Validation of an OpenFOAM/interFoam wave-modelling toolchain against two published wave-flume benchmarks

Harmonic transfer over a submerged bar (Beji-Battjes) and surf-zone breaking (Ting & Kirby)

Sergei Zvarich (ORCID 0009-0004-6960-7000) - Wavemorison, wave-engineering studio

wavemorison.com - hello@wavemorison.com

Technical report, v1.0 - August 2026

DOI 10.5281/zenodo.21776521 (Zenodo)

Abstract. *Before a numerical wave model is used for design decisions about a surf pool - a basin whose bottom profile and wave generation are fixed in concrete - the model should be shown to reproduce waves that were physically measured. We validate an open-source VOF toolchain (OpenFOAM v2512, interFoam, built-in wave generation and active absorption) against two published laboratory benchmarks. (1) Beji-Battjes / Luth Case A, non-breaking transformation over a submerged bar: mean surface-elevation RMSE of 2.7 mm at wave heights of 2-4 cm with correlation r = 0.944 across all 10 gauges; the second-harmonic amplitude behind the bar reaches 91% of the measurement, and a controlled numerical study shows the time-integration scheme matters about as much as spatial resolution for harmonic transfer. (2) Ting & Kirby 1994 spilling breaker on a 1:35 slope, run as a characterisation rather than a tuning exercise: wave trough, mean-water-level transition, breaking location and undertow are reproduced (the k-omega SST closure clearly improves the undertow), while the stock closure over-produces turbulent kinetic energy by a factor of 4-6 below the wave trough and up to ~30 at the surface roller - the documented RANS failure mode, reproduced and quantified. Open items, including an 18-24% incident-wave deficit in the breaking case, are stated explicitly. The result is a toolchain with known error bars rather than a gallery of agreements.*

1. Introduction

Computational fluid dynamics is routinely shown as proof that a wave can be engineered; renders of plausible water are easy to produce. The engineering question is narrower: does the specific toolchain - solver, wave generation, turbulence closure, mesh practice, post-processing - reproduce waves for which the true answer is known, and by how much does it miss where it misses? For surf-pool work the question has a deadline, because the geometry that shapes the wave is eventually poured in concrete.

This report documents two benchmark validations run in the production configuration of our toolchain. The benchmarks are published wave-flume experiments with open geometry, wave parameters and measured data, so the comparison can be re-run and challenged by anyone. Two principles were fixed in advance: **no tuning to fit -** model settings were chosen from standard practice, not adjusted to chase the measurements; and **divergence is a result -** where the model fails is reported with the same care as where it succeeds, because quantified limits are what design margins are built from.

2. Toolchain

OpenFOAM v2512 (ESI/openfoam.com), native arm64 build; solver interFoam (two-phase volume-of-fluid, water/air). Wave generation and active absorption use the built-in waveModels boundary conditions (Stokes II for the bar case; stream-function/Fenton coefficients, independently verified to <0.1% against a reference solution, for the breaking case). Adaptive time stepping is Courant-limited. Post-processing is scripted (Python/NumPy): fixed stationary averaging windows chosen by inspecting the full time series, per-wave robust statistics that reject spray artifacts, a single global phase alignment per case, and metrics computed identically for model and measurement.

3. Benchmark 1 - Beji-Battjes: harmonic transfer over a submerged bar

3.1 Case and data

A regular wave propagates over a submerged trapezoidal bar, steepens on its crest and releases higher harmonics on the lee side (Beji & Battjes 1993; measured gauge series by Luth et al. 1994, Delft Hydraulics, compiled by Dingemans 1994). This is the canonical test of whether a solver can carry the energy hand-off between harmonics that any non-trivial bottom profile - including a surf-pool bathymetry - imposes on a wave. Measured elevation series at 10 gauges were taken from the ERDC FUNWAVE-TVD validation testbed (real laboratory measurements, not model output); geometry and parameters were independently cross-checked against two further open implementations of the case.

