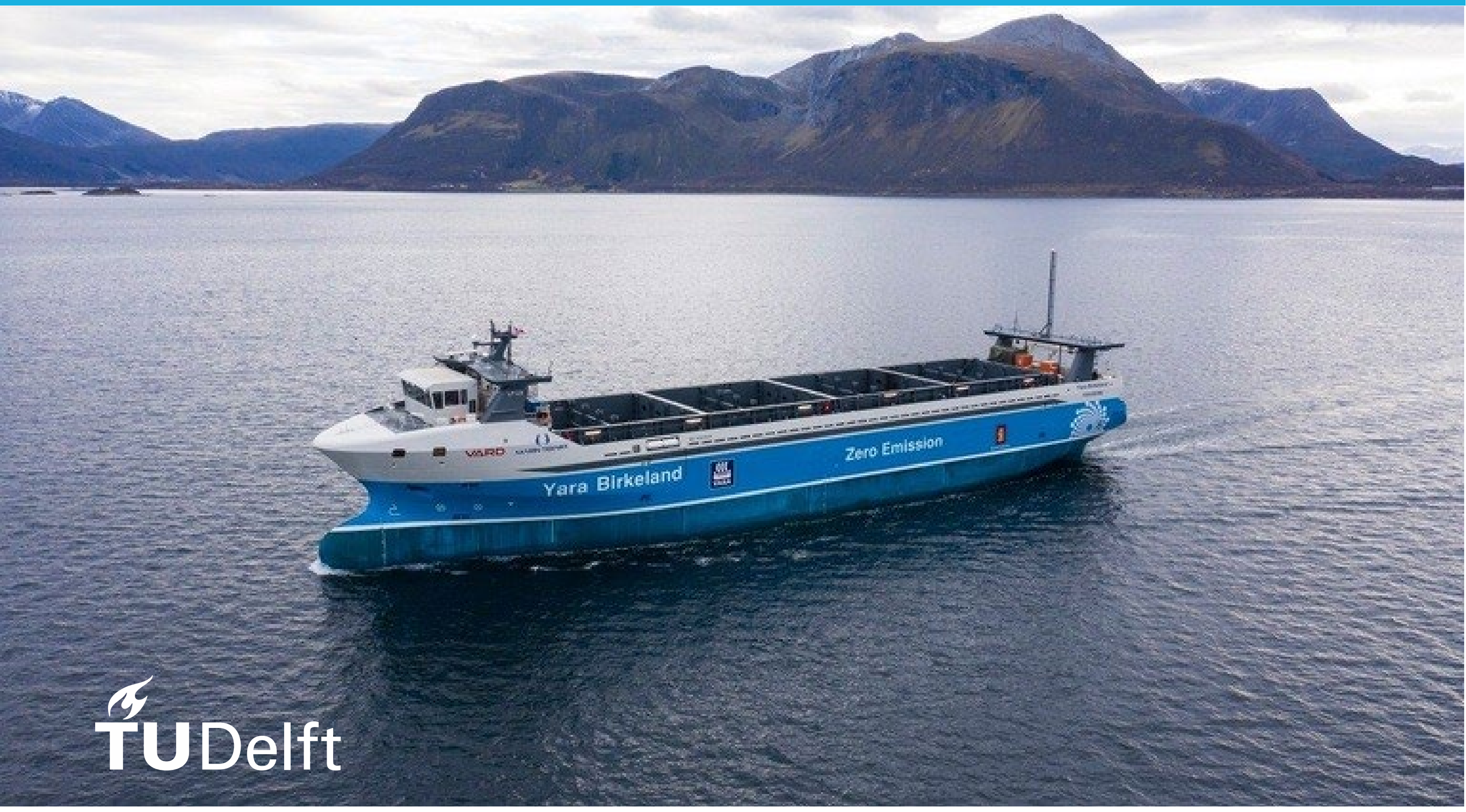
Brandon Daver

Risk-aware control for the fail-safe navigation of Autonomous Surface Vessels under thruster faults



https://sailbuoy.no/news/uncrewed-surface-vessels-usv

Risk-aware control for the fail-safe navigation of Autonomous Surface Vessels under thruster faults

By

B. Daver

Master Thesis

in partial fulfilment of the requirements for the degree of

Master of Science
in Mechanical Engineering

at the Department Maritime and Transport Technology of Faculty Mechanical, Maritime and Materials Engineering of Delft University of Technology
to be defended publicly on Tuesday June 16, 2026 at 10:00 AM

