Open access: https://journal.icpa-banyuwangi.ac.id/index.php/skyhawk

SKYHAWK: Jurnal Aviasi Indonesia
Volume. 6, Number. 2, 2026, pp. 751-769

P-ISSN: 2745-6102
E-ISSN: 2776-3854

Development and Internal Validation of a Risk-Based Safe-Manning Framework for Wing-in-Ground Craft

Dedi Kurniawan[1], Andi Fiardi[2], Baihaqi[3], Antoni Arif Priadi[4]
[123]Malahayati Maritime Polytechnic, Greater Aceh, Indonesia
[4]Directorate General of Sea Transportation, Ministry of Transportation, Jakarta, Indonesia

Email Address: andi.fiardi@poltekpelaceh.ac.id[2], baihaqi@poltekpelaceh.ac.id[3], antoni.pip.smg@gmail.com[4]
Email Correspondence: dedikurniawan@poltekpelaceh.ac.id[1]

ARTICLE INFO

Article History

Received: 15 July 2026
Revised: 01 September 2026
Accepted: 17 September 2026

Keywords

Crew competence, Human factors, Safe manning, Task-network simulation, WIG craft

ABSTRACT

Purpose. This study developed and computationally verified a function-based method for determining core operational crew for wing-in-ground (WIG) craft. **Methods.** A frozen 37-source corpus was mapped to six safety-critical functions, six mission phases, a competence matrix, and phase loads. Bounded integer compositions were tested using an exact scalar max-flow/min-cut model, 20% reserve, utilization caps, role-removal screening, and 10,000-trial simulation. Robustness analyses varied reserve and utilization, demand distributions, supporting competence coefficients (±20%), ten random seeds, and every present-role removal. **Results.** Baseline complements were three, four, five, and six crew for restricted Type A, coastal Type A passenger, demanding Type B, and remote Type C scenarios. In the primary triangular run, S1-S3 had no failures (Wilson 95% lower bound 99.962%); S4 had one (99.99%; 95% CI 99.943-99.998). The crew sequence persisted in 45/48 reserve-utilization cells and all coefficient perturbations. Across ten seeds, triangular feasibility was 100% for S1-S3 and 99.98%-100% for S4. Under lognormal-tail stress, ranges were 100%, 99.95%-100%, 99.68%-99.84%, and 96.94%-97.25% for S1-S4. No present-role removal was more favorable than the designated comparators. **Conclusions.** Concurrent transition/emergency workload, qualified takeover, fault management, and prolonged-survival capability should drive WIG manning. The 3-6-person outputs are internally consistent engineering baselines, not empirically validated or statutory minima; Type B/C tail sensitivity requires external validation.

1. INTRODUCTION

Wing-in-ground (WIG) craft obtain aerodynamic support by operating close to the water surface. The reduction in induced drag offers a transport regime between displacement vessels and conventional aircraft, but the operational system is more complex than either analogy suggests. A mission may include harbour manoeuvring, waterborne acceleration, transition to aerodynamic support, high-speed cruise near an irregular surface, temporary or sustained out-of-ground-effect operation for Types B and C, and a return to displacement mode. Each regime changes the control margins, visual cues, applicable navigation rules, and available emergency responses (IMO, 2002, 2018; Rozhdestvensky, 2006; Yun et al., 2010). Immediately before and during transition, the team must retain a COLREG-compliant lookout and collision-avoidance picture while shifting to aeronautical threat-and-error management (TEM), energy and flight-path control, mode awareness, and escape-option planning (ICAO, 2018, 2022b; IMO, 1972). This dual-domain overlap explains why transition can create a sharper workload peak than routine cruise.

The hybrid operating envelope creates a distinct manning problem. A complement that can manipulate the controls in steady cruise is not necessarily sufficient during transition, approach, collision avoidance, propulsion failure, automation reversion, or passenger evacuation. During these periods, flight-path control, lookout, route monitoring, radio communication, fault diagnosis, and emergency coordination can become simultaneous rather than sequential tasks. Wickens's four-dimensional multiple-resource model distinguishes processing stage, perceptual modality, visual channel, and processing code; interference rises when concurrent tasks draw on common resources (Wickens, 2002, 2008). Visual-auditory-cognitive-psychomotor (VACP) coding is used here as a practical task-screening lens, but it is not treated as identical to the four-dimensional theory and is not explicitly solved as four separate resource commodities. Situation awareness and recovery capacity can also deteriorate when an operator must supervise opaque automation while diagnosing an abnormal state (Endsley, 1995, 2021; Parasuraman et al., 2000). Under abnormal conditions, work can shift from skill- and rule-based performance toward knowledge-based problem solving, while latent system conditions shape the error opportunities presented to the crew (Rasmussen, 1983; Reason, 1990).

International requirements provide relevant principles but not a universal numerical answer. IMO guidance defines WIG craft categories and recommends craft-specific knowledge and training; the High-Speed Craft Code contributes technical-reliability and passenger-protection logic; STCW and the Maritime Labour Convention address competence, watchkeeping, fitness, work and rest; and aviation standards contribute type-specific proficiency, checking, crew coordination and safety-management concepts (ICAO, 2018, 2022a, 2022b; ILO, 2022; IMO, 2005, 2011, 2020, 2022). Resolution A.1047(27) is itself function- and task-oriented and calls for peak-workload and hours-of-rest considerations. The inadequacy arises when implementation is reduced to static proxies such as gross tonnage, propulsion power or passenger count. Those descriptors may support initial screening, but they cannot represent the short, highly coupled transition and emergency peaks of WIG operation.

Earlier research on WIG craft concentrated mainly on aerodynamic performance, stability and configuration (de Divitiis, 2005; Halloran & O'Meara, 1999; Nebylov, 2002; Rozhdestvensky, 2006). A recent commercialization study confirms renewed interest in WIG transport but also identifies certification, infrastructure and operational barriers that remain craft- and route-specific (Kerem et

al., 2025). Recent deck-officer workload assessment and a human-centred review of maritime autonomy likewise show that manning and automation credit must be grounded in task demand, operator workload, interface transparency and credible fallback arrangements (Li & Yuen, 2024; Pathak & Bhardwaj, 2024). These research streams still do not represent the coupled waterborne-airborne phases of WIG operation or convert them into a qualified complement. The unresolved problem is methodological: how can a regulator or operator convert mission functions into a minimum qualified complement without hiding task conflicts behind a simple crew table?

This study addresses that problem by developing and internally validating a risk- and workload-based safe-manning framework. In this article, internal validation means computational verification of implementation, reproducibility and robustness within the specified model; it does not mean empirical agreement with real WIG operations. Risk-based denotes scenario- and consequence-informed screening with explicit uncertainty and degraded-state tests; it is not a calibrated probability-by-consequence risk estimate. The contributions are: (1) a frozen, auditable evidence corpus mapped to six safety-critical functions; (2) a phase-function allocation model with integer role multiplicity, competence limits, utilization caps and role-removal survival testing; (3) a stochastic task-network simulation of environmental demand, automation reversion and fatigue-related capacity; (4) reserve-utilization and distributional robustness analyses; and (5) second-round checks covering competence-coefficient perturbation, independent-seed replication, all present-role removals and binding-cut failure diagnostics.