Parameter	Value
Still-water depth	0.40 m
Wave period / height	T = 2.02 s, H = 0.02 m
Bar	height 0.30 m (0.10 m over crest), slopes 1:20 front / 1:10 rear
Gauges	10, from flat approach to lee side
Regime	non-breaking; Ursell ~ 4.4

3.2 Numerical set-up

2D channel, terrain-following mesh draped over the bar with refinement held in a fixed band around the free surface (coarse: dx = 5 cm, 36k cells; fine: dx = 2.5 cm, 130k cells). Stokes II generation with soft ramp-up; active absorption at both boundaries. Reflection was verified directly with a two-gauge Goda-Suzuki analysis: Kr = 2.6-4.5% depending on the run, so the comparison is not contaminated by a standing-wave pattern.

3.3 Results

Incident wave and shoaling. At the approach gauges the first-harmonic amplitude matches the measurement within ~5% with correlation r = 0.99, and the computed wavelength matches the dispersion prediction. **Across all 10 gauges** the best configuration reaches a mean RMSE of 2.71 mm with mean r = 0.944 (per-gauge RMSE 1.5-4.1 mm), at measured wave amplitudes of roughly 1-2.5 cm (Fig. 1).

Harmonic transfer. The hard part of this benchmark is the second and third harmonic on the lee side. At baseline numerics (first-order Euler time scheme, coarse mesh) the second harmonic behind the bar is under-predicted by ~40%. A controlled study separated the causes:

Run	Mesh / time scheme	Mean RMSE [mm]	Mean r	H2 behind bar [cm] (meas. 0.86)
coarse	dx 5 cm / Euler	3.61	0.886	0.39
fine	dx 2.5 cm / Euler	3.59	0.887	0.49
coarse_co025	dx 5 cm / Euler, halved dt	3.21	0.916	0.47

coarse_CN	dx 5 cm / Crank-Nicolson	2.99	0.925	0.62
fine_CN	dx 2.5 cm / Crank-Nicolson	2.71	0.944	0.78 (91%)

Switching the time scheme from first to second order (Crank-Nicolson) on the same coarse mesh lifted the lee-side second harmonic from 0.47 to 0.62 cm - comparable to, and slightly larger than, the effect of halving the mesh size (0.49 cm). A dedicated control run isolated a configuration confound (the VOF sub-cycling that Crank-Nicolson forbids): the sub-cycling change alone moved the harmonic the *other* way (-0.03 cm), so the entire gain is attributable to the scheme. The remaining 9% deficit on the far lee side is documented numerical dissipation of the short third-harmonic wavelengths, not a physics error: on the bar crest the second harmonic matches the measurement (Fig. 2). The practical rule - second-order time integration is co-dominant with spatial resolution for harmonic transfer - is now standing practice in our production runs.

3.4 Independent review

The case passed an independent adversarial audit (four reviewers tasked with breaking the data provenance, the numerical set-up, the analysis code and the physics against the files). The audit confirmed the data are genuine laboratory measurements and the metrics reproduce from the code; two presentation issues it found (an FFT-normalisation error in spectrum plots that did not affect metrics, and an over-strong wording of the scheme conclusion) were corrected before this report.

