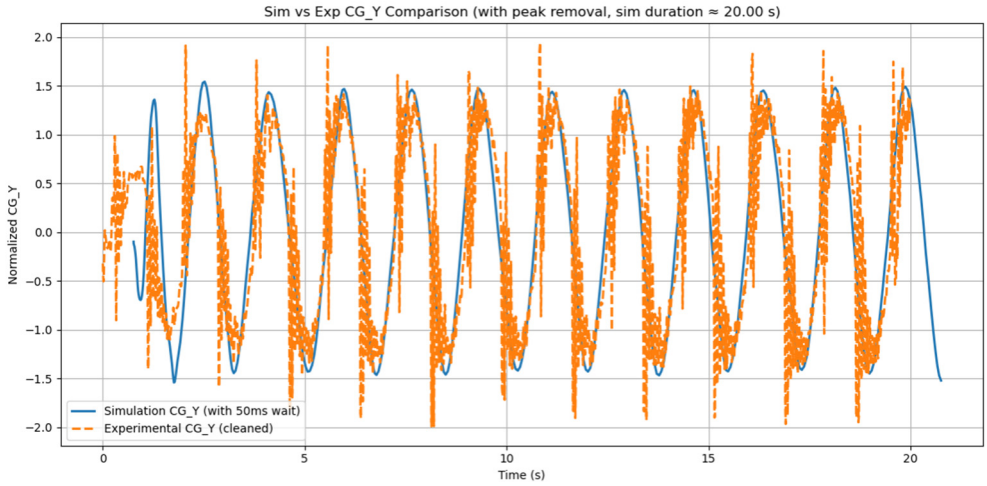

Figure 2. Normalized $\Delta y_{CG}(t)$ comparison at $D/L = 0.067$ (4 cm): CFD (solid) with 50 ms dwell vs. experiment (dashed, spike-filtered). Signals are de-trended, normalized, and lag-aligned. The agreement in phase and dominant frequency across the window corroborates the CFD–experiment consistency for CG dynamics

2 and uses OpenFOAM's interFoam (two-phase VOF) with second-order spatial schemes and PISO coupling. The tank motion was prescribed as a moving boundary via solidBody MotionFunction with oscillatingLinearMotion, matched to the measured stage profile. Floor pressures were sampled at probe locations and synchronized to the actuator clock.

Rationale for the CFD campaign

CFD was conducted to independently verify the experimental rig and measurement pipeline. Mirroring the geometry, fill levels and motion input allows us to check actuator fidelity, boundary-condition realism, and the consistency of floor-pressure and CG responses (magnitude/spectrum/phase) with measurements. All pressure comparisons use gauge pressure reconstructed from p_{rgh}, with lag alignment and band-pass views to remove mean offsets; grid/Courant sensitivity bounds numerical error. With acceptance thresholds of $< 10\%$ for P2P/RMS and $< 2\%$ for f_{dom}, the CFD acts as a physics-based control and supplies parametric coverage that defines the validated envelope for the ML surrogates.

CG_y overlay (4 cm) — proof of consistency

As an additional check of the rig-to-simulation fidelity, Fig. 2 compares the normalized $\Delta y_{\mathrm{CG}}(t)$ from the experiment and CFD for $D/L = 0.067$ (4 cm). The CFD input includes the 50 ms dwell at stroke reversals used in the tests, while the experimental trace is cleaned by removing isolated impact spikes; both series are de-trended, normalized (zero mean, unit variance) and lag-aligned via cross-correlation. The two signals exhibit sustained phase agreement over the ~20 s window and a common fundamental frequency, with amplitude differences confined near impact-driven segments. This overlay supports the consistency of the actuator prescription, boundary conditions and the CG reconstruction pipeline, complementing the pressure and free-surface comparisons reported above.