The research questions were: (1) which core operational complements satisfy robust concurrent coverage in four representative WIG scenarios; (2) whether those complements remain feasible under bounded uncertainty; (3) which phases expose the safety deficit after removal of one person; (4) whether the selected crew sizes persist when reserve and utilization assumptions change; (5) whether conclusions persist under beta-PERT and long-tailed lognormal demand factors across independent seeds; and (6) whether crew selection persists when supporting competence coefficients are perturbed and every present role is removed in turn. The outputs are engineering baselines for craft-specific safety cases, not statutory minima.

2. METHODS

2.1. Research design and evidence protocol

The research combined structured evidence synthesis, functional task analysis, exact bounded integer enumeration, max-flow allocation testing, role-removal verification, and stochastic task-network simulation. The design was selected because publicly available WIG accident, workload and crew-performance datasets are not sufficiently mature for direct statistical estimation of manning coefficients. A transparent engineering model is preferable to an opaque regression or an unqualified expert score when assumptions, source code and claim boundaries are reported.

The analytical corpus was frozen on 15 July 2026 at 37 sources. Protocol-driven targeted retrieval covered official IMO, ICAO, ILO, ISO and IEC repositories and publisher-indexed literature. Four concept families were used: ("wing-in-ground" OR WIG OR ekranoplan) AND (crew OR manning OR competence OR training OR safety); (minimum safe manning OR reduced crew) AND (workload OR task allocation OR automation); (ship bridge OR flight deck) AND (simulator OR workload OR situation awareness); and (single-pilot OR autonomous ship) AND (incapacitation OR fallback OR human factors).

Sources were included when they were primary international instruments, peer-reviewed articles, scholarly monographs or official technical reports that defined at least one operational function, competence requirement, workload mechanism, automation/fallback issue, redundancy principle or validation metric. Commercial promotion, generic crew-management blogs, sources without traceable authorship, and aerodynamic studies with no operational or human-system implication were excluded. Each retained source was coded against the six functions and the constructs phase demand, competence capacity, concurrency, automation credit, fatigue, redundancy and validation. The row-level audit trail and exact search protocol are supplied in Supplementary Material S1. This was a structured targeted synthesis rather than a systematic review, so no claim of exhaustive retrieval or PRISMA-complete screening is made.

Seven post-freeze references were added across the two revision rounds for context and validation design: BMKG operational weather information, Indonesian search-and-rescue legislation, Wickens's 2008 synthesis, the SPAR-H report, and three recent studies on WIG commercialization, maritime-autonomy human factors and safe-manning workload (Kerem et al., 2025; Li & Yuen, 2024; Pathak & Bhardwaj, 2024). None was used to alter the frozen 37-source numerical coding, phase loads or competence coefficients. This separation preserves the analytical freeze while updating the manuscript's academic and regulatory context.

Table 1. Evidence domains and their analytical contribution

Evidence domain	Representative sources	Contribution to the model
WIG craft regulation	IMO 2002, 2005, 2018; Kerem et al. 2025	Craft categories, operating envelope, officer competence, emergency philosophy
Safe manning and labour	IMO 2011, 2022; ILO 2022; Pathak & Bhardwaj 2024	Functional coverage, peak workload, watchkeeping, work-rest and fatigue constraints
High-speed and passenger safety	IMO 1974, 2020	Technical reliability, evacuation, survival and passenger protection
Aviation licensing and SMS	ICAO 2018, 2022a, 2022b	Type proficiency, checking, TEM, crew coordination and safety-case logic
Human performance	Endsley 1995, 2021; Wickens 2002, 2008; Salas et al. 2006	Situation awareness, resource conflict, monitoring and CRM
Automation and reduced manning	Parasuraman et al. 2000; IMO 2021; Li & Yuen 2024	Conditional automation credit, fallback, role redistribution and human responsibility
Risk and validation	ISO 2018; IEC 2019; Liu et al. 2023	Robustness, uncertainty treatment and simulator measurement

The evidence synthesis supports the architecture of the model but does not transform normalized task loads into empirical fleet coefficients. No human-participant or expert-consensus data are claimed. A preregisterable two-round Delphi and human-in-the-loop simulator protocol is supplied as Supplementary Material S2 for subsequent external validation.

2.2. Mission phases and safety-critical functions

The mission network used six phases: preparation; surface taxi; waterborne acceleration and transition; ground-effect or authorized out-of-ground-effect cruise; approach and alighting; and abnormal or emergency response. Type B and Type C out-of-ground operation was represented by higher control, navigation and emergency loads within cruise and approach rather than by an

additional average phase. This arrangement preserves the model's focus on peak concurrent demand.

The transition phase explicitly includes the overlap between maritime collision-avoidance duties and aeronautical TEM. A task was treated as concurrent when delaying it or assigning it sequentially to the same individual could reduce the control margin, collision-avoidance margin or emergency response. The emergency phase includes both immediate stabilization and the initiation of sustained survival management when rescue is not immediate.

Six functions were required: command, flight-path control, navigation and collision avoidance, communications, technical-system management, and emergency response. Passenger and cabin duties were represented within emergency demand and through a safety and emergency officer role; additional cabin attendants remain outside the core complement and must be determined by exit geometry, passenger distribution, survival equipment and evacuation trials.

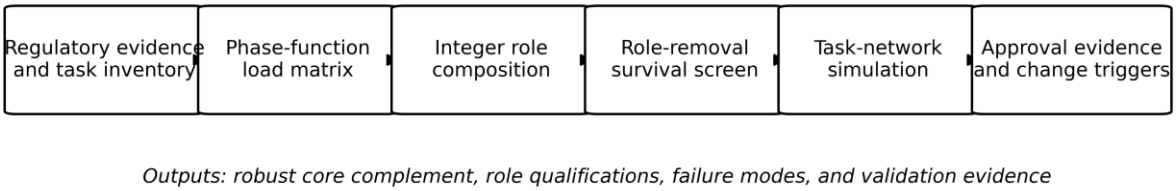

Figure 1. Reproducible workflow for WIG crew determination and internal validation.

Table 2. Mission-phase decomposition and concurrency triggers

Phase	Representative tasks	Peak functions	Principal trigger
Preparation	Route, weather, mass, systems and briefings	Command; navigation; technical	Latent errors propagate through the mission
Surface taxi	Lookout, COLREG decisions, manoeuvring and radio	Control; navigation; communications	Mixed traffic and restricted visibility
Transition	Maintain COLREG lookout while shifting to TEM, energy/height control, mode monitoring, surface scan and restraints	Control; navigation; technical; emergency	Dual-domain compliance, highest concurrency and smallest correction margin
Cruise	Flight path, wave/traffic scan and automation monitoring	Control; navigation; communications	Automation monitoring and degraded references
Approach/alighting	Traffic clearance, descent, surface-state assessment and landing decision	Control; navigation; emergency	Late correction has high consequence
Abnormal/emergency	Stabilize, diagnose, distress, divert/alight, evacuate and sustain survivors	All six functions	Concurrent technical, communications and occupant demand; duration may be prolonged

2.3. Task demand and competence capacity

For scenario s, phase p and function f, baseline demand d_spf was expressed in person-equivalent workload. A value of 1.0 denotes the amount one fully competent person could cover if no competing task used the same resources. VACP screening informed identification of visually,

auditorily, cognitively and psychomotorically competing tasks, but the numerical demand passed to the optimizer remains a homogeneous scalar. The current algorithm therefore does not impose an explicit within-dimension interference penalty—for example, when two simultaneous tasks are both visual-manual. Phase-specific loads, utilization caps and the reserve factor partially protect against such conflicts, but they do not replace a dimension-specific workload model. This is a deliberate and now explicit limitation.