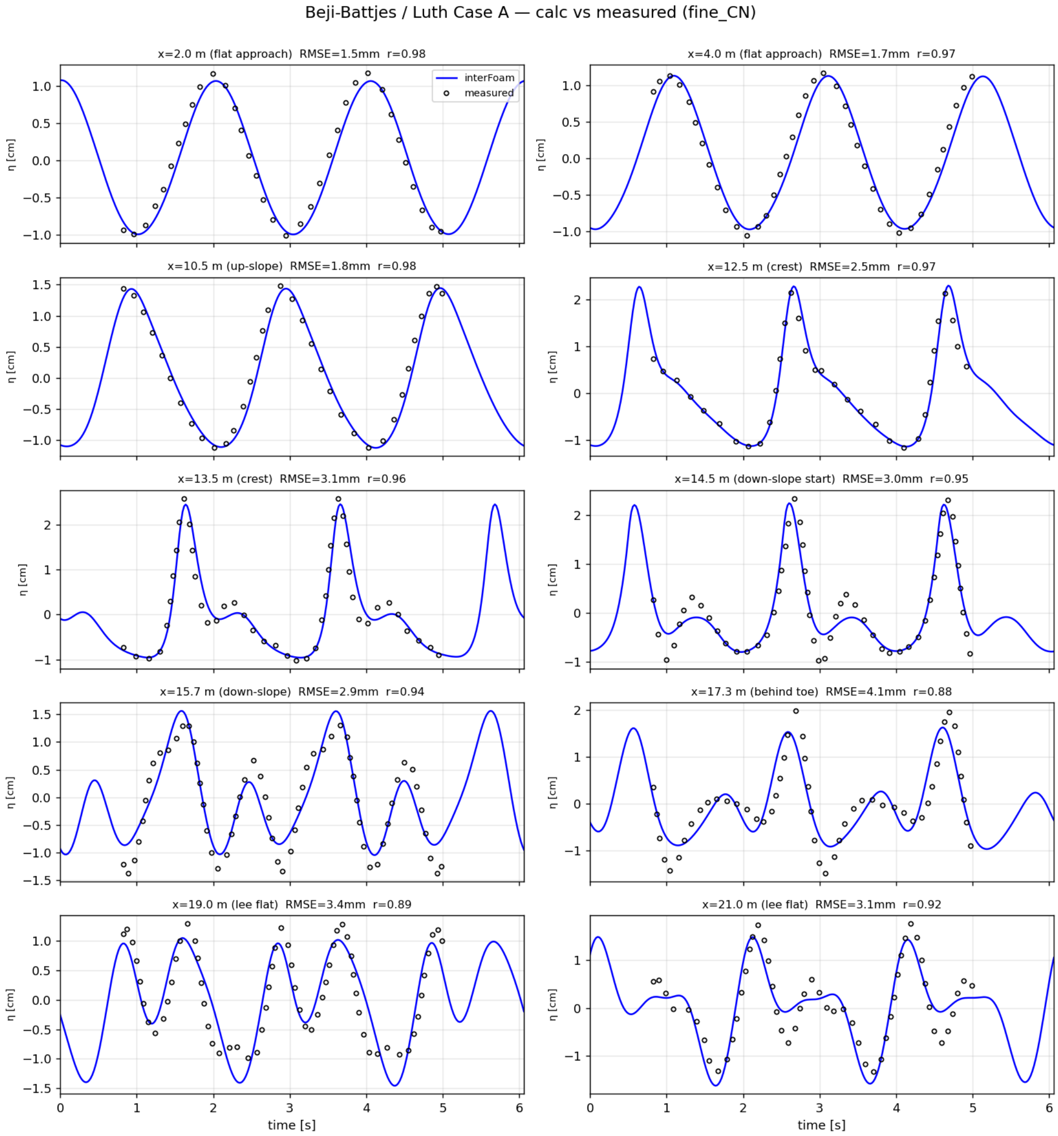

Figure 1. Beji-Battjes / Luth Case A: computed surface elevation (line) against measured series (points) at all 10 gauges, from flat approach to lee side; per-gauge RMSE and correlation annotated. Best configuration (fine mesh, Crank-Nicolson).