Student number:	4878183	
MSc track:	Multi-Machine Engineering	
Report number:	2026.MME.9171	
Supervisors:	Dr. V. Reppa A. Dhyani	
Thesis committee:	Dr. V. Reppa,	TU Delft committee Chair, Maritime And Transport Technology
	Prof.dr. R.R. Negenborn,	TU Delft committee member, Maritime and Transport Technology
	A. Dhyani,	TU Delft Supervisor, Maritime and Transport Technology
Date:	May 20, 2026	

List of Abbreviations

Table 1: List of abbreviations.

Abbreviation	Full Form
ASV	Autonomous Surface Vehicle
MPC	Model Predictive Control
DOF	Degree of Freedom
ODE	Ordinary Differential Equation
KPI	Key Performance Indicator
RK4	Fourth Order Runge-Kutta
MT	Multi-Trajectory
ST	Single-Trajectory
CTE	Cross-Track Error

List of Symbols

Table 2: Summary of key symbols used in this work.

Symbol	Description
State and Kinematic Symbols	
$\boldsymbol{x}$	Complete state vector $[\eta^T, \nu^T]^T$
$\boldsymbol{\eta}$	Earth-fixed position and heading vector $[x, y, \psi]^T$
$\boldsymbol{\nu}$	Body-fixed velocity vector $[u, v, r]^T$
$\boldsymbol{J}(\psi)$	Jacobian rotation matrix from body to earth frame
Dynamic and Force Symbols	
$\boldsymbol{M}$	Total mass matrix (Rigid-body + Added Mass)
$\boldsymbol{D}$	Damping matrix
$\boldsymbol{\tau}$	Control force and moment vector from thrusters
$\boldsymbol{d}$	External disturbances
$\boldsymbol{\tau}_{wind}$	Wind force and moment vector
l_x	Lever arm for rudder
V_w, ψ_w, γ_w	Relative wind velocity, wind direction, relative wind angle
c_x, c_y, c_n	Wind force coefficients
A_{fw}, A_{lw}	Frontal/lateral wind area
L_{oa}	Vessel length
ρ_a	Air density
m	Rigid body mass of the vessel
I_z	Moment of inertia of the vessel
X_u, Y_v, N_r	Linear damping coefficients
$X_{\dot{u}}$, $Y_{\dot{v}}$, $N_{\dot{r}}$	Added mass coefficients
Control and MPC Symbols	
$\boldsymbol{u}$	Control input vector $[f, a]^T$
k	Prediction horizon step index, $k \in \{0, \ldots, N\}$
t	Simulation time step index, $t \in \{1, \ldots, T\}$
Δt	Time step duration
B	Binary switch variable for nominal/fail-safe mode
$\boldsymbol{r}$, r_x, r_y	Reference functions
$\boldsymbol{q}$	Waypoint position vector
i	Waypoint index, $i \in \{1, \ldots, M\}$
α	Path advancement variable, $\alpha \in [0, L_{total}]$
$\boldsymbol{w}$	Vector of all decision variables
$\boldsymbol{\theta}$	Vector of all parameters
s_{ref}	Reference speed
α_{step}	Advancement rate
$\boldsymbol{\sigma}$	Grounding obstacle vector, $j \in \{1, \ldots, J\}$
Cost Function Terms and Weights Symbols	
$\varphi(\boldsymbol{w}, \boldsymbol{\theta})$	Objective function
$\xi(\cdot)$	Path progression cost term for nominal trajectory
$\epsilon(\cdot)$	Control input cost term
$\rho(\cdot)$	Risk cost term
$\zeta(\cdot)$	Terminal velocity cost function for fail-safe trajectory
$\boldsymbol{\kappa}$	Weight vector for path following terms
$\boldsymbol{\Lambda}$, $\boldsymbol{\Delta}$	Weights for control input terms (magnitude and rate of change)
μ_1, μ_2, Ξ	Weights for risk terms (base, wind and sensitivity)
Ω	Weight for the terminal velocity penalty
Supervisory System Symbols	
E	Solver convergence status
G	Steering and propulsion fault status
$\boldsymbol{P}$	Trajectory position matrix
$\mathcal{S}(\cdot)$	Grounding evaluation function
$v_{terminal}$	Terminal velocity of the fail-safe trajectory
$\Delta\rho$	Risk gradient along the nominal trajectory

Acknowledging use of AI

During the graduation assignment, the author used Google Gemini to assist with the refinement of written text for clarity and coherence. Additionally, the tool was utilized to assist in debugging and optimizing the MATLAB scripts used for the simulations presented in chapter 4, as well as generating the LaTeX code required to generate tables and figures in Overleaf. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the final content of this report.

