Validation of phase-resolving wave models against the Berkhoff elliptic-shoal basin benchmark

Serre-Green-Naghdi and layered non-hydrostatic solvers (Basilisk) against the 1982 laboratory sections, with a note on comparison methodology

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.21804625 (Zenodo)

Companion report (CFD tier): DOI 10.5281/zenodo.21776521 (Zenodo)

Abstract. *Surf-pool design needs wave physics at two scales: the wave-forming zone, resolved with 3D volume-of-fluid CFD, and the basin as a whole, where engineering practice uses phase-resolving depth-integrated wave models - resolving a full basin with VOF would take tens of millions of cells for a result a wave model delivers on a laptop. This report validates the wave-model tier on the benchmark created for exactly this purpose: the Berkhoff, Booy & Radder (1982) elliptic-shoal basin, in which a monochromatic wave crosses a 1:50 slope rotated 20 degrees to the wave direction with an elliptic shoal superimposed and focuses into a beam with local wave heights up to ~2.2x the incident. Two independent open-source solvers of the Basilisk framework are run as separate arms against the four openly available digitised measurement sections. The Serre-Green-Naghdi arm reproduces the normalised wave height with mean RMSE 0.093 (in H/H0 units), mean correlation r = 0.969 and a net bias of +1.6%, including the position, width and side minima of the focal beam; the layered non-hydrostatic arm reproduces the interference pattern (r = 0.942) but overshoots the focal wake by ~12% and is reported as the honest bracket. A methodological note quantifies a comparison pitfall: normalising crest maxima by a nominal incident amplitude - the easy default - biases the whole normalised field high by 10-25% on these waves; the comparison here uses the measured quantity (mean wave height) normalised by the incident wave actually delivered in each run, measured by numerical wave gauges. Model-class limits (depth-integrated physics, kh ~ 1.9 near the Green-Naghdi dispersion edge, no breaking) are stated explicitly. Together with the companion CFD-tier report this closes the two-tier validation of the toolchain.*

1. Introduction

Before the wave of a surf pool ever reaches its breaking zone, the basin as a whole has already acted on it: the full bathymetry refracts, diffracts and focuses the wave field, sets where energy concentrates and where shadow zones form. This basin-scale problem is not a job for 3D volume-of-fluid CFD - resolving a laboratory-scale basin alone would take tens of millions of cells - and engineering practice treats it with phase-resolving wave models instead, reserving CFD for the wave-forming zone where vertical flow structure matters. Our toolchain follows that standard two-tier split. The CFD tier is validated in a companion report against two wave-flume benchmarks (harmonic transfer over a submerged bar; surf-zone breaking) with quantified error bars and open caveats (Zvarich 2026, DOI 10.5281/zenodo.21776521). This report validates the second tier - the wave models - and thereby closes the basin-scale item announced in the companion's continuation section.

The benchmark is the one created for exactly this purpose. Berkhoff, Booy & Radder built the elliptic-shoal experiment to verify wave-propagation models - the 1982 paper is literally titled 'Verification of numerical wave propagation models for simple harmonic linear water waves' - and it has been the standard basin test of combined refraction-diffraction ever since. The same two principles as in the companion report were fixed in advance: **no tuning to fit** - both arms run the published configuration for this case, and nothing was adjusted to chase the measurements; and **divergence is a result** - the arm that misses is reported with the same care as the arm that lands, because quantified limits are what design margins are built from.

2. The benchmark

A monochromatic wave (T = 1.0 s) propagates across a plane 1:50 slope whose gradient direction is rotated 20 degrees against the wave direction, with an elliptic shoal (semi-axes 3 x 4 m) superimposed on the slope. Behind the shoal the 20 x 22 m measurement basin develops the classic focusing pattern: a narrow beam with local wave heights up to ~2.2x the incident, flanked by interference minima. Normalised mean wave height was measured along eight sections; the four openly available in digitised form - transverse sections at x = 3, 5 and 9 m and the longitudinal centreline, distributed with the Basilisk example of the case and traceable to Lannes & Marche (2015), Fig. 19-20 - are used here (97 measurement points in total).