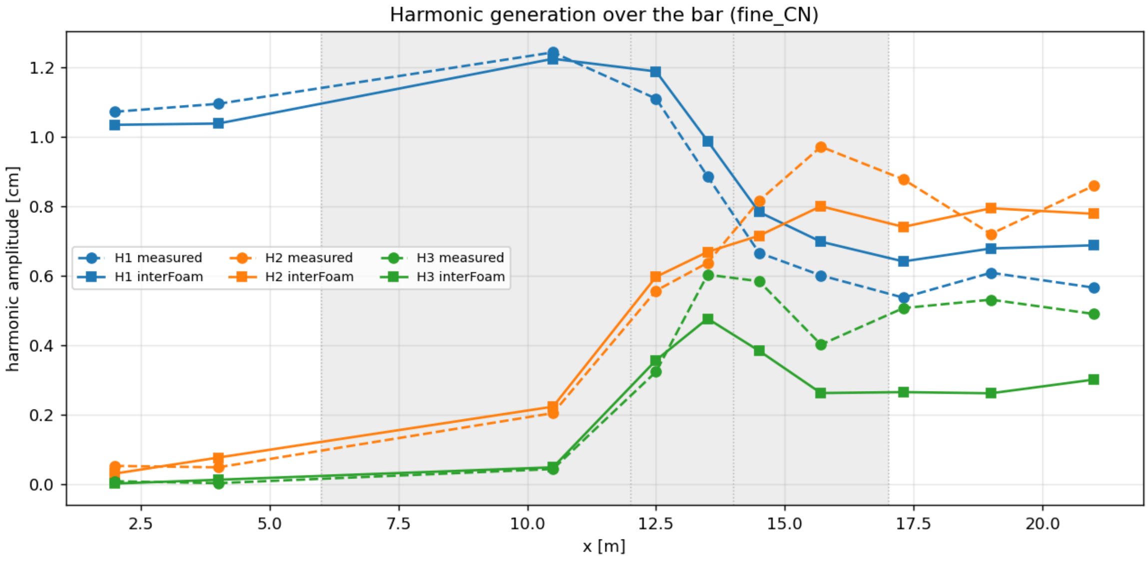

Figure 2. Harmonic amplitudes along the flume: measured (dashed) vs computed (solid) first, second and third harmonics. The shaded band marks the bar. Energy hand-off from the primary wave into higher harmonics over the bar and their release on the lee side are reproduced; the second harmonic reaches 91% of the measured amplitude behind the bar and matches on the crest.

4. Benchmark 2 - Ting & Kirby: spilling breaker on a 1:35 slope

4.1 Case, data and honesty constraints

The reference surf-zone experiment (Ting & Kirby 1994): a regular wave (T = 2.0 s, H = 0.125 m, still-water depth 0.40 m) shoals on a plane 1:35 beach and breaks as a spilling breaker at x_b = 6.4 m (breaker height 0.162 m). Measurements cover the crest/trough envelope, the mean water level, undertow profiles at four cross-sections and turbulent kinetic energy (TKE) at five. For a surf pool this is the regime that matters: it is the breaking wave.

Two provenance caveats are stated up front. First, the primary Ting & Kirby PDFs are paywalled; every measured number used here is a Ting & Kirby measurement read through open secondary sources, primarily the ICCE 2014 comparison study by Brown et al., with citations Crossref-verified. Second, the comparison points were digitised from published figures; the digitisation was checked against tabulated values where available (breaker height 0.163 digitised vs 0.162 tabulated, 0.5%).

4.2 Set-up and post-processing discipline

2D sigma-type mesh on the 1:35 beach (157k cells; bed elevation verified against the published profile to the millimetre), stream-function wave generation with independently validated Fenton coefficients, active absorption. Two arms differ only in the turbulence closure: no closure (laminar) and stock k-omega SST. A third arm (buoyancy-stabilised k-omega, the literature fix for surf-zone TKE over-production) is numerically unstable in our current build and is excluded rather than tuned into submission (Section 4.5). All statistics use a common stationary window of 10 wave periods chosen by inspection: the closure-free run develops a late-time reflected long-wave drawdown that would double its apparent undertow if the nominal longer window were used naively. Surface heights use per-wave medians, which reject detached-spray spikes that inflate naive means.