Abstract

Driven by the need to improve safety, the maritime industry is undergoing a transformative change with the development of autonomous surface vessels, which can aim to reduce the likelihood of accidents by incorporating risk awareness directly into motion planning and control to account for uncertainties in maritime operations. However, existing literature does not consider risk-aware control capabilities that can enable fail-safe actions. To address this limitation, this work proposes a Multi-Trajectory (MT) Risk-Aware Model Predictive Control (MPC) scheme designed to simultaneously optimize two distinct trajectories: a mission-driven nominal trajectory and a fail-safe contingency trajectory. The fail-safe trajectory is computed with a risk cost term in the MPC objective function to guide the vessel towards minimum-risk conditions. In addition, a terminal velocity cost term ensures that the velocity of the vessel can be reduced at each optimization step, to bring it to a near-stop at the end of the fail-safe trajectory. By enforcing a shared initial control input constraint, the system reduces the possibility that the nominal trajectory leads to a state from which an escape is impossible. A supervisory system monitors these trajectories, evaluating them for predicted grounding, the physical capability to bring the vessel to a stop and if the risk is predicted to increase along the nominal trajectory. If a hazardous situation which requires a fail-safe action is detected, defined as a predicted grounding event or a loss of stopping capability while risk is predicted to increase, the supervisory system activates the fail-safe action. The fail-safe action is activated by using the binary variable (B), which accordingly modifies the MT MPC objective function. Verification is performed in a simulation environment by comparing the MT scheme against a Single-Trajectory (ST) MPC scheme during no-fault and stuck rudder and loss of effectiveness fault scenarios in conjunction with drifting forces due to wind. Results demonstrate that the MT scheme is able to able to reduce the likelihood of grounding compared to the ST scheme. Finally, this enhanced safety imposes no penalty on travel time, energy consumption, or path-tracking accuracy during scenarios that do not require a fail-safe action. The only significant trade-off is a two- to three-fold increase in average solver time compared to the ST scheme.

Contents

Chapter 1

Introduction

Driven by the urgent need to improve safety, the maritime industry is undergoing a transformative change with the development of Autonomous Surface Vessels (ASVs) [1]. The recent grounding of the MT Cheng Xian Feng 168 in January 2026 highlights this need. While navigating near Siargao Island, the vessel suffered a sudden loss of propulsion. Despite deploying the anchors, the crew could not intervene in time. In conjunction with strong currents and waves, this inevitably resulted in the grounding of the vessel (see figure 1.1). Grounding occurs when the vessel makes unintentional contact with the seabed, shoreline or underwater obstacles. [2]

Figure 1.1: Grounding of the MT Cheng Xian Feng 168 in January 2026 [2]

This accident highlights a critical limitation in standard navigation: the failure to account for uncertainties and reliance on human assessment [3]. Risk-aware control of ASVs can account for the risk of grounding or collision with other vessels and uncertainties in maritime operations [1]. Furthermore, these risk-aware control systems are capable of reacting instantly based on real-time data rather than relying on human assessment, which can come with a delay. When detecting hazardous situations, such as grounding, these systems can execute a fallback strategy. These hazardous situations can arise from unsafe actions taken during vessel operations, originating from factors such as steering and propulsion faults and harsh weather conditions [4]. Fault detection and isolation is a function of the fault diagnosis system. The appropriate response is dictated by the supervisory system (see figure 3.7), which uses data provided by the fault diagnosis system, classifies the severity of the situation and determines if the vessel can safely continue the operation or if the vessel must cease the operation. If the supervisory system determines that the vessel can safely continue the operation, possibly with reduced functionality, a fail-operational action is executed [5]. A fail-operational action includes for example rerouting, reducing velocity and selecting a fault-tolerant control strategy [1, 4]. In contrast, if the supervisory system determines that it cannot guarantee safe operation due to the severity of the fault or environmental conditions, it must execute a fail-safe action. In this scenario, the focus shifts completely from following the mission to guiding the vessel towards a safe state and cease the operation. A fail-safe action includes for example moving away from obstacles [4].