Parameter	Value
Still-water depth / min. depth over shoal	0.45 m / ~0.13 m
Wave period / incident amplitude	T = 1.0 s, a0 = 2.32 cm (experiment)
Bathymetry	1:50 slope at 20 deg to the wave + elliptic shoal (semi-axes 3 x 4 m)
Compared sections	x = 3, 5, 9 m (transverse); y = 0 (longitudinal), 97 points
Regime	non-breaking refraction-diffraction; kh ~ 1.9 offshore

3. Models and numerical set-up

Two independent solvers of the open-source Basilisk framework (basilisk.fr) are run as separate arms on the same bathymetry and generation configuration: the **Serre-Green-Naghdi** solver (fully nonlinear, weakly dispersive depth-integrated equations; Popinet 2015) on a 1024 x 1024 grid with the gradient limiter off - the configuration the published example uses for this case, doubling resolution to compensate the dissipation of its second-order scheme - and the **vertically-Lagrangian layered non-hydrostatic** solver (Popinet 2020) with two layers on 512 x 512. The two arms share no discretisation machinery, so their agreement or disagreement is informative about model class, not about one code.

The set-up derives from the published Basilisk example of the case (bathymetry formulas, radiation-type wave generation at the inlet, friction-ramp absorption behind the measurement area; domain 25 x 25 m). Our additions serve measurement honesty rather than physics: nine numerical wave gauges on the constant-depth approach and three in the focus record surface elevation at 50 Hz; the surface-elevation minimum field is recorded alongside the maximum over the stationary window t = 40-50 s; and all four sections are written out directly. The full benchmark is laptop-scale: 74 minutes (Green-Naghdi) and 32 minutes (layered) wall-clock on an 11-core Apple M3 Pro, no HPC involved. Run-to-run repeatability of the section curves under thread-level nondeterminism is 0.11% at worst.

4. Comparison methodology - a note worth stating

What the experimental numbers are matters as much as what the model computes. The digitised sections are the **normalised mean wave height**: at each point the gauge record was analysed by zero up-crossing into individual waves and the mean wave height taken, normalised so the incident wave is unity (Lannes &

Marche 2015, Sec. 4.7 and Fig. 20). The model quantity is therefore built the same way: H = eta_max - eta_min per point over the stationary window, halved and divided by H0/2 **as actually delivered in the run** - the median per-period (crest - trough)/2 across the nine offshore gauges: 22.3 mm in the Green-Naghdi run, 21.2 mm in the layered run, with gauge-to-gauge spread under 2% (so shoal reflection is negligible at the gauges) and the harmonic-fit amplitude agreeing within 0.5 mm. The normalisation leans on nothing nominal: radiation-type generators do not deliver their setting exactly, and the delivered amplitude differs between solvers on the same setting.

The pitfall this avoids is worth quantifying, because it is the easy default: comparing the crest maximum |eta|_max (a field any solver records cheaply) normalised by a nominal incident amplitude. On these waves that combination biases the entire normalised field high by 10-25% - worst in the focus, where crest-trough asymmetry peaks - which is large against the ~10% discrimination the benchmark offers between model classes. Measured on our own runs: the crest-max comparison degrades the Green-Naghdi mean bias from +0.016 to +0.116 and the layered from +0.113 to +0.248. Two smaller effects were also measured rather than assumed: taking the window maximum instead of the per-wave mean contributes only +1.7-2.0% (checked at gauges placed in the focus - the wave train there is stationary, not modulated), and digitisation noise is visible but unquantified, as for any digitised benchmark.

5. Results

Section	GN RMSE	GN r	GN bias	ML RMSE	ML r	ML bias
2 (x = 3 m)	0.097	0.932	+0.015	0.167	0.897	+0.065
3 (x = 5 m)	0.097	0.985	+0.018	0.134	0.984	+0.047
5 (x = 9 m)	0.093	0.987	+0.026	0.203	0.954	+0.129
7 (centreline)	0.084	0.971	+0.006	0.264	0.932	+0.210
mean	0.093	0.969	+0.016	0.192	0.942	+0.113

The Green-Naghdi arm lands on the measurements (RMSE and bias in units of H/H0): mean RMSE 0.093 - i.e. 5.5-11.6% of the measured range per section - with mean correlation r = 0.969 and a net bias of +1.6% across all 97 points. The position, width and both side minima of the focal beam are reproduced (Fig. 1); the longitudinal profile through the focus (section 7) carries essentially zero net bias (+0.006). The arm under-predicts the two sharpest transverse peaks by 7-9% (sections 3 and 5) while holding near-zero net bias - consistent with the residual dissipation of its second-order scheme, the same property the published example compensates with doubled grid resolution.

