Marine Structures 106 (2026) 103966

Contents lists available at ScienceDirect

Marine Structures

journal homepage: www.elsevier.com/locate/marstruc

Research Paper

Evaluating numerical simulation accuracy for full-scale high-strength steel ship structures: Insights from the ISSC 2025 Ultimate Strength Committee benchmark on transversely stiffened panels

Marco Gaiotti, Ph.D. [a,*], Lars Brubak [b], Bai-Qiao Chen [c], Ionel Darie [d], Dimitris Georgiadis [e], Daisuke Shiomitsu [f], Mihkel Kõrgesaar [g], Yining Lv [h], Ken Nahshon [i], Marcelo Paredes [j], Jani Romanoff [k], Ingrid Schipperen [l], Akira Tatsumi [m], Murilo Vaz [n], Yikun Wang [o], Albert Zamarin [p], Zhihu Zhan [q], Jonas W. Ringsberg [r]

[a] University of Genova, Polytechnic School, DITEN, via Montallegro 1, Genova, 16145, Italy
[b] DNV Høvik, Norway
[c] Centre for Marine Technology and Ocean Engineering, Lisbon, Portugal
[d] DNV Hamburg, Germany
[e] Technical University of Munich, Munich, Germany
[f] Nippon Kaiji Kyokai (ClassNK), Tokyo, Japan
[g] Tallinn University of Technology, Kuressaare, Estonia
[h] China Classification Society, Beijing, China
[i] Naval Surface Warfare Center, Carderock Division, West Bethesda, USA
[j] Texas A&M University, Galveston, USA
[k] School of Engineering, Department of Energy and Mechanical Engineering, Marine and Arctic Technology, Aalto University, Espoo, 00076, Finland
[l] TNO Delft, Netherlands
[m] Osaka University, Osaka, Japan
[n] Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
[o] Lloyd's Register, Southampton, UK
[p] University of Rijeka, Faculty of Engineering, Rijeka, Croatia
[q] China Classification Society, Shanghai, China
[r] Chalmers University of Technology, Department of Mechanics and Maritime Sciences, Division of Marine Technology, SE-412 96 Gothenburg, Sweden

ARTICLE INFO

Keywords:
Buckling criteria
Experimental validation
Finite element analysis benchmarking
Regulatory compliance
Ultimate strength assessment

ABSTRACT

The demand for sustainable ship design has driven the use of high-strength steel to reduce structural weight, although this introduces buckling challenges due to unchanged elastic properties. Supported by the ISSC 2025 Ultimate Strength Committee, this study evaluated the ability of numerical simulations to predict the nonlinear response and ultimate strength of stiffened panels subjected to transverse compression. The benchmark consisted of full-scale blinded experimental tests that were conducted in parallel using a deck-like structure with thin plating prone to elastic buckling. The finite element models produced by participating researchers were compared, focusing on the complete end-shortening curve rather than just ultimate strength.

* Corresponding author at: University of Genova, Polytechnic School, DITEN, via Montallegro 1, Genova, 16145.
E-mail address: marco.gaiotti@unige.it (M. Gaiotti).

https://doi.org/10.1016/j.marstruc.2025.103966
Received 9 June 2025; Received in revised form 6 November 2025; Accepted 6 November 2025
Available online 12 November 2025

Despite identical input geometry and minimal modeling guidance, results varied widely, revealing the significant influence of user-defined assumptions. The inclusion of additional data on material properties in the second study phase led to greater result dispersion due to the different strategies adopted for the hardening model. Key variability sources included the modeling of initial imperfections, material constitutive laws, and residual stresses from welding. The study highlights the need for consistent modeling and improved experimental data collection, particularly regarding boundary conditions and residual stress effects. While including welding stresses improved stiffness predictions, uncertainty in boundary behavior limited the assessment of ultimate strength impacts. The study also evaluated compliance with classification society rules (e.g., CSR, DNV, UR-S35), offering insights into how nonlinear numerical analyses complement or challenge regulatory frameworks based on closed-form expressions. Recommendations are made for improving simulation reliability and result validation.

1. Introduction

The increasing demand for sustainability in ship design is driving the pursuit of structural weight reduction. Concurrently, climate change is expected to lead to more extreme environmental conditions, necessitating greater attention to the ultimate strength of hull girders under high design loads. This dual challenge highlights the critical need to accurately assess both as-designed and as-built structural performance. Achieving this requires a combination of experimental investigations and realistic numerical simulations, with a strong emphasis on ensuring consistency and agreement between the two approaches.

The shift in the maritime sector toward using high-strength steels helps reduce hull weight but also brings challenges in designing thin structures that are vulnerable to elastic buckling. This happens because, even though the yield stress increases, the Young's modulus broadly stays the same, which complicates the design process [21]. The plate slenderness parameter β can be defined as follows [19]:

$$\beta = \frac{b}{t}\sqrt{\frac{\sigma_y}{E}}, \tag{1}$$

where b is the length of the loaded edge, t is the plate thickness, σ_y is the yield strength, and E is the material's Young's modulus. While transverse loads are not typically severe in ships with longitudinal frames, the low resistance to elastic buckling due to the aspect ratio of the plating panels remains an issue, especially in the presence of hull openings or transverse framing that alter the stress field. International discussions are currently centered on establishing a regulatory framework for verifying buckling with transverse stress components [6].

Since Smith [19], numerous studies have been conducted to determine the ultimate strength of axially loaded ship structures at various levels of complexity, including grillages, stiffened panels, transverse ring frames, and entire hull blocks. The effects of cracks, corrosion and temperature on the ultimate strength have also been investigated. (e.g., [5,8,14,16–18]). Such experimental investigations are necessary to compare as-designed and as-produced structural properties, as the multi-stage production process introduces imperfections that may differ substantially from the idealized imperfections used to design the structure. Manco et al. [12,13] investigated the effects of corrosion and repair on scaled ship panels (1:3.5 scale) typical of floating production storage and offloading vessels. Their methodology involved detailed measurements of initial imperfections and material properties to validate finite element (FE) models that describe how these imperfections redistribute loads and affect failure mechanisms. Such investigations aim to assess the progressive collapse of the hull girder analytically and numerically, focusing on the sequence of plates, stiffeners, and web frames. The challenge with scaled experiments is that relevant physical mechanisms do not scale in the same way. This is why full-scale validations are needed. The [9] Ultimate Strength Committee's benchmark study [9] focused on longitudinally stiffened ship structures in which the stiffeners are positioned in the direction of hull girder bending-induced stresses. The study underscored the need to utilize numerical methods to analyze the full end-shortening curve rather than only the ultimate strength at the peak of the load-end-shortening curve. This approach is especially important in situations where energy dissipation within the structure is a key factor, such as in the failure of a hull girder subjected to wave-induced loading [11].

Limited research has focused on the effects of transverse loads on relatively thin shell plating, which is particularly susceptible to buckling even at low loads. As part of a research project funded by Fincantieri's R&D Department, a full-scale experimental test was carried out on a structure representing a deck portion with thin plating, designed to induce early local elastic buckling and a wide post-buckling range before reaching the ultimate load [1]. The authors carried out numerical simulations to evaluate the influence of initial plating imperfections and welding residual stresses, assessing their impact on both the end-shortening curve and the ultimate strength of the panel.