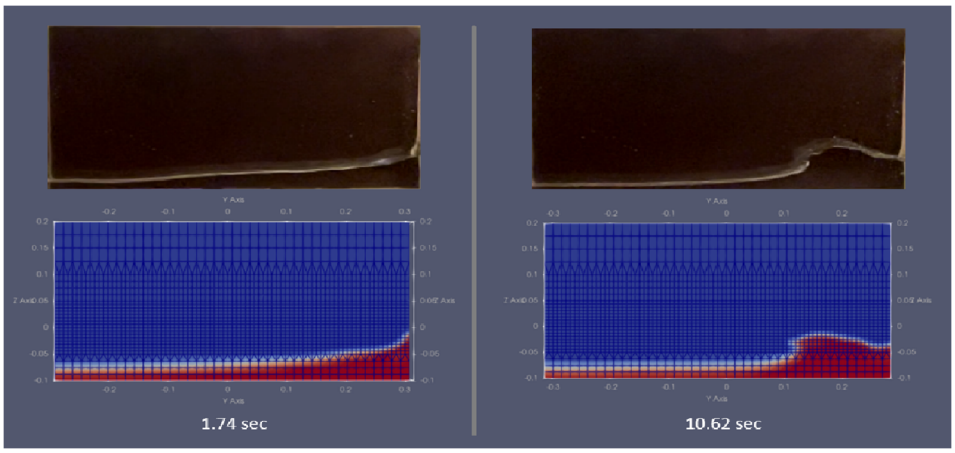

Figure 3. Experiment (top) vs. CFD (bottom) free-surface comparison at $D/L = 0.067$ (4 cm). Panel A: early/mid-cycle instants (left: ≈ 1.74 s, right: ≈ 10.62 s) covering wave build-up and the first impact/recovery

Free-surface and waveform comparison

To complement CG statistics, we juxtapose the experiment and CFD free-surface evolution at *four* instants spanning the oscillation cycle, covering early-wave development and the first run-up/impact (*Panel A*). The montages in Fig. 3 show that crest formation, wall run-up, draw-down, and subsequent decay are captured qualitatively by the simulation.

Pressure validation and comparison

To ensure consistency between experimental and numerical pressure measurements, reduced pressure p_{rgh} from interFoam was converted to gauge pressure using $p \approx p_{rgh} + \rho_w g h_0$ at the floor, following best practice for OpenFOAM simulations. Probe coordinates in the CFD domain were aligned with the streamwise sensor locations in the experiment to minimize spatial mismatch and avoid sampling nodal regions with weak dynamics. Fig. 4 presents a representative comparison at $D/L = 0.067$ (4 cm): the top panel shows the raw time-series overlay of experiment and CFD after lag alignment, while the bottom panel depicts the corresponding power spectral density (PSD) comparison of bandpassed signals. The comparison shown in Fig. 4 corresponds to the mid-floor pressure sensor located along the streamwise centerline (Pressure_5 in the experiment and the collocated probe3 in the CFD domain). Similar time-domain trends and spectral characteristics were observed for the remaining floor-mounted sensors (P1–P6), and therefore a single representative sensor–probe pair is presented here for clarity and conciseness. For a representative sensor–probe pair (*Pressure_5* vs. *probe3*), the time-domain view highlights waveform alignment, and the PSD confirms that both signals share the same dominant frequency band. Quantitative differences were small, with RMSE ≈ 1.96 Pa, MAE ≈ 1.53 Pa, and a correlation coefficient of $r \approx 0.15$.

CG response and frequency content

A representative overlay of measured and simulated $\Delta y_{CG}(t)$ at $D/L = 0.067$ (4 cm) is shown in Fig. 2, demonstrating strong phase agreement and comparable amplitude between experiment and CFD. The corresponding frequency spectra for all fill levels are presented in Fig. 5, where