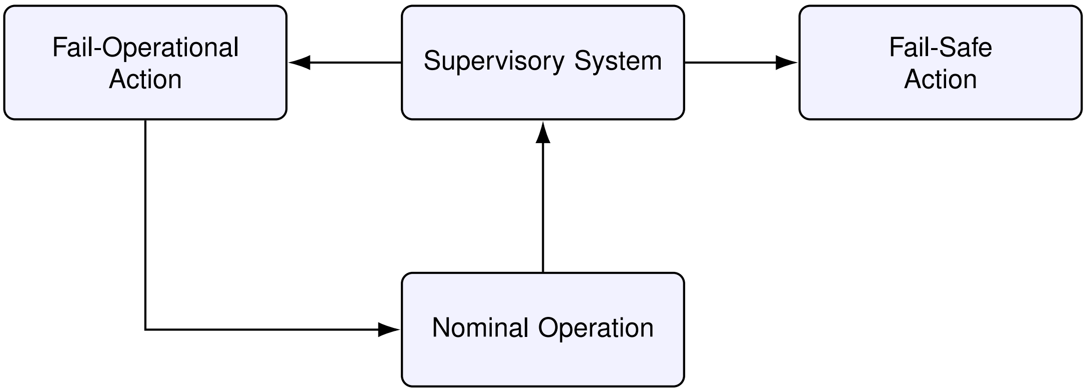


Figure 1.2: Approach for managing the fallback strategy.

In the case of the grounding of the MT Cheng Xian Feng 168, using risk-aware control, which anticipates the non-zero probability of propulsion loss, could have identified the high risk of grounding for the intended trajectory, given the prevailing onshore currents. Instead of operating in a high-risk area, actions could have been taken, such as rerouting the vessel or holding position until the currents have changed (see figure 1.3). In case of a propulsion loss, the vessel would not have grounded.