This paper presents an extensive benchmark study organized by the [10] Ultimate Strength Committee that focuses on the same experimental test reported by Barsotti et al. [1]; see ISSC [10] for more details. The test was designed as a blind benchmark study to assess the participating researchers' ability to predict the ultimate strength and nonlinear response of the panel structure tested up to its collapse. Since the experimental test was conducted in parallel with the numerical benchmark study, no observations or results from the experiment were available for comparison with the numerical results. Barsotti et al. [1] concentrated on describing the underlying motivations and objectives that led to the implementation of the large-scale experimental test, as well as the experimental results that

stemmed from it. By contrast, the present study emphasizes the numerical simulation results and compares the observations and results from the experiment and the numerical simulations.

1.1. CSR and UR-S35

One of the key motivations for the experimental test was to demonstrate that a stiffened plate, loaded in its transverse direction, retains strength beyond its elastic buckling limit, as is the case in a longitudinally loaded panel. Current buckling formulae in CSR [7] and UR-S35 [6] acknowledge that local elastic plate buckling can occur while still allowing for reserve strength. However, global elastic buckling is not permitted, as it leads to stress redistribution over a large area and represents an inherently unstable failure mode.

In CSR strength evaluations, for relatively slender panels, the yield strength of the material does not play a role, as global elastic buckling is primarily a function of elastic properties rather than yielding behavior. This aspect was examined in several stages of the benchmark study. Certain regulatory standards, such as the DNV Ship Rules [3], accept the use of nonlinear FE analysis for strength assessment but mandate an independent eigenvalue analysis to verify the global elastic buckling of stiffeners. Ultimately, the structure's capacity is determined by selecting the lower value between the global buckling load from eigenvalue analysis and the load capacity obtained from nonlinear FE simulations. In reality, the lowest eigenvalue may not represent the shape of the production-induced distortions, resulting in a difference in strength that warrants further research.

2. Experimental setup

A comprehensive description of the experiment is presented by Barsotti et al. [1]. The panel's geometry was defined within an extensive planning phase (Fig. 1a), which considered various specimen shapes and possible geometric imperfection patterns. The objective was to induce early local elastic plate buckling, followed by a significant post-buckling phase before reaching the ultimate load, while also accounting for the laboratory's limited capacity for full-scale testing. In particular, the geometry was designed to analyze the loss of stiffness in the structural component following the onset of local elastic instability and to assess its effects on load redistribution to the surrounding structural members of the plate. Additionally, the study aimed to observe the structural component's behavior during the transition from local elastic instability to collapse, as well as in the post-collapse phase.

The panel was designed to be representative of a typical ship deck structure. To prevent collapse at the panel's ends and ensure a clear experimental investigation area, different plating thicknesses were used. It is important to note that the primary focus of the study was on pre- and post-buckling behavior, particularly on stress redistribution within the structure, rather than solely on its ultimate strength. The thicknesses of the shell plating and beams were also measured, and 3D laser scans allowed for reconstructing the geometry of the panel, including the geometric imperfections introduced by the manufacturing process (Fig. 1b).

In bay 2, the geometric imperfection measured on the shell plating coincides with the shape of the first and lowest bifurcation buckling mode (Fig. 1c). This suggests the stresses introduced by welding have caused the elastic buckling of the specimen and points to the need to investigate the effect of such residual stresses on the specimen's structural response. Bays 1 and 3 also exhibit non-negligible initial imperfections; note that the shell plating in those areas was thicker than in the central bay to force the collapse to take place in bay 2.

Fig. 2 shows the tested panel structure—hereafter the specimen—positioned on the test bench (Fig. 2a), the edge opposite from the actuators (Fig. 2b), and the section where the collapse occurred (Fig. 2c), which indicates the collapse resulted from the onset of global instability in the specimen. The experimental data collected during the test indicate that the loaded side maintained a good degree of

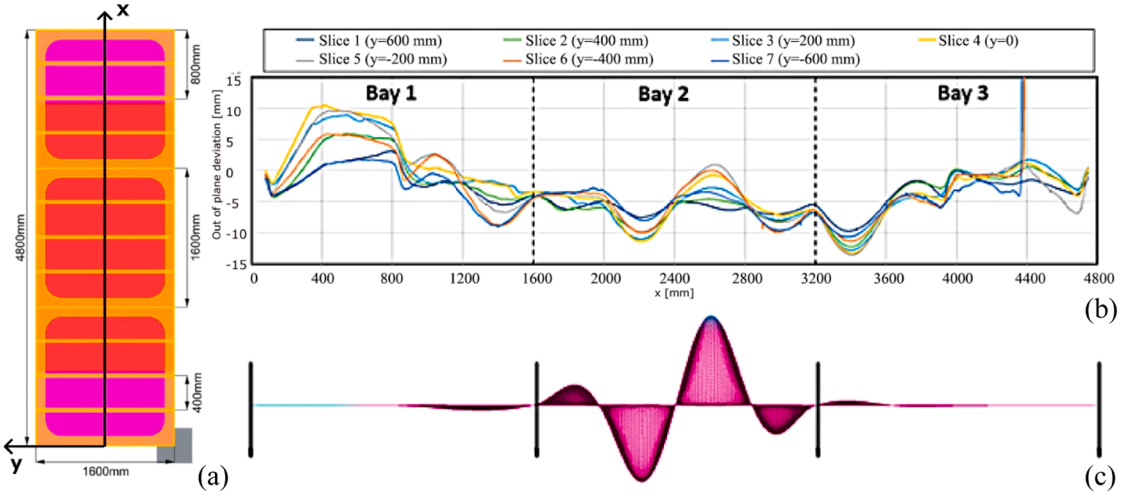


Fig. 1. **(a)** Top view of the panel's geometry. **(b)** Laser-scanned shell plating imperfections.
(c) First buckling mode shape at the panel's longitudinal symmetry axis.

Fig. 2. **(a)** Top view of the specimen on the laboratory test bench. **(b)** System for connecting the specimen to the bench on the unloaded side through the load cells. **(c)** Collapsed specimen section.

constraint with respect to the rotation around the *y*-axis of the specimen (Fig. 10). The opposite side, where the constraint included the connection to the load cells, proved to be less strong, partly due to the localized failure of some supports. All experimental data are available upon request from the University of Genova.

3. Description of the benchmark study

3.1. Overview

Bringing together experienced professionals facilitated the consideration of diverse perspectives and allowed for comprehensive statistical analysis. This included the evaluation of systematic uncertainties related to boundary conditions and the stress–strain curves and uncertainties introduced by participants selecting modeling parameters. The study compared different regulatory guidelines, participant expertise, numerical methods, and simulation methods to determine their effectiveness in predicting the full end-shortening curve of stiffened panel structures under transverse compression. The subsequent discussion focused on the influence of modeling uncertainties, material properties, initial imperfections, residual stresses, and the analysts' assumptions on the results to produce clearer, more rigorous guidelines.

Like the ISSC [9] study, this study was divided into three phases. The study outlines how prediction accuracy evolves with increasing data availability, which is relevant to researchers working on complex modeling and numerical simulations. The study offers valuable insights into the challenges of making reliable predictions with limited information and the potential benefits of incorporating additional data and refining key parameters. By comparing the three phases of the study, this paper provides a framework for evaluating the impact of data availability on model performance, which is relevant to various scientific and engineering disciplines. The final phase goes beyond standard validation approaches by encouraging researchers to critically assess discrepancies between numerical and experimental results, fostering a deeper understanding of modeling limitations and potential improvements.