Five functional role types were modeled: master/WIG pilot (MP), first officer/co-pilot (FO), navigation and communications officer (NCO), technical systems officer (TEO), and safety and emergency officer (SEO). Normalized competence capacity a_rf ranges from 0 to 1. A value of 1 indicates independent capability for a function, conditional on a current craft-specific qualification; lower values indicate supporting capacity. The matrix is an authored design input, not an estimate from expert elicitation, operational observations or statistical calibration. The research team qualitatively cross-walked primary accountabilities and supporting duties against the frozen evidence corpus and task inventory using an ordinal normalization rubric: 1.00 for independent primary capability; 0.75-0.95 for near-independent alternate capability; 0.40-0.70 for transferable partial capability; 0.20-0.35 for limited support; and 0.05-0.15 for marginal contingency support. Intermediate values encode judgment within those bands. This provenance makes the matrix auditable but does not establish external construct validity, which requires expert and simulator calibration.

Table 3. Normalized role-to-function competence-capacity matrix

Role	Command	Control	Navigation	Communications	Technical	Emergency
MP	1.00	1.00	0.75	0.60	0.25	0.40
FO	0.70	0.95	0.90	0.80	0.30	0.50
NCO	0.35	0.35	1.00	1.00	0.20	0.40
TEO	0.20	0.10	0.25	0.30	1.00	0.65
SEO	0.15	0.05	0.20	0.55	0.20	1.00

Note. MP = master/WIG pilot; FO = first officer/co-pilot; NCO = navigation and communications officer; TEO = technical systems officer; SEO = safety and emergency officer.

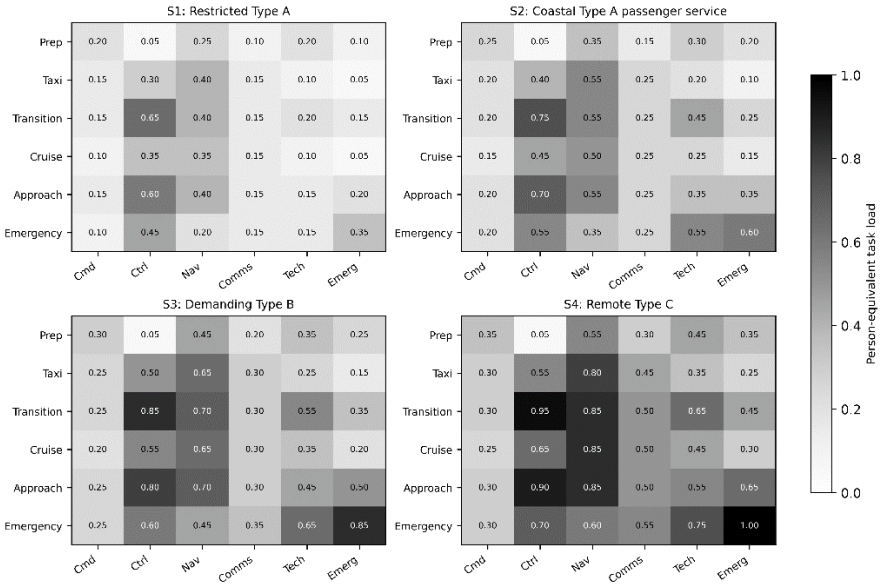

Figure 2. Phase-function task-load matrices for the four methodological verification scenarios.

2.4. Robust integer role-allocation model

Let n_r be the non-negative integer number of persons assigned to role r. The design objective minimizes the core operational complement N. Unlike a binary role selector, the formulation permits duplicate roles where the mission requires more than one independently qualified person.

$$\text{minimize } N = \Sigma_r n_r \quad (1)$$

Let y_rpf denote the workload allocated from role r to function f in phase p. Coverage was imposed on robust demand ρd_spf, with ρ = 1.20. Demand inflation and utilization caps address different failure modes: ρ protects against under-estimated demand, whereas u_p preserves unallocated capacity for monitoring and unplanned events.

$$\Sigma_r y_{rp\ f} \geq \rho\, d_{s\ p\ f} \text{ for every phase p and function f} \quad (2)$$

Preparation, surface taxi and cruise used u_p = 0.90; transition, approach and emergency used u_p = 0.85. Relative to nominal demand and full utilization, the combined structural pressure is 1.20/0.90 = 1.333 in normal phases and 1.20/0.85 = 1.412 in critical phases. This compounding is a deliberate engineering barrier against under-manning, not an empirically estimated psychological constant.

$$\Sigma f\, y_{rp\ f} \leq u_p\ n_r \text{ for every role r and phase p} \quad (3)$$

Competence limits prevented allocation to a role without relevant capacity. Each phase was evaluated as a bipartite network: Equation (3) limits total utilized role capacity, and Equation (4) limits each role-to-function edge by both competence and phase utilization. This is the same network implemented in the released code. The max-flow/min-cut theorem is exact for this published scalar network and prevents one person from being credited with full capacity on several functions simultaneously. Nevertheless, capacity remains fungible within the edges allowed by the role matrix. The model cannot represent separate visual, auditory, cognitive and psychomotor commodities or cross-dimensional interference. A dimension-specific extension would require a materially more complex multi-resource formulation and, for indivisible tasks, mixed-integer or unsplittable assignment constraints.

$$0 \leq y_{rp\ f} \leq u_p\ a_r f\, n_r \quad (4)$$

The bounded design space allowed zero, one or two persons in each role and no more than eight core crew. Every scenario required one MP, at least one FO and a TEO; passenger service required an SEO; demanding Type B and Type C scenarios required an NCO; and the remote Type C scenario required a second FO. All $3^5 = 243$ role combinations were enumerated exactly and ranked by total crew, duplicate roles and residual capacity.

2.5. Role-removal survival screen

Every selected composition was re-evaluated after removing one person from each present role in turn. The remaining team had to cover a reduced survival vector for stabilization, distress communication, immediate navigation, technical isolation and initiation of a safe alighting, evacuation and survival sequence. For remote service this screen does not imply that the emergency ends after evacuation: it requires preserved capability to begin passenger accountability, casualty triage, survival-craft management, exposure protection, rationing, watchkeeping and repeated communications until external rescue becomes available. Duration is route-specific and is not simulated by the present vector.

$$\text{Feasible}(n - e_q, h_s) = 1 \text{ for every present role q} \quad (5)$$

SPAR-H is cited only as a conceptual analogy for interpreting performance-shaping conditions; it is not a method applied in this study. The role-removal screen verifies whether the remaining crew can retain specified functional coverage. It does not estimate the probability that incapacitation will occur, calculate a human-error probability (HEP), or quantify the probability of mission failure due to incapacitation. Loss of a crew member may reduce available time, raise stress and complexity, and increase dependence between successive errors; SPAR-H represents such mechanisms through calibrated performance-shaping factors and dependency rules applied to base HEPs (Gertman et al., 2005). A WIG-specific error model requires separate empirical calibration, and no exponential HEP relationship is claimed here.