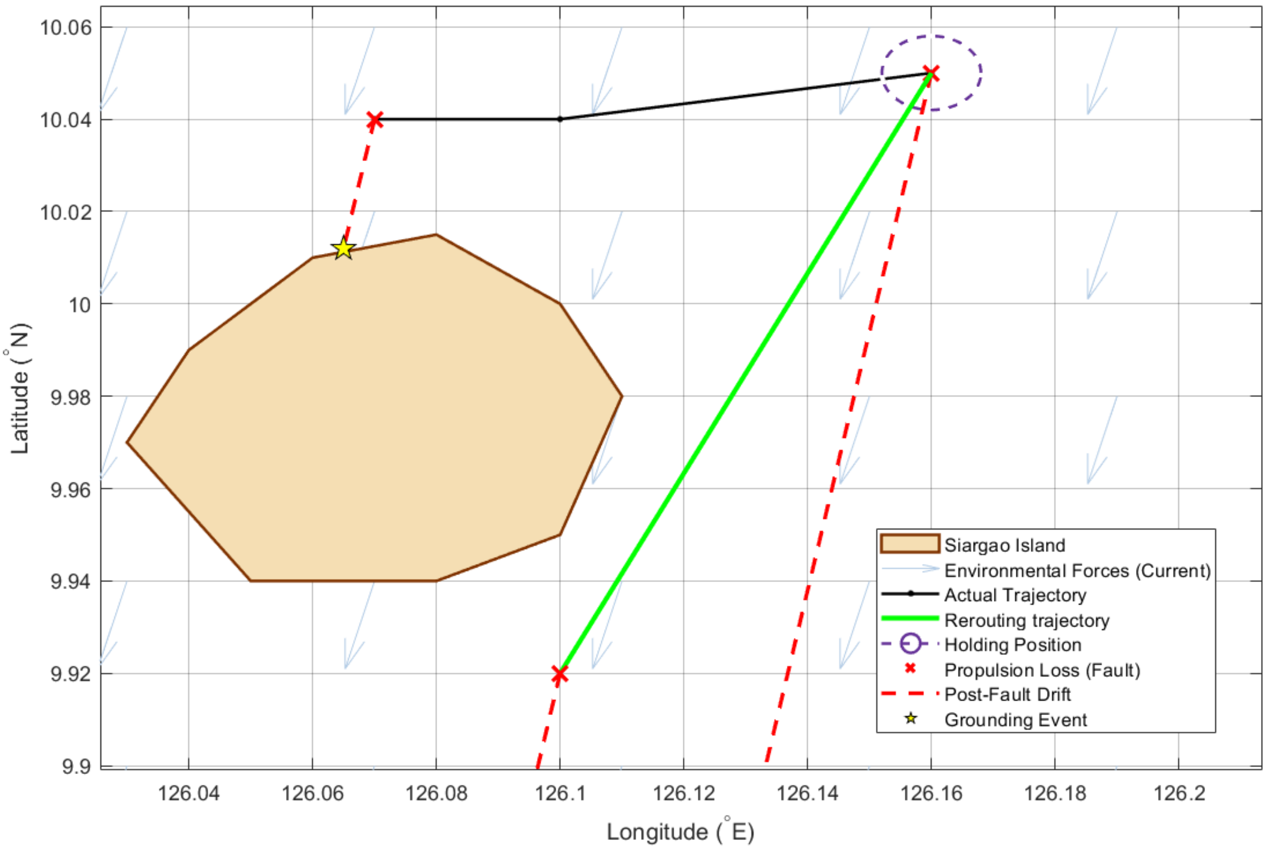


Figure 1.3: Potential rerouting and holding position of the MT Cheng Xian Feng 168

1.1 Related Work

Previous work into risk-aware control for ASVs reveal the current state-of-the-art. In [6], the Multi-objective Particle Swarm Optimization scheme is proposed, which utilizes dynamic safety domains to assess the risk of grounding in coastal waters. The dynamic safety domain accounts for steering and propulsion faults. The risk is defined by the overlap of the safety domain with grounding obstacle polygons. In this work, the optimal, goal-directed trajectory is determined, while considering safety domain violation. However, it lacks fallback strategies in the event of a hazardous situation. In [7], a Model Predictive Control (MPC) scheme is proposed where the risk is modeled using an exponential term in the MPC objective function to account for the vulnerability to steering and propulsion faults. The risk is dependent on the proximity to grounding obstacles and drifting forces due to wind. In this work, the goal-directed trajectory is optimized, but it lacks fallback strategies in the event of a hazardous situation. Also in [1], an exponential risk term is included in the objective function of the MPC to account for the vulnerability to steering and propulsion faults. The risk is dependent on the proximity to grounding obstacles and drifting forces due to wind. Besides, the objective function includes terms for mission progression and control effort. This work considers fail-operational actions such as rerouting and holding position along the reference path to wait for the situation to improve. However, no fail-safe actions are included. In [8], a Multi-Trajectory (MT) active fault-tolerant control scheme is proposed, specifically the Contingency-Robust-Adaptive Model Predictive Control approach. The primary advantage of MT optimization is the dual-plan strategy, which simultaneously optimizes a primary trajectory and a contingency trajectory. The primary trajectory is adaptive, utilizing real-time fault diagnosis results to optimize the mission-driven trajectory. In contrast, the contingency trajectory optimizes the mission-driven trajectory for the worst-case single fault realization considered in this work. This work defines risk using tubes around obstacles to account for disturbances and uncertainties. The vessel should operate outside these tubes. To guarantee a smooth transition from the healthy to the faulty model without infeasibilities, the initial control input for both trajectories is identical. By enforcing this constraint, the system can safely transition to a fault-tolerant control strategy, which is a fail-operational action, even when an unexpected fault occurs. However, no fail-safe actions are included.

The gap in the state-of-the-art is that existing literature does not consider risk-aware control capabilities that can enable fail-safe actions. Single-trajectory risk-aware control schemes are not well suited for this task, as it balances safety and progress in the optimization of a single trajectory. Even if the scheme was designed to not consider the mission progression objective when needed, it remains reactive. As it is uncertain if a feasible fail-safe action exists until there is an attempt to find one. Furthermore, simply relying on an anchor drop, is insufficient. The vessel does not stop immediately after dropping the anchor. The exact time and distance required to fully stop the vessel are uncertain [2]. These uncertainties creates a blind spot where the vessel may unknowingly enter a state of no return. A MT MPC scheme provides a solution to this problem. By optimizing the mission-driven and fail-safe trajectory simultaneously under the constraint that the initial control input is identical, the optimizer tries to answer: *What is the best action the vessel can take right now that not only guides the vessel along the reference path, but also provides the ability to execute a fail-safe action in the next optimization time-step?* This approach reduces the possibility that the nominal trajectory leads to a state from which an escape is impossible, since the previous fail-safe action is executed if no new, safe solution can be found [8, 9]. Simultaneously, the optimization of the fail-safe trajectory provides the ability to check if a valid fail-safe action exists.