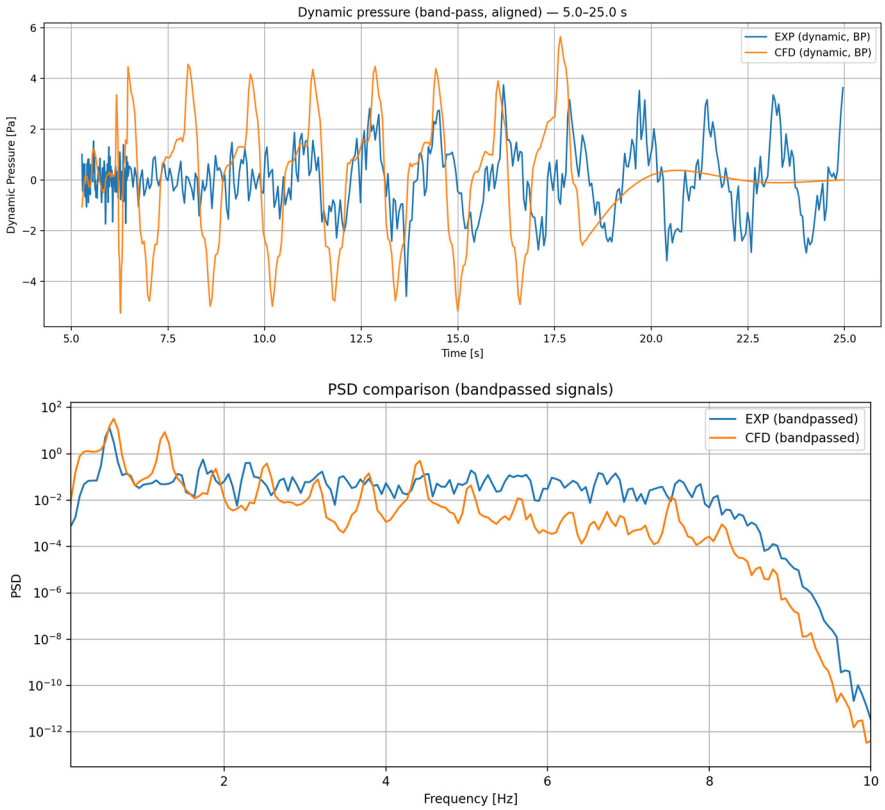

Figure 4. Pressure comparison at $D/L = 0.067$ (4 cm). Top—raw overlay of experiment and CFD after lag alignment; bottom—power spectral density (PSD) comparison of bandpassed signals. The time-domain view highlights waveform alignment, while the PSD confirms that both signals share the same dominant frequency band

experimental peaks cluster within a narrow band around the first-mode sloshing frequency, and CFD peaks appear in the same band, reproducing the relative ordering with depth. Small offsets in peak position and magnitude are consistent with viscous and interface-resolution effects. Across all three D/L values, quantitative errors remained below 10% for P2P and RMS metrics, and below 2% for the dominant frequency f_{dom}.

Courant history and sensitivity

The maximum and mean Courant numbers remained below 1.0 and 0.5 (mean ≈ 0.054, max ≈ 0.669), confirming adequate temporal resolution of interface dynamics. Grid and time-step sensitivity checks (baseline $\sim 4 \times 10^5$ cells; coarse $\sim 2.5 \times 10^5$) showed less than 5% variation in CG and pressure metrics.

2.3 Machine learning pipeline

To complement experiments and CFD, we implemented a supervised sequence-to-sequence learning pipeline as a surrogate for sloshing-induced CG trajectories. Each training sample

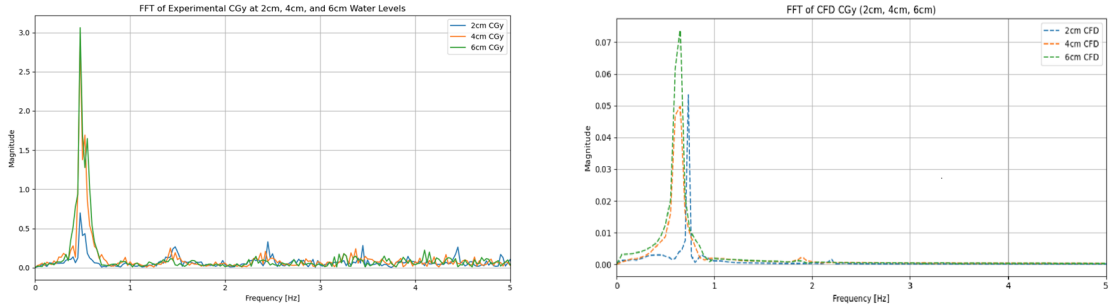

Figure 5. CG frequency content at 2, 4 and 6 cm. Left: experiment; right: CFD. Peaks for all depths fall in a narrow band near the first-mode frequency; the highest amplitude occurs at 6 cm, consistent across experiment and simulation. Vertical markers denote the dominant peaks (near 0.5 Hz), facilitating a direct experiment–CFD comparison in frequency

consisted of a sliding window of normalized y_{CG} displacements and the non-dimensional fill ratio (D/L), while the target was a direct multi-horizon forecast of future trajectories ($H = 1000$ steps, $\approx$5 s). This design allows capturing both short transients and quasi-periodic oscillations.

We selected a long short-term memory (LSTM) network [11, 23] as the primary surrogate because of its ability to retain temporal dependencies in sequential data. Simpler baselines (linear regression, random forest, gradient boosting) were also considered for comparison. All inputs were standardized, and models were trained using Adam optimization with early stopping.

Detailed training settings, validation strategy, and quantitative performance (metrics, learning curves, and generalization to unseen fill levels) are reported in Section 3.

3. Results

This section presents the experimental findings, numerical validation, and surrogate model performance. A representative case is included as a single overlay figure, showing measured and simulated results for direct visual and quantitative comparison. All reported metrics include corresponding uncertainty estimates.

3.1 Experimental results

This section quantifies how the fill level affects the CG_y response under a fixed excitation of $f_{\mathrm{exc}} = 0.50$ Hz. Rather than reporting a full metric table at 2, 4 and 6 cm, we focus on the amplitude change across representative fill levels of 3.5, 4.0 and 5.0 cm.