2.6. Verification scenarios and stochastic task-network simulation

Four archetypal scenarios were used. S1 represented restricted Type A daylight operation in sheltered water with eight persons on board. S2 represented coastal Type A passenger service in moderate traffic with 30 persons on board. S3 represented demanding Type B operation at night and in open water with 50 persons on board. S4 represented a remote Type C mission with sustained out-of-ground capability, limited rescue support and 100 persons on board. For any Indonesian S4 application, the emergency-planning duration must be set from route-specific communications, survival and rescue assumptions agreed with the relevant authorities, including BASARNAS under Indonesia's search-and-rescue framework (Government of Indonesia, 2014).

The cases are ordered methodological verification scenarios, not descriptions of a specific commercial craft. Complete phase-function matrices are supplied in Supplementary Material S1.

The primary internal-consistency run used 10,000 stochastic trials for each selected complement and one designated N-1 comparator, formed by removing the last scenario-triggered dedicated role (TEO in S1, SEO in S2, NCO in S3, and one FO in S4). The designated cut represented an operationally intuitive crew-reduction decision, not an assertion that it was the strongest or weakest possible removal. A separate second-round analysis therefore tested every distinct present-role removal. Baseline loads were multiplied by scenario-level environmental, phase and function factors drawn from bounded triangular distributions. One environmental draw was shared across all phases and functions in a trial, each phase draw across functions, and each function draw across phases. These common draws induced structured positive dependence, while the underlying environmental, phase, function and fatigue variates were sampled independently. Triangular distributions were selected because only transparent minimum-mode-maximum engineering anchors were available and mature WIG fleet distributions do not exist. The choice is not inherently conservative. Automation-reversion events increased control, navigation and technical demand, while a bounded fatigue factor reduced available capacity. The primary fixed seed was 20260715.

$$D_{s\ p}\ f^{(m)} = d_{s\ p}\ f\ E_s^{\ (m)}\ P_{s\ p}^{\ (m)}\ F_s\ f^{(m)}\ R_{s\ p}\ f^{(m)} \qquad (6)$$

The primary endpoint was whole-mission feasibility across all phases. Wilson 95% confidence intervals were calculated. A priori checks required monotonic crew size, passage of reserve and role-removal screens, at least 99.9% simulated feasibility for selected teams, and materially lower feasibility for the designated comparator. The 99.9% rule was applied to the simulated point estimate; Wilson intervals describe sampling precision and were not separate pass/fail criteria. First-failure phases were recorded.

2.7. Sensitivity, replication and diagnostic analyses

The first added analysis varied ρ over 1.00, 1.10, 1.20 and 1.30 and varied the transition/approach/emergency utilization cap over 0.80, 0.85 and 0.90 while retaining 0.90 for preparation, surface taxi and cruise. All 48 scenario-parameter cells were re-enumerated with unchanged phase loads, survival vectors and scenario-specific role floors, using the same role-removal test. The ρ = 1.30/critical-cap = 0.80 cell is the mathematically highest-pressure corner by construction—highest demand inflation combined with lowest available critical-phase capacity—not a corner selected after inspecting the results; every cell in the 4 × 3 grid was evaluated. The analysis tests robustness of the complete decision rule, not an unconstrained workload-only model.

The second added analysis compared three demand-factor families using 10,000 trials per design and common random uniforms. The baseline triangular model was compared with bounded beta-PERT distributions (shape parameter 4) and a deliberately severe right-tailed lognormal stress test. The lognormal mean was matched to the corresponding triangular mean, and its 99th percentile was set at twice the triangular upper excursion above 1.0. The factor of two was selected as a transparent illustrative severity multiplier to expose a tail-sensitive failure boundary; it was not calibrated to a measured WIG overload event or treated as a comparable aviation or maritime fleet distribution. Fatigue and automation-reversion mechanisms were otherwise unchanged. Supplementary Material S1 reports every minimum-mode-maximum triplet, the deterministic lognormal calibration, dependence construction, fixed seeds, code and complete outputs.

A third analysis perturbed every supporting competence coefficient below 1.00 simultaneously by factors of 0.80, 0.90, 1.10 and 1.20, capped at 1.00, while primary independent-capability coefficients remained fixed at 1.00. Each scenario was re-enumerated with unchanged loads, reserve, utilization caps, role floors and role-removal criteria. This global bounded perturbation tests local robustness of the authored matrix; it does not replace empirical construct validation.

A fourth analysis repeated each selected-team distributional simulation for 10,000 trials at ten fixed seeds (20260715-20260724) and reported the replication mean, standard deviation and range. A fifth analysis applied identical triangular stochastic draws to every distinct present-role removal; removal of either one of two identical FO positions in S4 was treated as the same role-type cut. For S4, failed trials were further classified by first infeasible phase and the most violated max-flow/min-cut functional subset.

3. RESULTS

3.1. Functional evidence synthesis

The synthesis confirmed that no single regulatory domain represents the complete WIG operating system. Maritime instruments define navigation, watchkeeping, technical, labour and survival obligations; aviation instruments contribute flight-crew proficiency, TEM, checking and safety-management logic; and WIG-specific guidance supplies craft categories, operational limitations and type-specific knowledge. The model therefore treats safe manning as a cross-domain safety case rather than as a direct conversion of either a ship safe-manning document or an aircraft cockpit complement.

The phase-function matrices showed two consistent workload concentrations. Transition combined high control demand with navigation, technical surveillance and emergency readiness. The emergency phase combined stabilization with fault diagnosis, communication and occupant

protection. Cruise loads were lower in aggregate, although automation and remote-route assumptions increased monitoring demand in S3 and S4.

3.2. Robust core operational complements

Exact enumeration produced a monotonic increase from three to six core operational crew. The selected compositions satisfied every phase at the 20% design load and passed removal of each present role against the survival vector. Table 4 reports the preferred functional role mix and the minimum reserve ratio; a ratio above 1 indicates that the smallest capacity cut remained above demand. The base reserve ratio declined monotonically from 1.500 in S1 to 1.308 in S4 (and from 1.250 to 1.090 at $\rho = 1.20$), showing that the larger, more complex scenarios operate with progressively thinner modeled margins.

S1 did not support a two-person conceptual baseline. A third person was required because the command pair could not simultaneously preserve control, navigation and fault-management capacity under robust transition demand. S2 added an SEO to prevent passenger and emergency duties from drawing control and technical crew away from their stations. S3 added an NCO, and S4 added a second FO to preserve control redundancy in remote Type C operation. Cabin attendants and mission specialists are additional to these core teams and require separate evacuation or mission analysis.

Table 4. Robust core operational complements

Scenario	Crew	Preferred functional roles	Base reserve	Reserve at ρ = 1.20	Role removal
S1	3	1 MP + 1 FO + 1 TEO	1.500	1.250	Pass
S2	4	1 MP + 1 FO + 1 TEO + 1 SEO	1.360	1.133	Pass
S3	5	1 MP + 1 FO + 1 NCO + 1 TEO + 1 SEO	1.349	1.124	Pass
S4	6	1 MP + 2 FO + 1 NCO + 1 TEO + 1 SEO	1.308	1.090	Pass

3.3. Simulation-based internal consistency testing

Under the primary triangular uncertainty run, selected teams had no failed trials in S1-S3; for each zero-failure estimate, the Wilson 95% interval was 99.962%-100.000%. S4 was feasible in 9,999 of 10,000 trials (99.99%; 95% CI 99.943%-99.998%). The designated one-person-reduced comparators failed frequently: S1 and S2 were feasible in only 12.61% and 26.80% of trials, while S3 and S4 were feasible in 70.57% and 62.87%.