The study followed a blind approach: in the first two phases, participants did not have access to the experimental results. The primary objective was to assess the researchers' ability to predict the ultimate strength and nonlinear behavior of a structural component up to its collapse with limited information available. By structuring the study into three phases, the impact of additional data on prediction accuracy could be systematically evaluated.

3.2. Study participants

The study was led and coordinated by the [10] Ultimate Strength Committee III.1. Researchers outside the committee were also invited to participate to make the benchmark study as successful as possible. The results were collected by the University of Genova (Italy), where the experimental test was carried out. Numerical results were provided by 16 participants from 15 research groups:

- CCS, China Classification Society, with two separate teams, China
- CENTEC, Centre for Marine Technology and Ocean Engineering, Portugal
- Chalmers University of Technology, Sweden
- ClassNK, Nippon Kaiji Kyokai, Japan
- DNV, with two separate teams, Germany and Norway
- LR, Lloyd's Register, UK
- NAOE, University of Osaka, Japan
- NSWC, Naval Surface Warfare Center Carderock Division, United States
- RITEH, University of Rijeka, Faculty of Engineering, Croatia
- TALTECH, Tallinn University of Technology, Estonia
- TAMU, Texas A&M University, United States
- TNO, Netherlands Organisation for Applied Scientific Research, Netherlands
- TUM, Technical University of Munich, Germany
- UFRJ, Universidade Federal do Rio de Janeiro, Brazil.

3.3. Benchmark phases

The three phases are outlined in Fig. 3. Building on insights from ISSC [9], this study used nominal mechanical properties and the specimen's actual geometry, including geometric imperfections and laser scan data, starting from Phase 1. In Phase 2, the actual mechanical properties were provided to the participants upon submission of the Phase 1 results. Phase 3 was different compared to the third phase in the ISSC [9] study. This is because the experimental test results from the present study indicated that the boundary conditions at the non-loaded end did not remain as originally defined, affecting the entire experiment. A workshop was arranged to inform the participants about the observations made during the actual experiment. After the workshop, the participants were invited to propose modifications to their Phase 2 models to capture the specimen's behavior during the experiment. Hence, in Phase 3, each participant was tasked with fitting the experimental curve using their skills.

The following information was provided to the participants in **Phase 1**:

- nominal CAD model of the specimen
- boundary conditions, load case, and speed

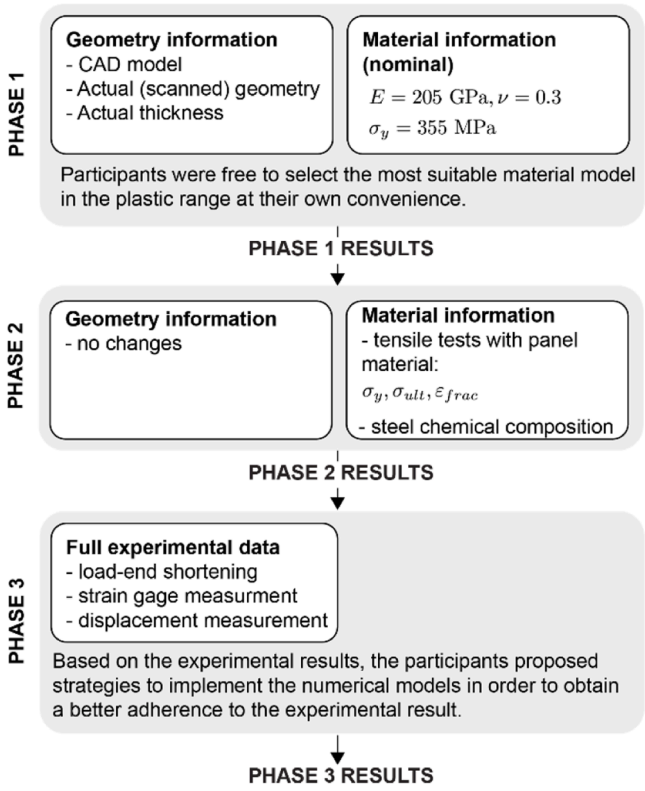


Fig. 3. Benchmark structure and phases.

- actual specimen geometry from laser scanning (geometric imperfections)
- actual specimen thickness from ultrasonic testing (UT) scanning
- nominal material properties, including the Young's modulus ($E = 205$ GPa), Poisson's ratio ($\upsilon = 0.3$), and yield strength ($\sigma_y = 355$ MPa).

The participants were allowed to select any strategy to model the measured geometric imperfections. They could also select the most suitable constitutive material model in the plastic range. In **Phase 2**, the participants were provided with data from experimental material tests carried out on small specimens extracted from the same sheets used for the construction of the test specimen. These data included (i) the yield strength, (ii) the engineering strength, (iii) the elongation at break, and (iv) the detailed steel chemical composition.

In **Phase 3**, the experimental data were fully revealed, including (i) the end-shortening curve, (ii) strain gauge measurements, and (iii) all measured axial displacements, which are useful for evaluating boundary conditions.

The participants were encouraged to recommend strategies to achieve a better fit with the experimental results, not limiting themselves to including parameters omitted in the previous phases. They were also encouraged to formulate hypotheses for the unforeseen problems that arose during the experimental test, such as variations in the boundary conditions of the specimen.

4. Results

Preliminary findings from Phase 1 were presented by Gaiotti et al. [4]. Therefore, this phase is discussed briefly in Section 4.1, while Phases 2 and 3 are covered in more detail in Sections 4.2 and 4.3, respectively.

4.1. Phase 1: Actual geometry and nominal material data

Gaiotti et al.'s [4] detailed analysis of Phase 1 results revealed that imperfection modeling plays a crucial role in the early stages of the load history. This is likely due to the test being designed to induce local elastic buckling at a low critical load compared to the specimen's expected ultimate strength. Most participants selected a strain hardening modulus of <0.5 % of the Young's modulus—a relatively low value. Its influence appears to extend beyond the ultimate strength; it also affects the post-collapse behavior when compared to a perfectly plastic material model.

Moreover, having provided the complete 3D scan of the model, it was assumed that all participants would use the real geometry for geometric modelling. Indeed, some participants opted for different strategies, ignoring the real imperfection and choosing alternatives such as eigenmode or sinusoidal, which led to scattered results, especially in the early stage of the end-shortening curve.

Fig. 4 summarizes the results from Phase 1. The participants conducted an in-depth analysis of the possible primary sources contributing to the scattered results from the numerical simulations. However, they could not identify trends describing how the choice of specific modeling parameters impacts the ultimate strength of the structure. As shown in Fig. 4a, establishing clear relationships between modeling strategies and the structural response is not straightforward. Moreover, the correlation between the combination and selection of modeling parameters seems to be affected by randomness, which affects the end-shortening curves (Fig. 4b).

4.2. Phase 2: Actual geometry and material data

Participants were instructed to update their numerical models from Phase 1 by incorporating the provided actual material properties, keeping all other parameters unchanged. The additional data included tensile test results from 16 specimens taken from the same steel sheets used to manufacture the test panel. Table 1 presents the shared data, which included various specimen thicknesses (4–12 mm) and material properties such as the yield strength and elongation. Ultimately, 15 participants completed Phase 2.