4.3 Where the model holds

Wave trough elevation tracks the measurement along the entire flume, offshore through the surf zone, in both arms. The setdown-to-setup transition of the mean water level is reproduced with the cross-over near the measured location. The undertow - the offshore-directed return current - comes out at the right magnitude (order 0.05-0.13 m/s) with the right mid-depth-maximum profile shape; here the turbulence closure clearly earns its keep, with NRMSE 0.20-0.54 of the measured range across the four sections against 0.33-1.23 for the closure-free run (Fig. 4). The onset of breaking is qualitatively right: the k-omega SST arm shoals to a sharp crest peak and collapses at x = 6.8-7.3 m against the measured 6.4 m, while the closure-free run breaks earlier and more diffusely.

4.4 Where the model diverges

TKE over-production is the headline. In the directly measured sub-trough range the stock k-omega SST closure produces 4-6 times the measured turbulent kinetic energy at 1.5-2.7 m past breaking, and its near-surface roller peak reaches k = 0.237 m2/s2 against a measured ~0.007 - a factor of ~30 (Fig. 5). The measured TKE is nearly uniform with depth; the model piles it at the surface. This is the known RANS over-production failure documented by Larsen & Fuhrman and quantified for this benchmark by Brown et al.; our contribution is reproducing it in the exact production configuration we use, so the bias is a known number rather than a literature anecdote. **Crest height is under-predicted** (measured peak +0.134 m about the local mean; k-omega SST +0.106 m, closure-free +0.088 m) - but most of this deficit traces to the incident-wave deficit described next, not to the closure: scaling the measured breaker height by the delivered incident wave predicts 0.132 m, almost exactly the k-omega SST value. **Setup magnitude** in the inner surf zone is under-predicted correspondingly (weaker modelled breaking, less momentum flux).

4.5 Open items

(C1) The wave delivered to the surf zone is 18-24% below target: the inlet is verified correct (Fenton coefficients, H = 0.125 m), and the loss occurs upstream of the measurement zone over ~2.6 wavelengths of propagation. It affects both arms equally, so the closure comparison stands, but absolute breaker heights inherit the deficit. Diagnosis (time-scheme and resolution study, as in Benchmark 1) is the next work item. (C2) The buoyancy-stabilised closure - the literature cure for the TKE over-production - blows up in our current build; it is reported as not yet usable rather than silently dropped. (C3) A mesh-sensitivity pass on the breaking case is planned; for surf-zone breaking the literature is explicit that classical grid convergence does not exist, so the honest deliverable is quantified sensitivity, not a convergence stamp. (C4) A handful of digitised setup points at the inner-most sections are digitisation artifacts (curve separation in the source figure) and are excluded.