Table 5. Task-network simulation results, 10,000 trials per design

Scenario	Selected crew	Selected feasible % (95% CI)	N-1 crew	N-1 feasible % (95% CI)
S1	3	100.00 (99.962-100.000)	2	12.61 (11.97-13.28)
S2	4	100.00 (99.962-100.000)	3	26.80 (25.94-27.68)
S3	5	100.00 (99.962-100.000)	4	70.57 (69.67-71.46)
S4	6	99.99 (99.943-99.998)	5	62.87 (61.92-63.81)

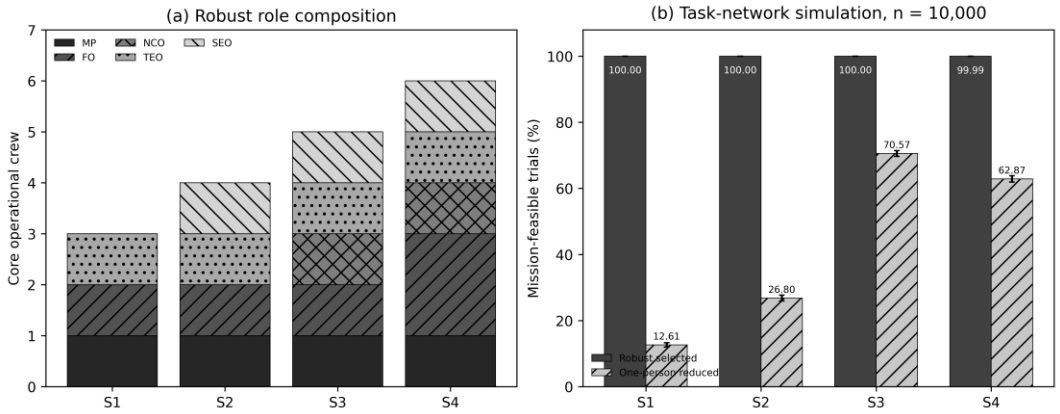

Figure 3. Selected role compositions and discrimination from the designated one-person-reduced comparators. Error bars show Wilson 95% confidence intervals.

3.4. Failure-phase discrimination

The first failed phase differed systematically by scenario. In the designated reduced S1 comparator, 98.01% of failed trials first became infeasible during transition. S2 failures were concentrated in transition and emergency response. S3 and S4 failures were dominated by emergency response, with additional approach and transition failures. No designated comparator first failed during preparation, surface taxi or cruise.

Table 6. Distribution of the first infeasible phase in designated one-person-reduced comparator trials

Scenario	Failed trials, n (% of 10,000)	Transition (% of failures)	Approach (% of failures)	Emergency (% of failures)
S1	8,739 (87.39)	98.01	1.99	0.00
S2	7,320 (73.20)	60.37	6.17	33.46
S3	2,943 (29.43)	17.87	8.05	74.07
S4	3,713 (37.13)	21.28	20.44	58.28

Note. Phase percentages use failed trials—not all 10,000 trials—as the denominator and may not total 100.00 because of rounding

3.5. Sensitivity analyses

The baseline crew size was retained in 45 of the 48 reserve-utilization cells (Table 7, Panel A). S1 remained at three in all 12 cells. S2, S3 and S4 each retained their baseline size in 11 cells and required one additional person only when the most demanding settings were combined ($\rho = 1.30$ and critical cap = 0.80). Because all scenario-specific role floors remained active, these counts describe robustness of the complete stated decision rule rather than workload capacity with the role logic removed.

The bounded beta-PERT alternative produced 10,000 feasible missions in 10,000 trials for every selected team (100.00%; Wilson 95% CI 99.962%-100.000%). Under the mean-matched lognormal-tail stress test, selected-team feasibility was 100.00% for S1 (95% CI 99.962%-100.000%), 99.95% for S2 (99.883%-99.979%), 99.72% for S3 (99.596%-99.806%) and 97.13% for S4 (96.784%-97.440%). Designated one-person-reduced comparators remained markedly weaker across all families. Under the prespecified 99.9% point-estimate rule, S1 and S2 pass, whereas S3 is a marginal failure and S4 a material failure.

The only primary-seed triangular failure in S4 first occurred in Emergency and exceeded the aggregate six-function cut by 0.00095 person-equivalent capacity. Under lognormal-tail stress, 287 S4 trials failed: 154 first failed in Emergency (53.66%), 70 in Approach (24.39%) and 63 in Transition (21.95%). In every failed trial, the most violated cut contained all six functions rather than one isolated role-function edge; the median aggregate overload was 0.124 and the maximum was 1.109 person-equivalent capacity.

Across ten seeds, triangular runs had no failures for S1-S3 (per-run Wilson lower bound 99.962%) and ranged from 99.98% to 100.00% for S4; beta-PERT runs had no failures for any selected team (same lower bound). Lognormal-tail ranges were 100.00% for S1 (same lower bound), 99.95%-100.00% for S2, 99.68%-99.84% for S3 and 96.94%-97.25% for S4. Thus all ten replications preserved the threshold classification.

Scaling all supporting competence coefficients from 0.80 to 1.20 of baseline did not change the 3-4-5-6 crew sequence. Simulating every distinct present-role removal yielded scenario-specific feasibility ranges of 12.61% (S1), 26.65%-26.80% (S2), 70.57% (S3) and 62.87% (S4); no alternative removal was materially more feasible than the designated comparator. Complete replication, coefficient-sensitivity, role-removal and diagnostic outputs are provided in Supplementary Material S1.

Table 7. Added sensitivity analyses

Panel A. Reserve-factor and critical-utilization sensitivity

Scenario	Baseline crew	Grid cells retaining baseline (of 12)	Resulting total crew at ρ = 1.30, critical cap = 0.80	Extreme-case change
S1	3	12	3	None
S2	4	11	5	+1 NCO
S3	5	11	6	+1 SEO
S4	6	11	7	+1 SEO

Note. Phase loads, survival vectors and all scenario-specific role floors were held fixed across the 4 × 3 grid.

Panel B. Distributional sensitivity; 10,000 trials per design

Scenario	Triangular feasible % (95% CI)	Beta-PERT feasible % (95% CI)	Lognormal-tail feasible % (95% CI)	N-1 feasibility range (%)
S1	100.00 (99.962-100.000)	100.00 (99.962-100.000)	100.00 (99.962-100.000)	12.61-24.65
S2	100.00 (99.962-100.000)	100.00 (99.962-100.000)	99.95 (99.883-99.979)	26.80-37.53
S3	100.00 (99.962-100.000)	100.00 (99.962-100.000)	99.72 (99.596-99.806)	53.40-80.31
S4	99.99 (99.943-99.998)	100.00 (99.962-100.000)	97.13 (96.784-97.440)	48.03-73.08

Note. Values are point estimates with Wilson 95% confidence intervals; the 99.9% decision rule applies to point estimates. Complete multi-seed results are reported in Supplementary Material S1.