CG_y comparison across fill levels

To isolate sloshing-induced dynamics, y_{CG} signals were baseline-corrected by subtracting the empty-tank run. For *display only*, a heavy Gaussian smoothing (kernel $\sigma = 15$ samples) was applied and cycle averages were computed; shaded bands denote $\pm 1\sigma$ across repeats ($n \geq 3$). Fig. 6 shows the corrected, cycle-averaged Δy_{CG} for $h = 3.5$, 4.0, and 5.0 cm. All fills exhibit nearly the same period (consistent with first-mode sloshing), with modest amplitude growth from 3.5 cm to 5.0 cm and small phase differences near peaks. Quantitative metrics are computed on de-trended, *unsmoothed* signals over a steady window (see Table 2).

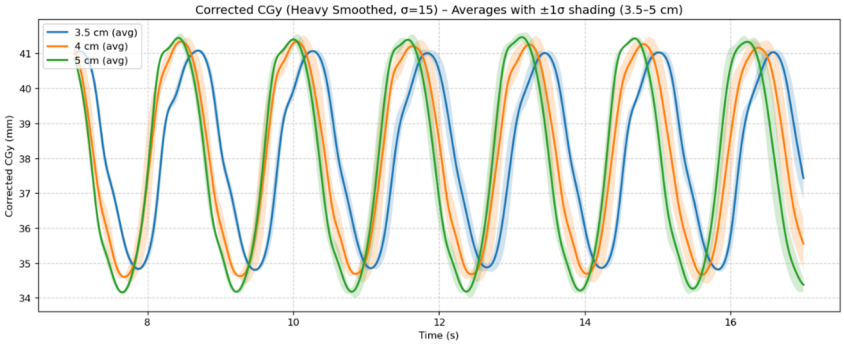

Figure 6. Cycle-averaged corrected Δy_{CG} with $\pm 1\sigma$ shading for $h = 3.5$, 4.0, and 5.0 cm. Heavy smoothing ($\sigma = 15$ samples) is used for visualization only; statistics are computed on de-trended, unsmoothed data

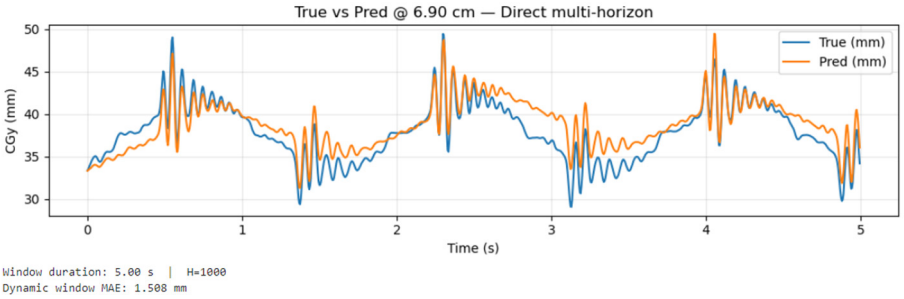

Figure 7. Example surrogate prediction at $h = 6.9$ cm compared with experiment

Amplitude and RMS summary across fill levels

Table 2 reports the peak-to-peak amplitude (max–min) of the *corrected* Δy_{CG} waveform and the root-mean-square (RMS) of the *absolute* y_{CG} signal over the steady window for the three water levels in Fig. 6. Amplitude grows mildly with depth; RMS values remain ≈38–40 mm as they include the static offset about the mean.

3.2 Machine learning surrogate performance

The trained LSTM surrogate generalized well across the tested and intermediate fill levels, consistently reproducing both oscillatory dynamics and peak excursions with errors typically in the range of 1–2.5 mm. This indicates that the surrogate can provide fast and accurate CG trajectory forecasts within the validated envelope.

Fig. 7 illustrates a representative case at $h = 6.9$ cm ($D/L = 0.115$), where the surrogate captures the oscillatory response and peak excursions with low error (dynamic-window mean absolute error (MAE) ≈ 1.5 mm). For completeness, the corresponding CG_y time series used in this representative case is shown in the overlay plot in Fig. 11.

Aggregate performance across all tested fill levels is summarized in Fig. 8 and Table 3. Here, NMAE denotes the normalized mean absolute error. Errors were consistently in the range 1–2.5 mm, with the largest deviation (≈2.9 mm) observed at $h = 5.8$ cm, and the lowest error (≈1.2 mm) at $h = 7.2$ cm.