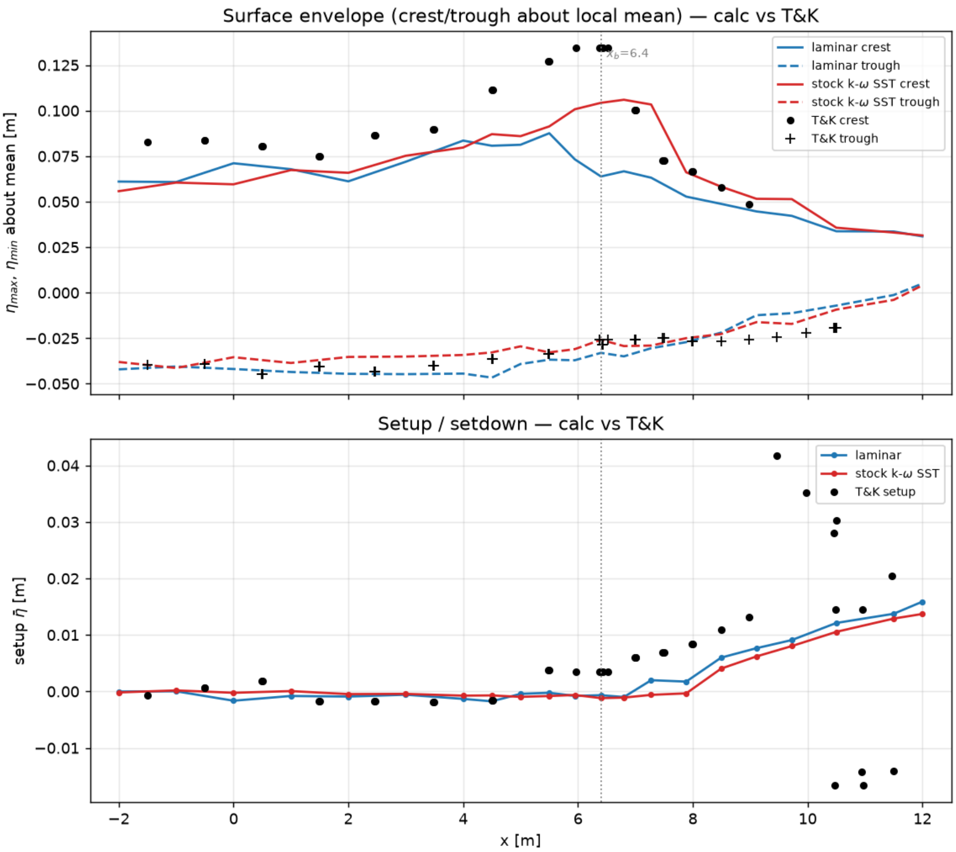

Figure 3. Ting & Kirby spilling case: crest/trough envelope about the local mean (top) and mean-water-level setdown/setup (bottom) - closure-free (blue) and stock k-omega SST (red) vs measured points. Trough and setup transition track the data; the crest peak inherits the incident-wave deficit (Section 4.5).

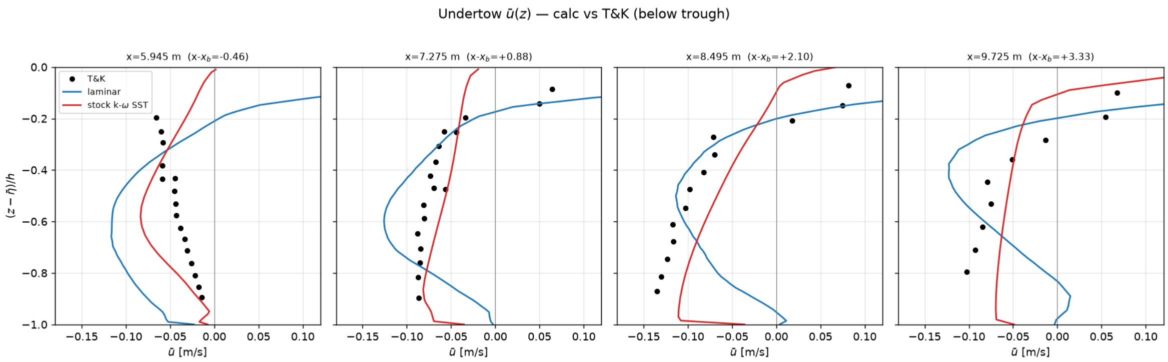

Figure 4. Undertow profiles at four cross-sections through the surf zone (depth normalised by local water column). The k-omega SST arm (red) reproduces the measured shape and magnitude better than the closure-free arm (blue): NRMSE 0.20-0.54 vs 0.33-1.23.