During Phase 1, all participants used the same yield strength with minimal changes to the work-hardening model. By contrast, Phase 2 introduced more detailed material data, leading to diverse interpretations of the material's plastic behavior. Some participants modified their models to reflect variations in sheet thickness. Notably, the 4 mm plating model resulted in a wide range of approaches, as illustrated in Fig. 5a. Fig. 5b displays the standard deviation of the end-shortening curves in Phases 1 and 2. As anticipated, the results show noticeable deviation only at high compressive loads, where the material undergoes plastic deformation.

The Phase 1 results suggested a certain dependence in the post-ultimate load response on the selected hardening model. However, analyzing the results of three participants who adopted similar constitutive material models (Fig. 6a) produced significantly different responses when plotting the normalized end-shortening curve with respect to the participant-estimated ultimate loads and relative displacements (Fig. 6b). One explanation for this result is that the ultimate load also depends on other modeling parameters.

Providing material data (Table 1) improved numerical accuracy but also increased result dispersion due to participants' differing material model interpretations. At the end of Phase 2, participants received experimental data including the measured ultimate compressive force. In Phase 1, the average ultimate strength was underestimated compared to the experimental value by 7.47 % (1295.7 kN, St.Dv. 80.3 kN); in Phase 2, it was overestimated by 1.98 % (1434.6 kN, St.Dv. 118.2 kN). A key insight from Phase 2 is that additional material data improved the accuracy of numerical predictions. However, differences in interpreting the constitutive material model increased result dispersion, highlighting the impact of modeling assumptions on consistency (Fig. 7).

(a)

Team	SW	Mesh size [mm]	nodes	Imperfection model: Shape	Amplitude [mm]	Ultimate strength: Load [kN]	Ultimate strength: Disp. [mm]	Deviation: Load [%]	Deviation: Disp. [%]	Solver	Hardening	Tangential modulus [Mpa]
#1	NX-Nastran	50x50	4	Eigen mode 1	2.0	1348.3	6.00	4.1	3.9	Implicit	yes	1000
#2	LS-Dyna	20x20	4	Actual	6.0	1366.3	5.60	5.4	-3.0	Implicit	yes	125
#3	Abaqus	10x10	4	Actual	6.0	1188.6	6.00	-8.3	3.9	Explicit	yes	1000
#4	LS-Dyna	20x20	4	Eigen mode 1	1.5	1398.2	5.30	7.9	-8.2	Implicit	yes	1000
#5	Abaqus	20x20	4	Sinusoidal	6.0	1256.7	5.80	-3.0	0.5	Implicit	no	-
#8	Abaqus	20x20	4	Actual	6.0	1351.8	6.10	4.3	5.7	Implicit	yes	1000
#11	Abaqus	20x20	4	Actual	6.0	1296.6	6.20	0.1	7.4	Implicit	yes (NL)	≈815
#12	Abaqus	19x19	4	Actual	6.0	1258.0	5.70	-2.9	-1.3	Implicit	yes	1000
#13	Abaqus	12.5 to 50	4	Only tilt	-	1362.7	5.70	5.2	-1.3	Implicit	no	-
#15	LS-Dyna	20x20	4	Not used	-	1393.8	5.40	7.6	-6.5	Explicit	no	-
#16	Abaqus	25x25	8	Actual	6.0	1276.7	5.60	-1.5	-3.0	Implicit	yes (NL)	≈815
#19	Abaqus	15x15	4	Eigen mode 1 to 5	0.8	1382.6	5.70	6.7	-1.3	Implicit	yes	975
#20	LS-Dyna	20x20	4	Actual	6.0	1195.6	6.10	-7.7	5.7	Explicit	yes	1000
#21	Marc	20x20	4	Actual with tilt	6.0	1212.2	5.90	-6.4	2.2	Implicit	no	-
#22	Abaqus	20x20	4 (3)	Eigen mode 1	2.0	1147.9	5.50	-11.4	-4.7	Implicit	yes	N/A
					AVG	**1295.73**	5.77					
					St.Dev.	**80.3**	0.26					

(b)

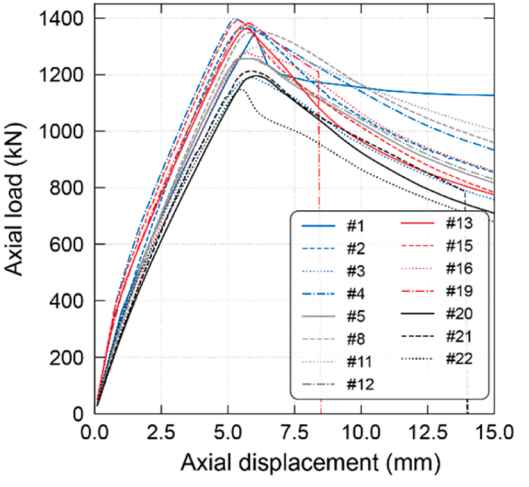


Fig. 4. **(a)** Salient features of the modeling strategies of participants. **(b)** Resulting end-shortening curves from Phase 1.

Table 1
Material properties disclosed to participants in Phase 2. The engineering strength is the ultimate strength of the material with engineering assumptions.

Specimen no.	Thickness [mm]	Yield Strength [MPa]	Engineering Strength [MPa]	Elongation [%]
1	4	417	505	38.2
2	4	394	495	40.2
3	4	416	499	34.1
4	4	382	502	34.7
5	4	385	493	33.5
6	4	412	495	36.9
Average		401	498	36.3
Standard deviation		15	4	2.4
7	5	398	494	39.4
8	6	387	497	38.3
9	8	399	529	31.8
10	10	395	546	28.2
11	10	406	532	34.1
12	10	397	529	35.5
13	10	396	526	30.1
14	10	389	533	28.7
Average		397	533	31.3
Standard deviation		6	7	2.9
15	12	425	554	32.8
16	12	392	550	29.8

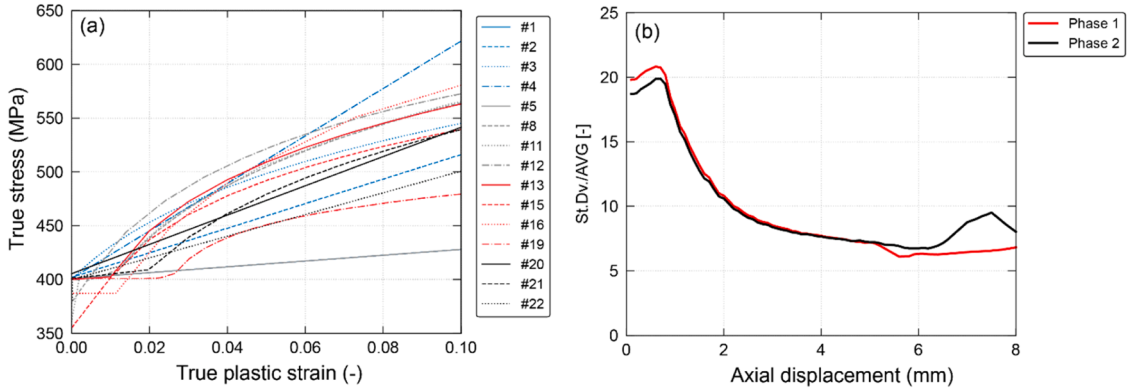


Fig. 5. **(a)** True plastic stress vs. true plastic strain as modeled by participants in Phase 2.
(b) Standard deviation of the end-shortening curve over the average values observed in Phases 1 and 2.