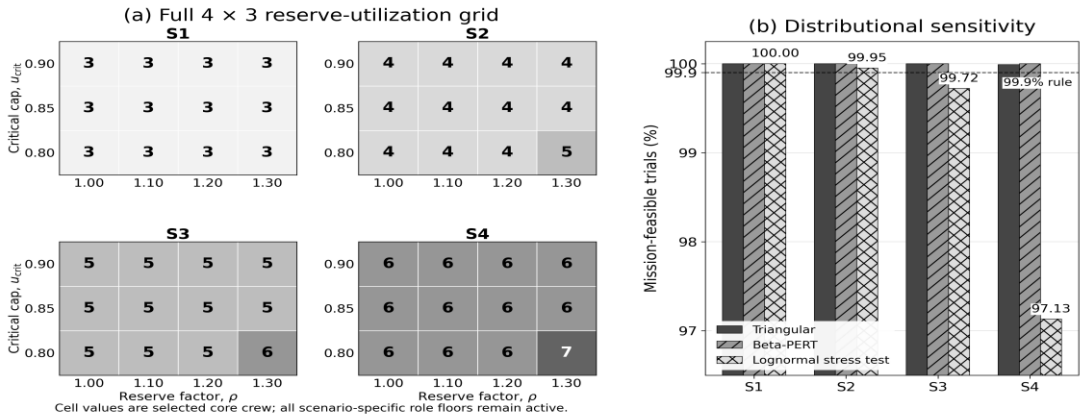

Figure 4. (a) Full reserve-by-utilization grid with scenario-specific role floors retained; (b) distributional sensitivity of selected teams. The dashed line marks the prespecified 99.9% point-estimate rule.

3.6. Role-based competence and assessment architecture

Mathematical role capacity is valid only when supported by demonstrable competence. Table 8 converts the role model into training and assessment evidence. A maritime certificate or pilot licence is prior learning, not automatic WIG equivalence. Assessment must cover the actual control philosophy, propulsion and electrical systems, approved envelope, emergency procedures and transfer-of-control arrangements. Low-frequency abnormal manoeuvres and prolonged-survival leadership require recurrent checking because nominal qualifications do not prevent skill decay.

Table 8. Proposed role-based competence and validation architecture

Role	Primary accountability	Minimum competence evidence	Critical validation scenarios
MP	Overall command; primary control; go/no-go; diversion or ditching	Advanced WIG type qualification; COLREG; meteorology; command CRM; emergency command	Full-mission check; unstable approach; degraded control; emergency command
FO	Independent monitoring; control takeover; threat-and-error management	WIG co-pilot qualification; monitoring; BRM/CRM; abnormal recovery	Takeover, incapacitation, rapid manual reversion and rejected transition
NCO	Lookout; route; collision avoidance; radio and traffic coordination	Advanced marine navigation; COLREG; marine and aviation communications; WIG trajectory awareness	Congested traffic, restricted visibility, radio saturation and navigation disagreement
TEO	Propulsion, electrical, control and automation supervision; fault isolation	Craft-systems type course; electrical safety; software/cyber awareness; fault management	Propulsion or control fault while the command team maintains trajectory
SEO	Occupant safety; fire; evacuation; survival; first response	Passenger safety; crowd/crisis management; firefighting; survival; first aid; survival-craft leadership	Smoke, hard alighting, water evacuation, casualty triage, communications loss and prolonged survival until rescue

4. DISCUSSION

4.1. The revised evidence changes the crew baseline

The most consequential limitation for the headline crew numbers is that phase loads, competence coefficients, reserve factors and utilization caps were authored within the same model and have not been calibrated against operational WIG data. The internal checks establish implementation consistency and robustness under stated perturbations; they do not establish external truth. Within that boundary, the reconstructed model does not retain the earlier conceptual assumption that two technical crew may be adequate for tightly restricted Type A operation once phase-specific allocation, a 20% reserve, technical-fault separation and role-removal survival are imposed. The specified S1 model selects three, and the designated two-person comparator failed 87.39% of triangular trials, almost entirely during transition.

Crew complement is therefore an output of the specified functional structure rather than a free assumption. Two people can appear adequate when control, navigation and technical duties are assessed separately, but become inadequate when those functions peak together and the second person must remain capable of immediate takeover. The framework converts that human-factors concern into a falsifiable allocation test, while leaving the underlying functional structure open to external challenge and recalibration.

4.2. Transition and prolonged emergency response define the safety boundary

Transition is brief, but duration is not a proxy for workload. The team must continue maritime lookout and collision-avoidance reasoning while taking on aeronautical TEM, power management, pitch/height control, surface scanning, technical monitoring and occupant restraint within a narrow correction margin. The S1 and S2 concentration of reduced-team failures in transition, together with the shift toward emergency failures in S3 and S4, supports the conclusion that average cruise workload is an inadequate basis for crew reduction.

Emergency demand becomes dominant with mission complexity. S3 and S4 require stabilization, fault diagnosis, navigation, distress communication, selection of an alighting/diversion option and occupant protection. Remote archipelagic operation is discussed here only as a potential application context; the Arafura Sea example is an illustrative regulatory anchor, not a completed craft-specific case study or external validation. In such service, external rescue cannot be assumed to be immediate. The SEO's competence should extend beyond evacuation to passenger accountability, casualty triage, survival-craft command, exposure and ration management, communications watches and repeated liaison until rescue. For Indonesian approval, the operator should document a route-specific rescue-time and communications assumption in coordination with BASARNAS; the current model represents initiation of this survival phase, not its full duration.

4.3. Automation receives conditional manning credit

Automation was represented as a modifier of demand and capacity, not as a replacement for responsibility. An autopilot may reduce manual control in steady cruise, but it does not eliminate monitoring, mode awareness, failure detection or takeover. Manning credit is defensible only when reliable detection, comprehensible alerts, manual reversion and recovery within available time are demonstrated (Parasuraman et al., 2000). Integrated software may also create common-cause failures that increase technical and control demand simultaneously.

Environmental volatility sharpens this constraint in Indonesia. The Badan Meteorologi, Klimatologi, dan Geofisika (BMKG) marine-warning service reports significant-wave categories and synoptic conditions, and the specifically dated 3 September 2026 bulletin cited here warns that cumulonimbus can generate strong winds and increase wave height (Badan Meteorologi, Klimatologi, dan Geofisika, 2026). For a WIG craft, a rapidly developing squall, wave-height increase or degraded visual reference can compress the available time for manual reversion and may force an unplanned change in ground-effect height or operating mode. Approval evidence should therefore define route-specific weather minima, forecast-update intervals, onboard alerting, reversion-time criteria and recurrent manual-control checks rather than awarding generic automation credit.

Remote support was not counted as immediate onboard capacity. Shore personnel can improve weather information, planning, maintenance advice and emergency coordination, but communication latency, connectivity loss, incomplete sensory information and uncertain authority limit their ability to replace a person at the control station (Man et al., 2015). Any future remote-support credit should be tested as a separate architecture rather than subtracted informally from the onboard complement.