The layered arm is the honest bracket. With two layers at 512 x 512 it reproduces the interference pattern (r = 0.942) but overshoots the focal wake by ~12% on average (sections 5 and 7) and the sharpest transverse peak by ~14% (Fig. 1, blue). It is reported alongside rather than averaged into the headline number: the point of running two independent arms is to see the model-class spread, and here it brackets the measurements from above while Green-Naghdi touches them from below at the peaks.

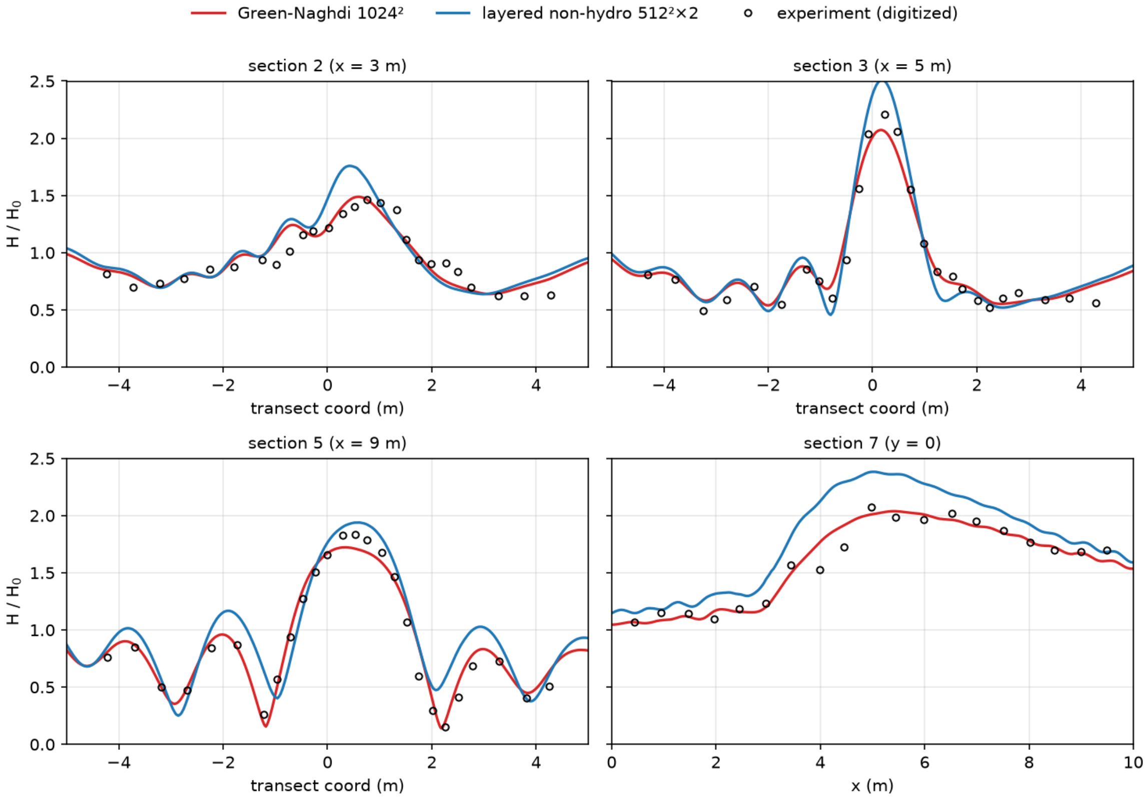

Figure 1. Normalised wave height along the four openly available measured sections - Green-Naghdi (red) and layered non-hydrostatic (blue) arms against the digitised laboratory points. Green-Naghdi: mean RMSE 0.093, r = 0.969, net bias +1.6%.