4.3. Phase 3: Simulation of the experiment

In Phase 3, participants analyzed numerical–experimental discrepancies after reviewing the experimental data. Building on the experience gained during the previous benchmark study [9], a discrepancy between the numerical results and the test was expected, especially regarding the specimen's stiffness. Thus, Phase 3 was established to investigate possible reasons for these differences. Discrepancies were confirmed after the disclosure of experimental results. A key observation from discussions was that the boundary conditions had changed during testing, as evidenced by four distinct changes in the experimental curve. These changes comprised an initial settling phase and three sudden load drops due to support failures, which had not been accounted for in Phases 1 and 2. To further investigate these factors, participants were divided into specialized groups. One group analyzed the influence of including welding residual stresses in the FE model (Fig. 8a), while another group focused on refining the constraint conditions to better match the experimental setup without introducing residual stresses in the FE model (Fig. 8b). This approach was based on the group's confidence in the reliability of the initial experimental data and the accuracy of their simulations.

A separate approach involved modeling the specimen using three layers of solid elements linked to shell elements via constraint equations and incorporating real panel imperfections. However, this method resulted in less than a 0.5 % increase in the ultimate load in Phase 2 and had a negligible impact on post-ultimate behavior. Given the complexity and high computational cost of solid element modeling, shell or plate elements remain the preferred choice for accurately capturing structural responses in thin plating analysis.

4.3.1. Effect of residual stresses

Three participants incorporated welding-induced residual stresses into their FE models using different approaches. Participant 8 applied a simplified thermal model based on CEN [2] targeting residual stress levels of 340 MPa and 400 MPa. Participants 5 and 11

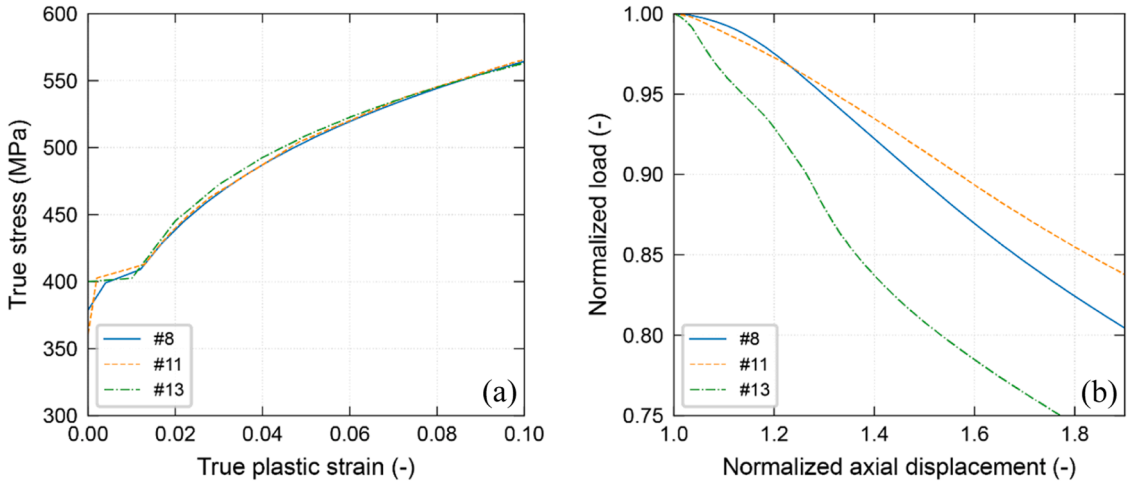


Fig. 6. **(a)** Normalized end-shortening curves produced by Participants 8, 11, and 13; **(b)** corresponding true stress–true plastic strain models.

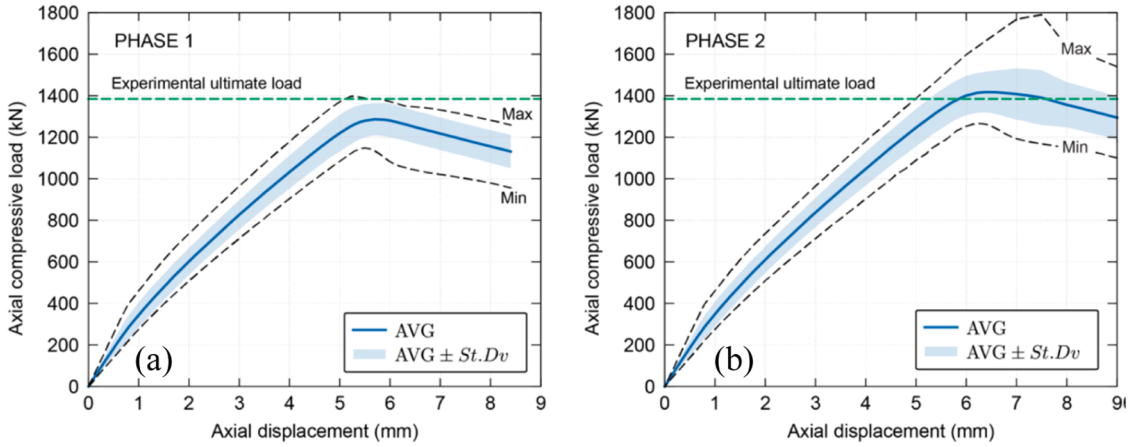


Fig. 7. Resulting average (AVG) end-shortening curves and standard deviations (St.Dv.) after **(a)** Phase 1 and **(b)** Phase 2.

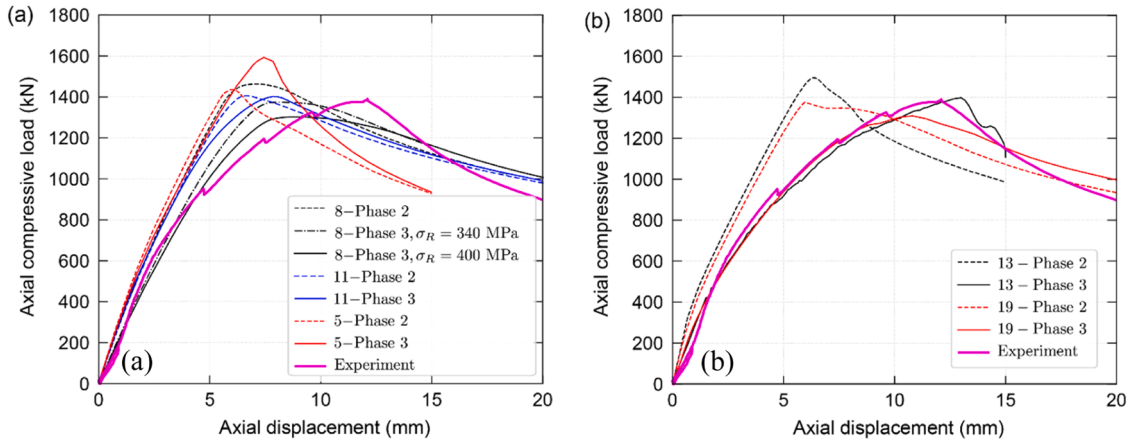


Fig. 8. **(a)** End-shortening curves after introducing residual stresses and the Phase 2 curves from the same participants. **(b)** The effect of including modified boundary conditions that mimic the test conditions.

followed Yi et al. [20] and Paik [15], directly inputting a predefined stress field and ensuring target values were met after convergence. While Participant 11 maintained a uniform residual-to-yield stress ratio of 0.85, Participant 5 applied residual stresses only to key areas, excluding the longitudinal frame. Participant 8′s results showed that residual stresses reduced the ultimate load compared to Phase 2. Similarly, Participant 11 observed a slight decrease in the ultimate load, along with lower stiffness and increased displacement at peak load. At a higher residual stress level of 400 MPa, the reduction became more pronounced. By contrast, Participant 5′s model unexpectedly showed an increase in the ultimate load, possibly due to an imbalance in the applied stress field, leading to residual tensile stress on the beams, which may have enhanced structural strength by influencing the beam buckling collapse mode. Despite these differences, all models exhibited reduced stiffness, aligning with experimental trends. However, as expected by the participants, residual stresses alone do not fully explain the discrepancies between the numerical and experimental results. In Fig. 9, the residual stresses are shown for longitudinal and transverse stresses with the maximum tensile stress in the welds equal to 340 MPa. These tension stresses are in equilibrium with compressive stresses in the interior of the plate, resulting in a softer response.