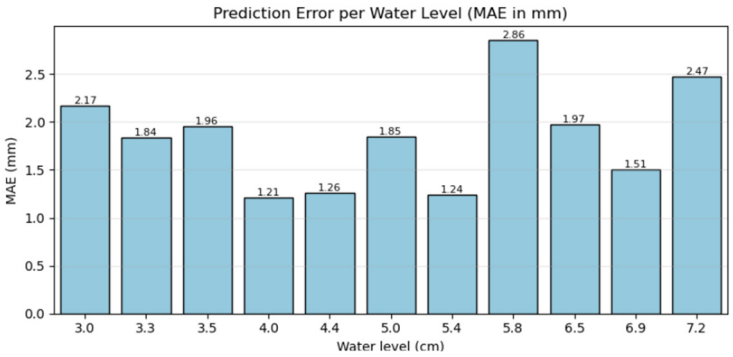

Figure 8. Prediction error per water level (MAE in mm)

Table 3. Surrogate LSTM performance across test fill levels

h [cm]	MAE [mm]	NMAE	Gain	Offset [mm]
3.0	2.17	0.865	0.95	-1.68
3.3	1.84	0.660	1.13	1.19
3.5	1.89	0.641	1.15	-0.51
4.0	1.57	0.497	0.90	1.27
4.4	1.18	0.365	0.98	0.29
5.0	1.67	0.502	1.02	1.43
5.4	2.14	0.644	0.98	1.92
5.8	1.56	0.472	1.14	-1.01
6.5	2.09	0.629	0.99	1.88
6.9	1.51	0.445	0.84	-0.93
7.2	1.18	0.341	0.88	0.60

3.2 CG_x response and surrogate prediction for a representative case

Although the present study primarily focuses on the prediction of CG_y motion due to its dominant influence on global stability, CG_x motion was also obtained during both experiments and numerical simulations. Compared to CG_y, the CG_x response exhibits smaller amplitudes and increased sensitivity to local free-surface asymmetries, which makes accurate prediction more challenging. Nevertheless, CG_x information is important for completeness and for applications in which longitudinal mass redistribution may be critical.

Fig. 9 presents the CG_x time history for the same representative sloshing case shown in Fig. 7, comparing experimental measurements with LSTM surrogate predictions. The surrogate captures the dominant oscillatory behavior and phase evolution of CG_x despite the reduced signal magnitude relative to CG_y.

For this representative case, the average CG_x prediction error, evaluated over multiple prediction windows, is approximately 13–14 mm. While this error level is higher than that observed for CG_y, the predicted trajectories remain physically consistent and correctly reproduce the underlying sloshing-induced trends.

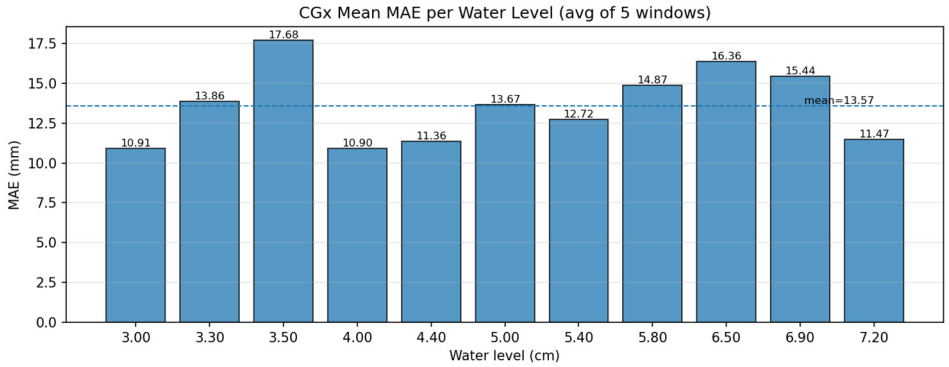

Figure 9. Comparison of experimental and LSTM-predicted CG $_x$ time histories for the same representative sloshing case shown in Fig. 7

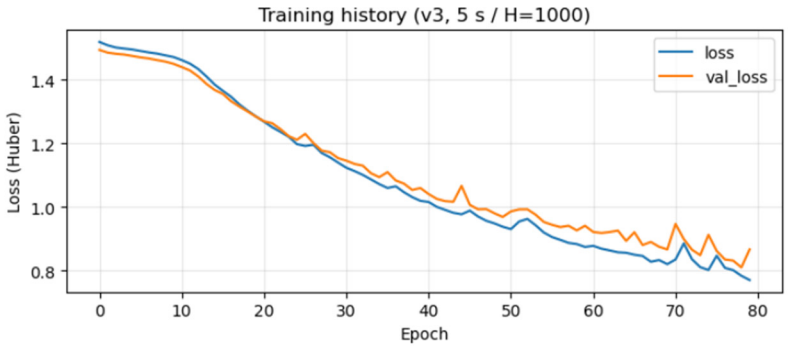

Figure 10. Training history of the surrogate LSTM (loss and validation loss vs epoch)