4.4. Incapacitation and human-reliability interpretation

The role-removal test verifies retained functional coverage after a specified role is removed; it does not calculate the probability that incapacitation will occur or cause mission failure. SPAR-H is cited only as a qualitative analogy, not as the method used. The second-round stochastic rerun removed every distinct present role and produced equal or near-equal feasibility within each scenario. This finding rejects the concern that the designated cut was uniquely easy to fail, but it also reveals a limitation: the aggregate six-function cut, rather than a role-specific edge, dominated these scalar-network failures. A WIG-specific HRA and dimension-sensitive task model should therefore be calibrated with observed response times, errors and resource conflicts from simulator trials.

4.5. Indonesian regulatory implementation

The institutional sequence proposed below is the authors' recommended WIG safety-case workflow; it is not presented as an existing, formally mandated Indonesian WIG certification procedure. Under this proposal, the Directorate General of Sea Transportation (Direktorat Jenderal Perhubungan Laut/Hubla) would use the phase-function task-load matrix as a safety-case annex rather than determine WIG manning primarily from static vessel descriptors. The proposal complements, rather than replaces, the functional principles in IMO Resolution A.1047(27). Biro Klasifikasi Indonesia (BKI) could support technical verification of system reliability, automation fallback and class-related assumptions, while the Administration would retain responsibility for manning approval; BASARNAS would inform route-specific rescue and communications planning.

A practical approval sequence is: (1) the applicant submits a craft- and route-specific phase-function matrix, VACP conflict screen and role-competence evidence; (2) Hubla verifies the operational envelope, duty/rest pattern, reserve assumptions and role separation; (3) BKI or another recognized technical assessor verifies automation, fault detection, manual reversion and common-cause protections; (4) the applicant demonstrates selected and degraded complements in human-in-the-loop scenarios, including incapacitation and prolonged emergency management; and (5) approval contains change triggers for route, passenger load, software, propulsion, weather limits or rescue coverage. Gross tonnage, propulsion power and passenger count may remain screening inputs, but they should not substitute for phase-specific concurrency evidence.

4.6. Claim boundaries and external validation

The added analyses sharpen the claim boundary. The 3-6 sequence was stable over most reserve-utilization combinations, all ±20% supporting-coefficient perturbations and all bounded triangular/PERT selected-team runs. Across ten independent seeds, every S3 lognormal run remained below the 99.9% rule (99.68%-99.84%) and every S4 run remained materially below it (96.94%-97.25%), while S1 and S2 passed in all replications. The mechanism is not a larger number of modeled phases, because all scenarios use six. Instead, S3 and especially S4 start with higher concurrent transition/emergency loads and thinner robust reserve ratios; correlated scenario, phase and function multipliers then compound against a fatigue-reduced total-capacity cut. This explains why the long tail exposes S3/S4 but not S1/S2. The numerical values remain conditional on the published model, and the triangular distribution is transparent rather than proven conservative.

The person-equivalent and max-flow formulation remains scalar. It is exact for the stated network but may overstate substitution where tasks compete for the same visual, auditory, cognitive or psychomotor resource. A future multi-resource mixed-integer model could constrain each resource dimension k and add a linearized penalty for co-assigned, mutually interfering functions. In Equation (7), b_pfk maps functional demand into resource k, c_rk is role-specific resource capacity, γ_fgk is the within-resource interference penalty, and binary z_rpfg identifies concurrent co-assignment. This formulation is illustrative; its coefficients require empirical calibration before use.

$$\Sigma f\, b_{p\ fk}\, y_{rp\ f} + \Sigma f<g\, \gamma fgk\, z_{rp\ fg} \le u_{p\ k}\, c_r k\, n_r,\ k \in \{V, A, C, P\} \qquad (7)$$

External validation should proceed in two stages. An indicative Delphi design would recruit 18-30 experts, balanced across WIG design, test piloting, master mariner practice, human factors, regulation and search-and-rescue, for at least two rounds with predefined consensus criteria (Beiderbeck et al., 2021; Diamond et al., 2014). A counterbalanced human-in-the-loop study should then target at least 12 independent crew sessions per scenario-complement condition, with the final sample determined by prospective precision or power analysis. Scenarios should include transition failure, propulsion fault, pilot incapacitation, navigation saturation, hard alighting and evacuation. Objective measures should include trajectory deviation, response time, missed alerts, communication accuracy, checklist completion and recovery success; heart-rate variability, electrodermal activity, eye-tracking fixation/dwell and pupil measures, and SAGAT probes should supplement NASA-TLX and behavioural observations (Endsley, 2021; Hart & Staveland, 1988; Liu et al., 2023). Supplementary Material S2 provides the protocol.

These limitations define the contribution's proper use. In descending order of consequence for the headline 3-6-person numbers, the principal limitations are: (1) absence of empirical calibration and external validation of loads and competence coefficients; (2) scalar, fungible workload representation; and (3) uncertainty about the operational demand-tail distribution. The selected teams are the smallest compositions satisfying the published model, reserve, functional-removal screen and stochastic checks. They are candidates for craft-specific external testing, not legal minima.

5. CONCLUSION

This study frames WIG safe manning as a phase-specific, competence-constrained allocation problem. Under the baseline 20% reserve and published utilization caps, the computationally verified complements are three persons for restricted Type A, four for coastal Type A passenger service, five for demanding Type B and six for remote Type C operation. Testing every distinct

present-role removal produced no more favorable N-1 comparator than the designated cut. Transition and emergency response, not average cruise, set the modeled safety boundary.

The reserve-utilization analysis retained the baseline sizes in 45 of 48 cells; only the joint 30% reserve/80% critical-cap setting added one person in S2-S4. The sequence also persisted under ±20% perturbation of supporting competence coefficients. Ten-seed replication preserved the original inference: bounded triangular/PERT runs passed, whereas every lognormal-tail replication failed the 99.9% rule for S3 (99.68%-99.84%) and S4 (96.94%-97.25%). The single primary-seed triangular S4 failure and the 287 lognormal failures were aggregate all-function capacity deficits, concentrated first in emergency, approach and transition.

For Indonesia, the authors propose that a WIG manning safety case demonstrate: (1) concurrent dual COLREG/TEM workload management; (2) qualified manual reversion under rapidly changing BMKG weather; (3) retained functional coverage after any single present-role removal, without interpreting the test as an incapacitation probability; and (4) SEO capability for prolonged survival, not merely immediate evacuation. Hubla would retain approval authority, with BKI supporting technical assurance and BASARNAS informing remote-route rescue assumptions. Operators whose craft, route, passenger load or operating conditions fall outside S1-S4 must rerun the complete phase-function matrix and validation sequence; they should not interpolate between Table 4 rows.

The numerical outputs remain internally consistent engineering baselines. Their most consequential limitation is the absence of empirical calibration and external validation of the authored loads and competence coefficients. Dimension-sensitive workload measurement, human-in-the-loop simulation, on-craft drills and evacuation/survival evidence are therefore required before regulatory adoption. The contribution is a reproducible decision architecture that makes assumptions and failure modes auditable.