4.3.2. *Effect of modified boundary conditions*

Further analysis of the end conditions focused on estimating the panel rotations around the transversal *y*-axis using axial displacements at multiple points, as shown in Fig. 10. The results indicate negligible rotation at the loaded edge (actuator side), while significant rotation occurred at the fixed edge. Initially stabilizing at 0.1 degrees, the rotation sharply increased after the first support failure at 930 kN and continued rising linearly until the next failure just before 1200 kN

Building on this observation, two Phase 3 participants examined the impact of boundary conditions and machine compliance on the predicted results. Participant 19 analyzed displacement sensor data from both ends of the test panel, leading to the proposal of constraint conditions that differed from the initially assumed conditions (Fig. 10). The imposed constraints were as follows:

- **Actuator constraints:** these constraints were determined based on the loaded end of the sample, specifically the region of the actuators. Tie constraints were applied to a rigid-link node (reference point) located at the centroid of this set of faces (Fig. 11a).
- **Load cell constraints:** these constraints were related to the regions where the load cells were experimentally placed. The constraints were numerically specified by the nodes of the *i* th load cell support geometry, which were constrained through ties to the corresponding *i* th reference point (Fig. 11b).

The analysis confirmed the absence of significant rotation at the loaded end (Fig. 11a). However, modifications were required for the end in contact with the load cells (Fig. 11b) to align with the experimental records, where one of the load cells supports failed during the experiment. These modifications were based on the following assumptions. First, the load cell acting on the support that failed during the test and consistently showed zero compressive force was considered non-contacting, leaving that area completely free. Second, one support was assumed to be perfectly clamped. Third, the remaining two supports were allowed to rotate around an axis perpendicular to the load direction. The degrees of freedom considered for the load cell region are summarized in Fig. 11c.

Additionally, loading in the model was applied by displacing the master reference point in the load region (Fig. 11a). This was achieved by replicating the experimental load history of the sample using the amplitude from the numerical model. The model closely matches the experimental curve up to approximately 1300 kN (Fig. 8b); beyond this point, it slightly underestimates the load.

In Phase 3, Participant 13 investigated the influence of the loading frame's elastic compliance on structural response. To replicate its spring-like behavior, load pads were introduced at the loaded end of the frame. These $100 \times 100 \times 2.5$ mm thin plates were modeled in ABAQUS/Explicit with normal-hard and surface-to-surface contact conditions. Using iterative reverse engineering, the best match with experimental data was achieved with a 300 GPa elastic modulus and elastic-perfectly plastic behavior (yield strength of 355 MPa). While this adjustment improved correlation, it did not capture support failures or sharp load drops, indicating that

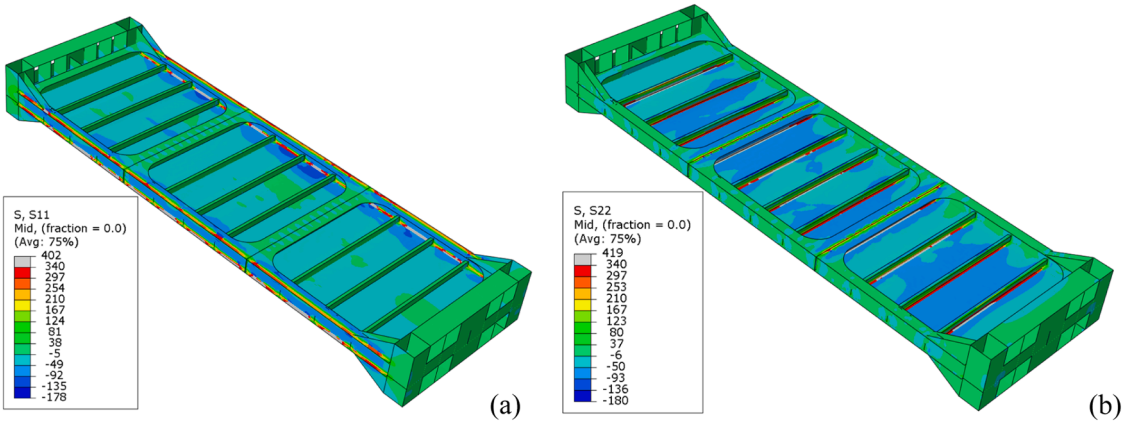


Fig. 9. Residual stresses for **(a)** longitudinal stresses (S11) and **(b)** transverse stresses (S22) with the maximum tensile stress in the welds equal to 340 MPa.

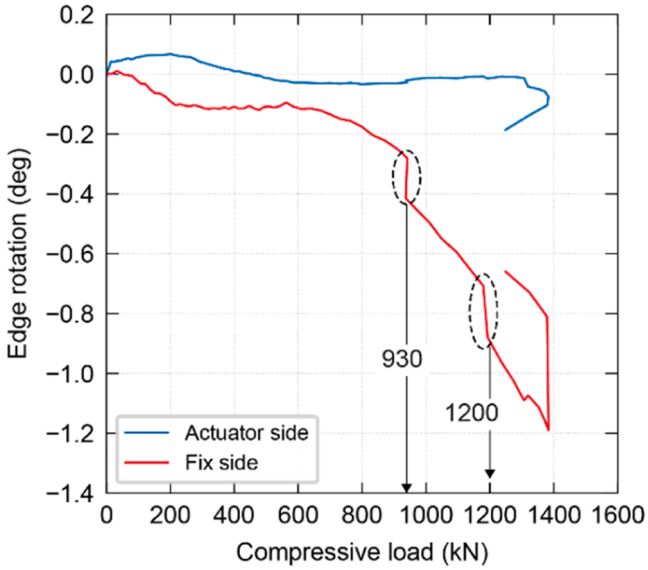
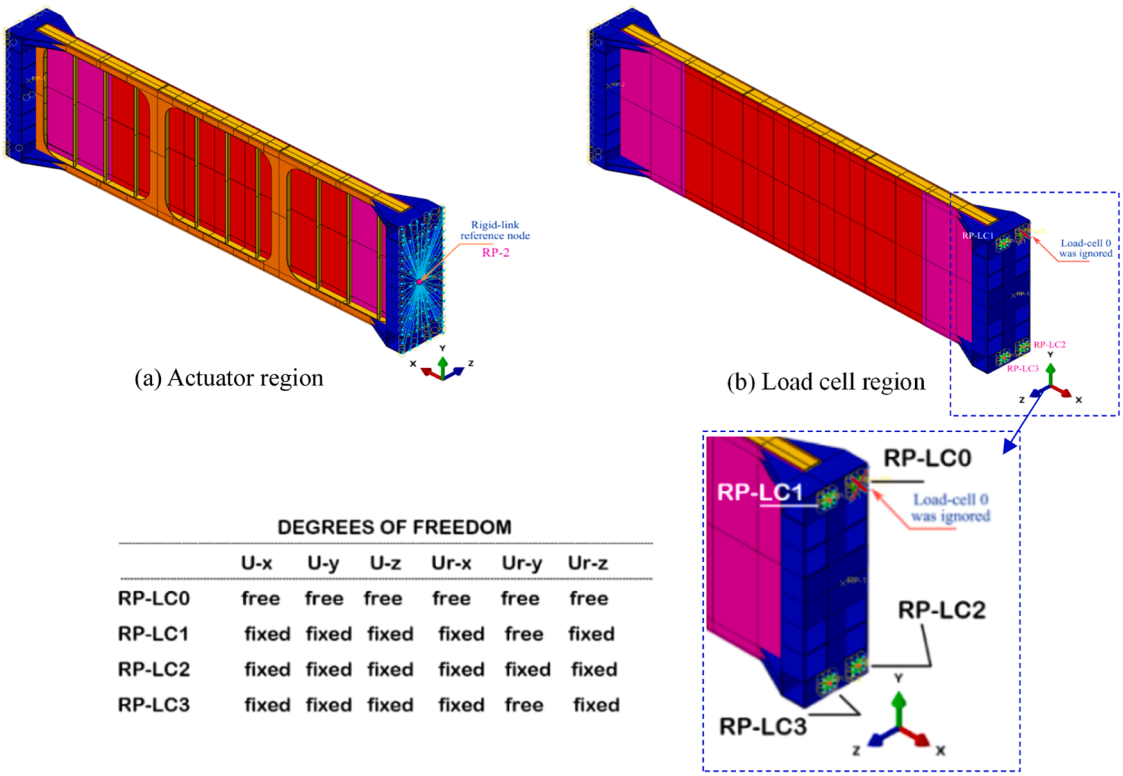