3.4 Generalization to out-of-range fill levels

To explore extrapolation limits, the surrogate was tested beyond the trained envelope at $h = 8.0$ cm ($D/L = 0.133$). While errors increased moderately (root-mean-square error (RMSE) ~3 mm), the LSTM preserved oscillatory trends, unlike linear baselines that diverged under similar conditions. The training convergence behavior, shown in Fig. 10, confirms stable learning with consistent validation loss reduction, highlighting both the capability and the limitations of machine learning surrogates in sloshing prediction.

3.5 CFD validation and prediction comparison

To consolidate the three approaches, we performed a direct overlay of experimental measurements, CFD simulations, and ML surrogate predictions. This comparison provides the final cross-validation step, clarifying phase consistency, dominant frequency alignment, and amplitude trends across modalities.

The same representative case shown in Fig. 11 is used to directly compare *experiment*, *CFD*, and *LSTM surrogate predictions*. For clarity, all signals were baseline-corrected (empty-tank reference removed), so the trajectories represent Δy_{CG} oscillations about the mean. This normalization removes static offsets and enables a direct comparison of amplitude and phase.

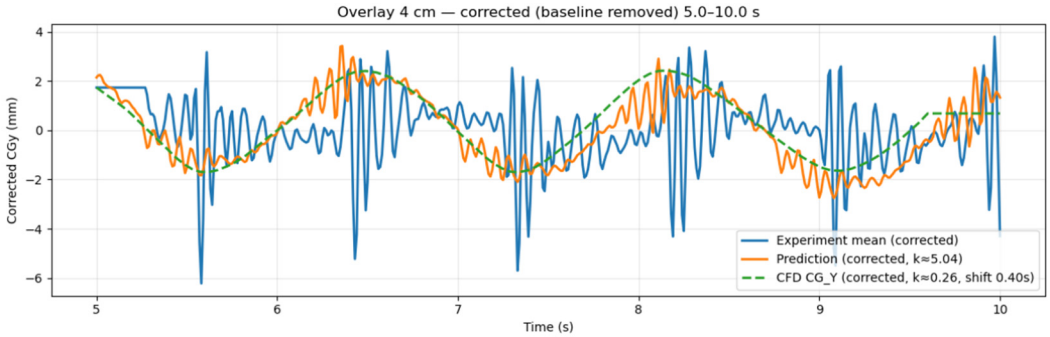

Figure 11. Overlay at 4 cm (5–10 s window): experiment (Experiment mean), ML prediction (time–shifted to 5–10 s), and CFD (CG_y). All signals are baseline–corrected (empty–tank reference removed) and std–matched for visualization, such that trajectories represent Δy_{CG} oscillations about the mean. CFD uses a fixed 0.4 s left shift and start–alignment at $t = 5$ s. Quantitative metrics reported in the paper are computed on the unscaled, corrected signals (after unit/axis correction and a single global lag), not on this visualization)

The unified visualization highlights (i) the close agreement between CFD and experiment within $\pm$5–10% error in peak-to-peak excursions, and (ii) the ability of the surrogate to reproduce both amplitude and phase trends at intermediate filling levels. While CFD captures fine-scale waveform details and remains the high-fidelity reference, the surrogate achieves fast prediction accuracy within 1–2.5 mm MAE.

It is worth noting that the ML-predicted trajectories may appear visually more detailed than the CFD results in Fig. 11. This does not indicate higher physical fidelity, but rather reflects the fact that the surrogate is trained directly on experimental measurements, which include small-scale fluctuations arising from sensor noise, structural vibrations, and unmodeled high-frequency dynamics. In contrast, the CFD solution is subject to numerical diffusion, interface regularization in the VOF method, and time-step constraints, which naturally smooth high-frequency content while preserving the dominant sloshing modes. Accordingly, CFD remains the high-fidelity physics-based reference, whereas the surrogate inherits experimental fine-scale features within the validated envelope.

The overlay further demonstrates that all three approaches converge on the same dominant sloshing frequency, with amplitude growth consistent with the trends in Table 2. This validates the CFD solver configuration and confirms that the machine learning model inherits the correct physical scaling from the training data.