6. REFERENCES

Ahvenjärvi, S. (2016). The human element in autonomous shipping. *Journal of Marine Science and Engineering*, *4*(4), 78. https://doi.org/10.3390/jmse4040078

Badan Meteorologi, Klimatologi, dan Geofisika. (2026). *Prakiraan harian tinggi gelombang 7 hari ke depan*: 04 September 2026 s/d 10 September 2026 [Seven-day daily wave-height forecast]. Updated September 3, 2026. https://maritim.bmkg.go.id/marine2026-data/doc/gelombang_7hari_kedepan.pdf?t=1771632000128

Beiderbeck, D., Frevel, N., von der Gracht, H. A., Schmidt, S. L., & Schweitzer, V. M. (2021). Preparing, conducting, and analyzing Delphi surveys: Cross-disciplinary practices, new directions, and advancements. *MethodsX*, *8*, 101401. https://doi.org/10.1016/j.mex.2021.101401

de Divitiis, N. (2005). Performance and stability of a winged vehicle in ground effect. *Journal of Aircraft*, *42*(1), 148-157. https://doi.org/10.2514/1.4830

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M., & Wales, P. W. (2014). Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology*, *67*(4), 401-409. https://doi.org/10.1016/j.jclinepi.2013.12.002

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32-64. https://doi.org/10.1518/001872095779049543

Endsley, M. R. (2021). A systematic review and meta-analysis of direct objective measures of situation awareness: A comparison of SAGAT and SPAM. *Human Factors*, *63*(1), 124-150. https://doi.org/10.1177/0018720819875376

Felski, A., & Zwolak, K. (2020). The ocean-going autonomous ship—Challenges and threats. *Journal of Marine Science and Engineering*, *8*(1), 41. https://doi.org/10.3390/jmse8010041

Gertman, D., Blackman, H., Marble, J., Byers, J., & Smith, C. (2005). *The SPAR-H human reliability analysis method* (NUREG/CR-6883). U.S. Nuclear Regulatory Commission.

Government of Indonesia. (2014). *Law of the Republic of Indonesia Number 29 of 2014 concerning Search and Rescue* [Undang-Undang Republik Indonesia Nomor 29 Tahun 2014 tentang Pencarian dan Pertolongan]. https://peraturan.bpk.go.id/Details/38691/uu-no-29-tahun-2014

Halloran, M., & O'Meara, S. (1999). *Wing in ground effect craft review* (Report No. DSTO-GD-0201). Defence Science and Technology Organisation.

Hart, S. G., & Staveland, L. E. (1988). *Development of NASA-TLX: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati* (Eds.), Human mental workload (pp. 139-183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

International Civil Aviation Organization. (2018). *Safety management manual* (Doc 9859, 4th ed.). ICAO.

International Civil Aviation Organization. (2022a). *Annex 1 to the Convention on International Civil Aviation: Personnel licensing*. ICAO.

International Civil Aviation Organization. (2022b). *Annex 6 to the Convention on International Civil Aviation, Part I: International commercial air transport—Aeroplanes*. ICAO.

International Electrotechnical Commission. (2019). *IEC 31010:2019 risk management—Risk assessment techniques*. IEC.

International Labour Organization. (2022). *Maritime Labour Convention, 2006, as amended*. ILO.

International Maritime Organization. (1972). *Convention on the International Regulations for Preventing Collisions at Sea*. IMO.

International Maritime Organization. (1974). *International Convention for the Safety of Life at Sea, as amended*. IMO.

International Maritime Organization. (2002). *Interim guidelines for wing-in-ground craft* (MSC/Circ.1054). IMO.

International Maritime Organization. (2005). *General principles and recommendations for knowledge, skills and training for officers on wing-in-ground craft* (MSC/Circ.1162). IMO.

International Maritime Organization. (2011). *Principles of minimum safe manning* (Resolution A.1047(27)). IMO.

International Maritime Organization. (2018). *Revised guidelines for wing-in-ground craft* (MSC.1/Circ.1592). IMO.

International Maritime Organization. (2020). *International Code of Safety for High-Speed Craft, 2000: 2020 edition*. IMO.

International Maritime Organization. (2021). *Outcome of the regulatory scoping exercise for the use of maritime autonomous surface ships* (MSC.1/Circ.1638). IMO.

International Maritime Organization. (2022). *STCW including 2010 Manila amendments: STCW Convention and STCW Code*. IMO.

International Organization for Standardization. (2018). *ISO 31000:2018 risk management—Guidelines*. ISO.

Kerem, K., Carjova, K., & Tapaninen, U. P. (2025). Success factors in commercialization of wing-in-ground crafts as means of maritime transport: A case study. *Future Transportation*, *5*(1), 13. https://doi.org/10.3390/futuretransp5010013

Li, X., & Yuen, K. F. (2024). A human-centred review on maritime autonomous surfaces ships: Impacts, responses, and future directions. *Transport Reviews*, *44*(4), 791-810. https://doi.org/10.1080/01441647.2024.2325453

Liu, Y. C., Figalova, N., Pichen, J., Hock, P., Baumann, M., & Bengler, K. (2023). *Workload assessment of human-machine interface: A simulator study with psychophysiological measures*. AHFE International. https://doi.org/10.54941/ahfe1004172

Man, Y., Lundh, M., Porathe, T., & MacKinnon, S. N. (2015). From desk to field—Human factor issues in remote monitoring and controlling of autonomous unmanned vessels. *Procedia Manufacturing*, *3*, 2674-2681. https://doi.org/10.1016/j.promfg.2015.07.635

Nebylov, A. V. (2002). *Ekranoplanes: Controlled flight close to the sea*. WIT Press.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, *30*(3), 286-297. https://doi.org/10.1109/3468.844354

Pathak, S. K., & Bhardwaj, S. (2024). Safe manning: Workload assessment of deck officers. *Journal of Maritime Research*, *21*(1), 106-113.

Porathe, T., Prison, J., & Man, Y. (2014). *Situation awareness in remote control centres for unmanned ships. In Proceedings of the Human Factors in Ship Design and Operation Conference*. The Royal Institution of Naval Architects.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *13*(3), 257-266. https://doi.org/10.1109/TSMC.1983.6313160

Reason, J. (1990). *Human error*. Cambridge University Press.

Rozhdestvensky, K. V. (2006). Wing-in-ground effect vehicles. *Progress in Aerospace Sciences*, *42*(3), 211-283. https://doi.org/10.1016/j.paerosci.2006.10.001

Salas, E., Wilson, K. A., Burke, C. S., & Wightman, D. C. (2006). Does crew resource management training work? An update, an extension, and some critical needs. *Human Factors*, *48*(2), 392-412. https://doi.org/10.1518/001872006777724444

Sawada, R., Miyauchi, Y., Wada, S., Tanigushi, T., Hamada, S., Koike, H., Wakita, K., & Maki, A. (2024). *Perspective on the marine simulator for autonomous vessel development* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2407.19673

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159-177. https://doi.org/10.1080/14639220210123806

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*(3), 449-455. https://doi.org/10.1518/001872008X288394

Wróbel, K., Montewka, J., & Kujala, P. (2017). Towards the assessment of potential impact of unmanned vessels on maritime transportation safety. *Reliability Engineering & System Safety*, *165*, 155-169. https://doi.org/10.1016/j.ress.2017.03.029

Yun, L., Bliault, A., & Doo, J. (2010). *WIG craft and ekranoplan: Ground effect craft technology*. Springer.