Fig. 10. Rotation at the edges about the transversal *y*-axis during the experiment.



DEGREES OF FREEDOM

	U-x	U-y	U-z	Ur-x	Ur-y	Ur-z
RP-LC0	free	free	free	free	free	free
RP-LC1	fixed	fixed	fixed	fixed	free	fixed
RP-LC2	fixed	fixed	fixed	fixed	fixed	fixed
RP-LC3	fixed	fixed	fixed	fixed	free	fixed

(c) Boundary conditions summary

Fig. 11. Boundary conditions simulated by Participant 19.

numerical–experimental discrepancies likely arise not only from elastic compliance but also from complex boundary conditions that are difficult to model precisely.

5. Comparison with CSR and UR-S35

In this section, the ultimate load predicted by IACS CSR buckling capacity formulas [7], same as in IACS UR S35 [6], is compared with the nonlinear FE analyses conducted in the benchmark study. These rules apply to over 90 % of ships worldwide, in terms of gross tonnage. CSR calculations estimate a maximum allowable stress of 47 MPa for the critical central stiffened plate (see UR-S35 in IACS [6]), corresponding to a force of 593 kN, which is conservative compared to the experimental and FE analysis results.

This discrepancy arises because CSR considers global elastic buckling—where the stiffener and its attached plate buckle together—as the critical failure mode, setting a cut-off limit in the rules. This design approach helps prevent usage of excessively slender stiffeners on ship structures. In this specific case, yield strength does not influence CSR's strength calculations, as global elastic buckling is independent of material yield strength, a factor explored in different phases of the benchmark study. Some regulatory frameworks, such as the DNV Ship Rules [3], allow for the use of nonlinear FE analysis but require separate verification of global elastic stiffener buckling through eigenvalue analysis. The ultimate capacity is defined as the lower value between the global elastic buckling load and the load determined by nonlinear FE analysis. In this study, the global eigenmode obtained through FE analysis yields a buckling load of 891 kN. Under these rules, global elastic buckling becomes the governing factor since it occurs at a lower load than the maximum reached in nonlinear FE analysis.

The higher buckling load computed in the FE analyses compared to CSR is partially due to differences in boundary conditions. CSR assumes simply supported stiffeners, whereas the FE models included welded ends, introducing rotational restraint. This boundary condition significantly impacts global elastic buckling, making the load more sensitive to support constraints. When the plates are modeled with simple supports, the results align more closely with CSR predictions.

6. Conclusions

This benchmark study compared participants' expertise, numerical techniques, and simulation methods to determine their effectiveness in predicting the full end-shortening curve of stiffened panel structures under transverse compression. It also compared FE analysis results against different regulatory guidelines. The key findings and observations are as follows.

Despite the increasing reliability of nonlinear numerical models and the availability of more precise experimental data—gathered from multi-point displacement sensors, strain gauges, and digital image correlation—the process of identifying discrepancies between numerical and experimental approaches has become more complex as the volume and accuracy of data grow. This challenge stems partly from the fact that even with extensive knowledge, it can be easy to concentrate on detailed analyses rather than considering the overall system response. In fact, even though all participants used the same input geometry and were provided with limited modeling parameters in Phase 1, the results varied significantly. This highlights the strong influence of user choices on nonlinear FE analysis outcomes. Establishing a clear correlation between specific modeling characteristics and their impact on the ultimate load value remains challenging.

The initial imperfection models significantly influenced the early end-shortening curve, particularly the local buckling response of the shell plating, though they had little effect on the ultimate strength. Variations in initial stiffness may alter load redistribution across structural components. To minimize abrupt stiffness changes due to local buckling, this study recommends using real imperfections or applying an eigenvalue-derived imperfection pattern with sufficient amplitude (at least 1.5 mm in this study).

Access to material yield strength and engineering ultimate strength data alone does not eliminate modeling uncertainties. Participants made differing assumptions regarding the constitutive material model, increasing result dispersion.

Determining the exact experimental boundary conditions is challenging. However, access to displacement data from multiple measurement points allowed for the formulation of hypotheses regarding actual boundary conditions, refining numerical–experimental data comparisons. To improve accuracy, future experiments should capture axial displacements at intermediate frames, or full-displacement components through optical digital image correlation, and at ends while monitoring all degrees of freedom at the panel ends.

Welding-induced residual stresses had a noticeable effect on the end-shortening curves, although their impact on the ultimate strength remains unclear. The optimal modeling strategy for incorporating these stresses could not be identified due to challenges with the boundary conditions that arose during the test. Nevertheless, Phase 3 revealed that incorporating welding residual stresses in nonlinear FE analysis produces results that closely match the experimentally measured stiffness (Fig. 8a). Yet, discrepancies arise around 950 kN, coinciding with the failure of a panel support and subsequent changes in boundary conditions. To accurately reflect experimental behavior beyond this point, the boundary conditions in numerical models should be adjusted accordingly (Fig. 8b). In conclusion, the analysis conducted in Phase 3 is not sufficiently broad to provide a clear recommendation on how to incorporate residual stresses into the model properly. However, the authors argue that, by introducing residual stresses shown to be self-equilibrating at the level of the individual elementary shell plating, these residual stresses improve the model's approximation of the experimental behavior.

In the current benchmark study, no clear trends were found concerning the choice of solver, element type, or mesh size, except for the initial geometric imperfection model identified in the first phase. However, it was found that introducing residual stresses had a significant impact on the outcomes. Correctly modeling constraint conditions emerged as the most influential factor.