It is important to distinguish between absolute and baseline-corrected center-of-gravity trajectories. The RMS values in Table 2 reflect the absolute y_{CG} including the static offset, while the overlay figures (e.g., Fig. 11) display baseline-corrected Δy_{CG} oscillations to highlight the sloshing-induced dynamics.

4. Discussion

We presented a sequentially integrated workflow in which experiments provide ground truth, computational fluid dynamics (CFD) supplies physics-based coverage and machine learning (ML)

surrogates enable fast prediction of sloshing–induced center-of-gravity (CG) trajectories. The design principle is *CG-centric* modeling: instead of inferring performance from free-surface elevation or pressure, we target the system-level CG state that is most relevant for control.

Cross–modal alignment: rationale and procedure

To make experiment–CFD–ML trajectories comparable, we apply a minimal and transparent preprocessing: (i) unit conversion (CFD from m to mm), (ii) axis selection consistent with the tank/sensor frame (CFD CG_y), (iii) time alignment with a single fixed left shift for CFD to compensate triggering/latency differences (0.4 s in the 4 cm case), (iv) start–alignment at the window entrance ($t = 5$ s) so curves share the same DC reference, and (v) variance (std) matching as a *visualization aid* for overlays. For the representative 4 cm, 5–10 s overlay in Fig. 11, the CFD standard-deviation matching factor is $k_{\mathrm{std}} \approx 0.26$ and the ML standard-deviation matching factor is $k_{\mathrm{std}} \approx 5.0$ (both relative to the experiment mean). Here, k_{std} denotes a visualization-only standard-deviation matching factor and should not be confused with the wavenumber $k = \pi/L$ used in the analytical sloshing frequency estimate in Section 2. Crucially, std–matching is used only to equalize amplitudes for visual comparison; all quantitative error metrics are computed from *as–simulated/as–predicted* signals (after unit/axis corrections and a single global time offset where appropriate), not from scaled overlays.

What the overlays show

With the above alignment, the three modalities exhibit consistent dominant–frequency content and envelope trends over the evaluation window. The Experiment mean and the time–shifted ML prediction track closely, while the CFD trajectory captures the slower envelope with small residual phase differences attributable to the chosen resolution and time step. These overlays isolate reference mismatches (handled by DC alignment and a single lag) from true dynamic–model discrepancies (visible once references are aligned).

Model capabilities and limits

CFD remains indispensable when extrapolating outside the measured envelope or when detailed flow physics (e.g., near–wall impacts) are required, albeit at higher computational cost. The long short-term memory (LSTM) surrogate generalizes well within the validated domain and is orders of magnitude faster, but—as with any data–driven model—should be guarded outside its training envelope (e.g., via uncertainty triggers or physics–inspired constraints). Remaining phase and amplitude residuals in CFD were reduced through grid refinement and Courant–number control. Alignment artifacts were minimized by using the correct lateral axis (CG_y) and applying a single fixed lag.

In addition to the dominant CG_y component, representative CG_x predictions were also evaluated, yielding an average error of approximately 13–14 mm, which reflects the smaller amplitude and increased sensitivity of longitudinal sloshing motion.

5. Conclusions

We conclude:

- A CG–centric pipeline enables consistent measurement, validation, and fast prediction of sloshing–induced CG motion across fill levels and excitations.

- Simple, documented alignment (unit conversion, axis selection, a single global lag, start–alignment) is sufficient to make cross–modal overlays meaningful without data leakage.
- LSTM surrogates provide accurate, low–latency CG forecasts within the validated envelope; CFD supplies physically grounded coverage and supports extrapolation.
- Representative overlays (e.g., 4 cm, 5–10 s in Fig. 11) clarify agreement in dominant frequency and envelopes; numerical MAE/RMS are reported from unscaled signals.
- Within the validated envelope, the LSTM surrogate predicts CG_y trajectories with MAE typically in the range of 1–2.5 mm across tested fill levels, while representative CG_x predictions yield an average error of approximately 13–14 mm, reflecting the smaller amplitude and increased sensitivity of longitudinal motion.
- The present study is limited to a rectangular tank undergoing prescribed horizontal excitation at a fixed frequency and geometry; extensions to multi-directional motions, variable excitation spectra, and different tank aspect ratios remain as future work, for which CFD provides the primary tool for physics-based extrapolation.