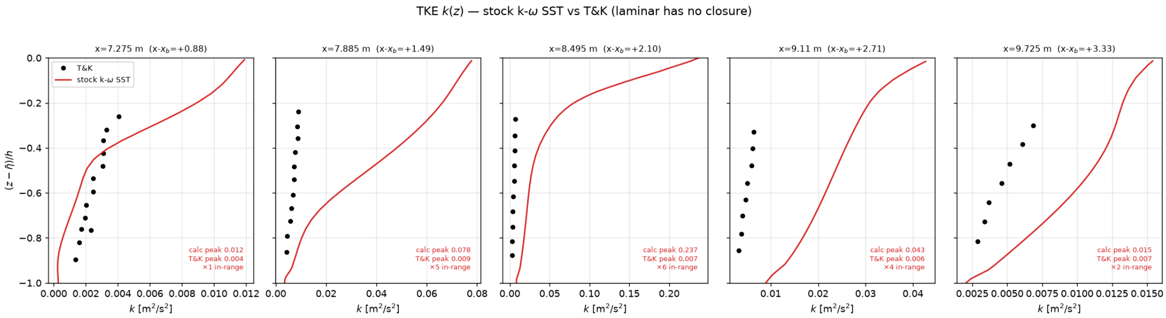

Figure 5. Turbulent kinetic energy at five sections: stock k-omega SST (red) against measured points. In-range over-prediction factors are annotated per section; the near-surface roller peak reaches ~30x the measured value. The documented RANS failure mode, shown rather than hidden.

5. What this means for surf-pool design work

The two benchmarks bracket the two halves of the design problem. Benchmark 1 validates the part of the toolchain that predicts what a bottom profile does to a wave before breaking - shoaling and harmonic transfer - and yields a concrete numerics rule (second-order time integration is co-dominant with resolution) that transfers directly to production runs. Benchmark 2 establishes where the standard closure can be trusted in the breaking zone (mean flow, undertow, trough, setup transition, breaking location) and where its output must be treated as biased by a known factor (turbulent kinetic energy near the roller). In design practice this is exactly the difference between quantities we state with confidence and quantities we bracket with margins or route to a physical scale prototype. A model with known error bars supports decisions about concrete; a gallery of agreements does not.

6. Continuation

In order of priority: stabilise the buoyancy-corrected closure and demonstrate (or refute) the literature fix for TKE over-production in this configuration; diagnose and remove the incident-wave deficit of the breaking case; add the plunging-breaker arm of the same experiment; quantify mesh sensitivity of the breaking metrics; and extend from flume sections to a fully three-dimensional wave-focusing benchmark over a shoal (Berkhoff-type), the step from 2D transects toward full surf-pool bathymetry. Updated versions of this report will accompany each closed item.

7. Reproducibility

The solver and wave boundary conditions are open source; both benchmarks use published geometry and openly available measured data (Benchmark 1 via the ERDC FUNWAVE-TVD validation testbed; Benchmark 2 via digitisation of open published figures, checked against tabulated values). Full run-by-run memos - case files, data provenance, per-gauge metrics, and the audit record - are available from the author on request.

References

Beji, S., Battjes, J.A. (1993). Experimental investigation of wave propagation over a bar. *Coastal Engineering* 19, 151-162.

Luth, H.R., Klopman, G., Kitou, N. (1994). Kinematics of waves breaking partially on an offshore bar. Delft Hydraulics, Report H1573.

Dingemans, M.W. (1994). Comparison of computations with Boussinesq-like models and laboratory measurements. Delft Hydraulics, Report H1684 (data compilation).

Shi, F. et al. FUNWAVE-TVD validation testbed, ERDC / University of Delaware (measured gauge series for the Beji-Battjes/Luth case).

Ting, F.C.K., Kirby, J.T. (1994). Observation of undertow and turbulence in a laboratory surf zone. *Coastal Engineering* 24, 51-80. DOI 10.1016/0378-3839(94)90026-4.

Ting, F.C.K., Kirby, J.T. (1996). Dynamics of surf-zone turbulence in a spilling breaker. *Coastal Engineering* 27, 131-160. DOI 10.1016/0378-3839(95)00037-2.

Brown, S.A., Greaves, D.M., Magar, V., Conley, D.C. (2014). An evaluation of RANS turbulence closure models for spilling breakers. In: *Coastal Engineering Proceedings*, 1(34), waves.5 (Proc. 34th Int. Conf. on Coastal Engineering, Seoul). [Primary open source of the digitised comparison data.]

Larsen, B.E., Fuhrman, D.R. (2018). On the over-production of turbulence beneath surface waves in Reynolds-averaged Navier-Stokes models. *Journal of Fluid Mechanics* 853, 419-460. DOI 10.1017/jfm.2018.577.