For future research and benchmark studies, it is recommended to record axial displacements at intermediate frames and at the ends

of the panels, as well as to monitor all degrees of freedom at the panel ends. This will help improve the reconstruction of boundary conditions. Clearer and unified modelling guidelines across the classification societies would also be beneficial.

CRediT authorship contribution statement

Marco Gaiotti: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Lars Brubak:** Writing – review & editing, Software, Methodology, Formal analysis, Data curation. **Bai-Qiao Chen:** Software, Formal analysis. **Ionel Darie:** Software, Formal analysis, Data curation. **Dimitris Georgiadis:** Writing – review & editing, Software, Methodology, Formal analysis, Data curation. **Daisuke Shiomitsu:** Software, Formal analysis, Data curation. **Mihkel Kõrgesaar:** Software, Methodology, Formal analysis, Data curation. **Yining Lv:** Software, Formal analysis, Data curation. **Ken Nahshon:** Software, Methodology, Investigation, Formal analysis, Data curation. **Marcelo Paredes:** Writing – review & editing, Software, Methodology, Formal analysis, Data curation. **Jani Romanoff:** Writing – review & editing, Methodology, Conceptualization. **Ingrid Schipperen:** Software, Formal analysis, Data curation. **Akira Tatsumi:** Software, Formal analysis, Data curation. **Murilo Vaz:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis, Data curation. **Yikun Wang:** Software, Methodology, Formal analysis, Data curation. **Albert Zamarin:** Software, Formal analysis, Data curation. **Zhihu Zhan:** Software, Formal analysis, Data curation. **Jonas W. Ringsberg:** Writing – review & editing, Supervision, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Marco Gaiotti reports financial support was provided by Fincantieri Spa. Mihkel Korgesaar reports a relationship with Estonian Research Council that includes: funding grants. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors wish to thank the Fincantieri Naval Vessels Business Unit, Innovation and Products Development Department, for funding the experimental activities reported in this paper. Mihkel Kõrgesaar acknowledges the financial support from Estonian Research Council grant no. PSG754.

Data availability

Data will be made available on request.

References

[1] Barsotti B, Battini C, Gaiotti M, Rizzo CM, Vergassola G. Experimental and numerical assessment of ultimate strength of a transversally loaded thin-walled deck structure. Mar. Struct. 2025;103:103793. https://doi.org/10.1016/j.marstruc.2025.103793.
[2] CEN. BS en 1993-1-14 eurocode 3: design of steel structures. part 1-14: design assisted by finite element analysis. Brussels: European Committee for Standardization; 2023.
[3] DNV. Class guideline DNVGL-CG-0128: buckling. 2021.
[4] Gaiotti M, Barsotti B, Brubak L, Chen BQ, Darie I, Georgiadis D, et al. ISSC 2025 Committee III.1 - Compressive test of a transversely stiffened thin-plated structure with expected early nonlinear response prior to the ultimate capacity: preliminary comparison of the numerical results. In: Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering. June 9–14, 20242; 2024. https://doi.org/10.1115/OMAE2024-126382. V002T02A087.
[5] Gordo JM, Guedes Soares C. Tests on ultimate strength of hull box girders made of high tensile steel. Mar. Struct. 2009;22(4):770–90. https://doi.org/10.1016/j.marstruc.2009.07.002.
[6] IACS. Unified requirements strength UR S-35. Buckling strength assessment of ship structural elements 2024. https://iacs.org.uk/resolutions/unified-requirements/ur-s [accessed October 27, 2025].
[7] IACS. Common structural rules for bulk carriers and oil tankers (CSR) - (Part 1 chapter 8). UK: Electronic Publication: London; 2025. 2025, https://iacs.org.uk/resolutions/common-structural-rules/csr-for-bulk-carriers-and-oil-tankers.
[8] ISSC. Committee III.1: ultimate strength. In: Proceedings of the 18th International Ship and Offshore Structures Congress. September 9–13, 20121; 2012. p. 285–363.
[9] ISSC. Committee III.1: ultimate strength. In: Proceedings of the 21st International Ship and Offshore Structures Congress. September 11–15, 20221; 2022. p. 395–500. https://doi.org/10.5957/ISSC-2022-COMMITTEE-III-1.
[10] ISSC. Committee III.1: ultimate strength. In: Wu Wenwei, Ding Jun, editors. Proceedings of the 22nd International Ship and Offshore Structures Congress. September 22–26, 2025. 1. Wuxi, China: Springer Nature Singapore Pte Ltd.; 2025. p. 419–526.
[11] Li S, Benson SD. A re-evaluation of the hull girder shakedown limit states. Sh. Offshore Struct. 2019;14(Suppl 1):239–50. https://doi.org/10.1080/17445302.2019.1573872.
[12] Manco MR, Vaz MA, Cyrino JCR, Liang DA. Mechanical behavior assessment of FPSO mid-ship repaired stiffened corroded side panels under combined loading. Eng Struct 2022;251:113441. https://doi.org/10.1016/j.engstruct.2021.113441.
[13] Manco MR, Vaz MA, Cyrino JCR, Ramos NM, Liang DA. Experimental and numerical study of uniaxially compressed stiffened plates with different plating thickness. Thin-Walled Struct 2019;145:106422. https://doi.org/10.1016/j.tws.2019.106422.
[14] Paik JK. Residual ultimate strength of steel plates with longitudinal cracks under axial compression–experiments. Ocean Eng 2009;35(17–8):1775–83. https://doi.org/10.1016/j.oceaneng.2008.08.012.

[15] Paik JK. Ultimate limit state analysis and design of plated structures. 2nd ed. Hoboken, NJ: John Wiley & Sons Ltd; 2018. https://doi.org/10.1002/9781119367758.
[16] Paik JK, Lee DH, Noh SH, Park DK, Ringsberg JW. Full-scale collapse testing of a steel stiffened plate structure under cyclic axial-compressive loading. Structures 2020;26(1):996–1009. https://doi.org/10.1016/j.istruc.2020.05.026.
[17] Paik JK, Lee DH, Noh SH, Park DK, Ringsberg JW. Full-scale collapse testing of a steel stiffened plate structure under axial-compressive loading triggered by brittle fracture at cryogenic condition. Sh. Offshore Struct. 2021;15(Suppl 1):S29–45. https://doi.org/10.1080/17445302.2020.1787930. a.
[18] Paik JK, Lee DH, Park DK, Ringsberg JW. Full-scale collapse testing of a steel stiffened plate structure under axial-compressive loading at a temperature of −80 °C. Sh. Offshore Struct. 2021;16(3):255–70. https://doi.org/10.1080/17445302.2020.1791685. b.
[19] Smith CS. Compressive strength of welded steel ship grillages. The Royal Institute of Naval Architects 1975:325–59.
[20] Yi MS, Hyun CM, Paik JK. Three-dimensional thermo-elastic-plastic finite element method modeling for predicting weld-induced residual stresses and distortions in steel stiffened-plate structures. World J. Eng. Technol. 2018;6(1):176–200. https://doi.org/10.4236/wjet.2018.61010.
[21] Zhang S, Khan I. Buckling and ultimate capability of plates and stiffened panels in axial compression. Mar. Struct. 2009;22(4):791–808. https://doi.org/10.1016/j.marstruc.2009.09.001.