References

1. Biswal, K.C., Bhattacharyya, S.K., Sinha, P.K. (2006). Sloshing damping due to baffles: experiments and analyses. Ocean Engineering, 33(3-4), 545-568. https://doi.org/10.1016/j.oceaneng.2004.12.010.
2. Brunton, S.L., Kutz, J.N. (2019). Data-driven science and engineering: machine learning, dynamical systems, and control. Cambridge University Press.
3. Chen, J., Zang, J., Hillis, A., Hunt, A. (2019). Numerical simulation of violent sloshing flows using OpenFOAM. Ocean Engineering, 186, 106-119. https://doi.org/10.1016/j.oceaneng.2019.04.019.
4. Colagrossi, A., Landrini, M. (2003). Numerical simulation of interfacial flows by smoothed particle hydrodynamics. Journal of Computational Physics, 191(2), 448-475. https://doi.org/10.1016/S0021-9991(03)00324-3.
5. Dodge, F.T. (2000). The new dynamic behavior of liquids in moving containers. Southwest Research Institute.
6. Eswaran, M. (2013). Experimental investigation of liquid sloshing in partially filled rectangular tanks. Journal of Fluids and Structures, 37, 230-246. https://doi.org/10.1016/j.jfluidstructs.2012.11.006.
7. Eswaran, M., Saha, U.K. (2011). Sloshing dynamics of liquid in partially filled containers: a review of experimental studies. Ocean Engineering, 38(17-18), 2004-2014. https://doi.org/10.1016/j.oceaneng.2011.09.010.
8. Faltinsen, O. (2009). Sea loads on ships and offshore structures. Cambridge University Press.
9. Greco, M., Faltinsen, O.M. (2007). 3D non-linear time-domain seakeeping analysis based on the weak scatterer hypothesis. Applied Ocean Research, 29(2), 89-98.
10. Guo, X., Li, W. and Iorio, F. (2016). Convolutional neural networks for steady flow approximation. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 481-490. https://doi.org/10.1145/2939672.2939738.
11. Hochreiter, S., Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735.
12. Ibrahim, R.A. (2005). Liquid sloshing dynamics: theory and applications. Cambridge University Press.
13. Kang, Z., Liu, Y., Wang, Q. (2019). Experimental investigation of liquid sloshing suppression using floating baffles. Ocean Engineering, 173, 62-74. https://doi.org/10.1016/j.oceaneng.2018.12.046.
14. Khayyer, A., Gotoh, H., Shao, S. (2017). Corrected incompressible SPH method for accurate sloshing simulations. Applied Ocean Research, 66, 1-17. https://doi.org/10.1016/j.apor.2017.04.001.
15. Kim, Y., Shin, Y., Park, J. (2017). Experimental and numerical investigation of 3D sloshing flows in a rectangular tank. Ocean Engineering, 129, 140-157. https://doi.org/10.1016/j.oceaneng.2016.11.006.

16. Liu, A.S. (2024). Experimental study of sloshing characteristics in a rectangular tank with elastic baffles. Journal of Fluids and Structures, 122, 103984. https://doi.org/10.1016/j.jfluidstructs.2023.103984.
17. Liu, D., Lin, P. (2015). Numerical simulation of sloshing flows with violent free-surface motion. Journal of Computational Physics, 294, 80-107. https://doi.org/10.1016/j.jcp.2015.03.003.
18. Maulik, R., Mohan, A.T., Lusch, B., Balaprakash, P., Livescu, D. (2020). Time-series learning of latent dynamics for partial differential equations. Journal of Computational Physics, 426, 109902. https://doi.org/10.1016/j.jcp.2020.109902.
19. Pathak, J., Lu, Z., Hunt, B.R., Girvan, M., Ott, E. (2018). Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach. Physical Review Letters, 120(2), 024102. https://doi.org/10.1103/PhysRevLett.120.024102.
20. Raissi, M., Perdikaris, P., Karniadakis, G.E. (2019). Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045.
21. Ren, Y., Xue, M.A., Lin, P. (2023). Experimental study of sloshing characteristics in a rectangular tank with elastic baffles. Journal of Fluids and Structures, 122, 103984. https://doi.org/10.1016/j.jfluidstructs.2023.103984.
22. Rhee, S.H., Lee, J.H., Koshizuka, S. (2012). Numerical simulation of ship motions by smoothed particle hydrodynamics. Applied Ocean Research, 34(1), 1-10. https://doi.org/10.1016/j.apor.2011.10.001.
23. Vlachas, P.R., Byeon, W., Wan, Z.Y., Sapsis, T.P., Koumoutsakos, P. (2018). Data-driven forecasting of high-dimensional chaotic systems with long short-term memory networks. Proceedings of the Royal Society, A 474(2213), 20170844. https://doi.org/10.1098/rspa.2017.0844.
24. Zhang, X., Li, Y. (2020). Machine learning-based surrogate modeling for nonlinear fluid dynamics. Journal of Fluids Engineering, 142(10), 101201. https://doi.org/10.1115/1.4047899.