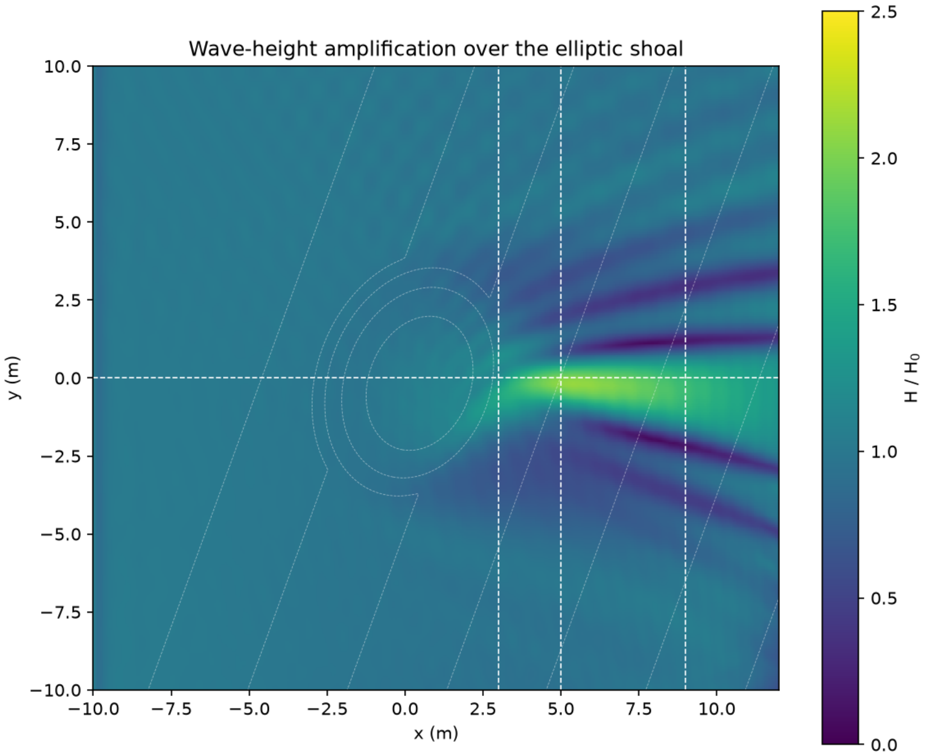

Figure 2. Computed wave-height amplification across the basin (Green-Naghdi arm): the shoal focuses the incident wave into a beam of up to ~2.3x H0 with interference minima on both flanks. Dashed lines mark the measured sections; thin contours the bathymetry.

6. Limits

These are depth-integrated models: vertical flow structure is not resolved - that is what the CFD tier of the toolchain is for. At kh ~ 1.9 offshore the Green-Naghdi class operates near its dispersion limit, and the +1.6% net bias should be read with that context; the layered solver has more dispersive headroom but, at the two-layer / 512 x 512 configuration published for this case, delivers the weaker quantitative agreement here. Sections 1, 4, 6 and 8 of the experiment are not openly available in digitised form and are not compared. No wave breaking occurs in this case, and nothing here speaks to breaking-zone validity - that regime is validated on the CFD tier in the companion report. Grid-refinement sensitivity of the arms was not varied beyond the published configurations.

7. Reproducibility

Both solvers are open source (basilisk.fr); the case configuration derives from the published Basilisk example of the benchmark, and the digitised measurement sections are publicly distributed with that example. The additions (wave gauges, minimum-field recording, section output) are minor and fully described above. Full case files, run logs, per-section metrics and the analysis script are available from the author on request.

References

Berkhoff, J.C.W., Booy, N., Radder, A.C. (1982). Verification of numerical wave propagation models for simple harmonic linear water waves. *Coastal Engineering* 6, 255-279.

Lannes, D., Marche, F. (2015). A new class of fully nonlinear and weakly dispersive Green-Naghdi models for efficient 2D simulations. *Journal of Computational Physics* 282, 238-268. [Source of the digitised sections and of the measured-quantity definition.]

Popinet, S. (2015). A quadtree-adaptive multigrid solver for the Serre-Green-Naghdi equations. *Journal of Computational Physics* 302, 336-358.

Popinet, S. (2020). A vertically-Lagrangian, non-hydrostatic, multilayer model for multiscale free-surface flows. *Journal of Computational Physics* 418, 109609.

Basilisk framework: example 'shoal.c' - periodic wave propagation over an ellipsoidal shoal. basilisk.fr/src/examples/shoal.c (accessed August 2026). [Case configuration and section-data distribution.]

Zvarich, S. (2026). Validation of an OpenFOAM/interFoam wave-modelling toolchain against two published wave-flume benchmarks. Technical report v1.0, Wavemorison. DOI 10.5281/zenodo.21776521. [Companion report: the CFD tier.]