1.2 Aim and Research Questions

To address the identified gap, this work proposes a Multi-Trajectory Risk-Aware Model Predictive Control scheme. This scheme optimizes the mission-driven nominal trajectory and fail-safe trajectory simultaneously. Integrated into this scheme is a supervisory system designed to identify hazardous situations that require a fail-safe action and manages the activation of the fail-safe action. This work will consider hazardous situations that arise from unsafe actions taken during vessel operations originating from steering and propulsion faults and drifting forces due to wind. In this work, the following research question will be addressed:

How can a risk-aware control scheme be designed for fail-safe actions of ASVs during steering and propulsion faults?

This question is supported by the following sub-questions:

1. Which steering and propulsion fault scenarios create hazardous situations that require a fail-safe action?
2. How to design a risk-aware MPC in which both the nominal and fail-safe trajectories are considered?
3. How can the performance of the proposed Multi-Trajectory Risk-Aware MPC be verified through simulations and key performance indicators?

The sub-questions are structured to systematically address the main research question by following a logical progression from identifying the fault scenarios to the design of the control scheme and, finally, verification. Sub-question 1 lays the foundation by identifying the specific steering and propulsion fault scenarios that can escalate into hazardous situations that require a fail-safe action. Sub-question 2 addresses the design of the Multi-Trajectory Risk-Aware Model Predictive Control scheme capable of simultaneously optimizing a mission-driven nominal trajectory and a fail-safe trajectory. Finally, Sub-question 3 serves as the verification phase, using simulations to verify effectiveness of the Multi-Trajectory Risk-Aware Model Predictive Control scheme (designed in sub-question 2) in hazardous situations that requires a fail-safe action (identified in sub-question 1) based on key performance metrics. Collectively, the answers to these sub-questions provide the steps required to answer the main research question.

1.3 Contribution

The key *contribution* of this work is the design and verification of a Multi-Trajectory Risk-Aware Model Predictive Control scheme that enables fail-safe capabilities of ASVs. This scheme enables guiding the vessel along the reference path, but also provides the ability to execute a fail-safe action in the next optimization time-step. This reduces the possibility that the nominal trajectory leads to a state from which an escape is impossible, since the previous fail-safe action is executed if no new, safe solution can be found. This scheme includes a supervisory system that is able to identify hazardous situations that require a fail-safe action and manage the transition to the fail-safe action.

1.4 Scope Limitations

This work focuses on the design and simulation-based verification of a Multi-Trajectory Risk-Aware Model Predictive Control scheme designed to enable fail-safe actions for ASVs. For the sake of limiting complexity to a manageable degree, some research scope limitations are present. The scope limitations can be summarized as follows:

1. Fault detection and isolation is not considered. It is assumed that the exact onset, type, and magnitude of any steering and propulsion fault is instantaneously and perfectly known.
2. Sensors and measurement filters are excluded. While a real-world deployment would require sensors to capture measurements and filters to reject noise to provide reliable data, the design of these systems are not considered in this work.
3. Sensor failures, communication delays, and cyber-attacks are not considered. This work assumes an uninterrupted, instantaneous flow of data and control inputs.
4. The ship model does not account for ocean currents, the Coriolis-centripetal forces, wave, ballast, buoyancy and gravitational forces. Furthermore, six degrees of freedom (DOF) motions (roll, pitch, and heave) are not considered. The mathematical ship model is a 3-DOF model (surge, sway, and yaw) subjected solely to wind-induced drifting forces.
5. Dynamic obstacles fall outside the scope of this work. Consequently, complex interactions with moving vessels, including compliance with dynamic collision avoidance rules (e.g., COLREGs), are excluded.

1.5 Approach

The research question and the sub-questions are addressed throughout this work and are addressed through literature research, model-based design and simulation-based verification. First, chapter 2 addresses the steering and propulsion fault scenarios (sub-question 1) that require a fail-safe action through literature research. Chapter 3 addresses the design of the Multi-Trajectory Risk-Aware Model Predictive Control scheme (sub-question 2) by formulating the 3-DOF non-linear dynamic model of the vessel as an Ordinary Differential Equation (ODE) based on literature research. Then the continuous-time ODE is discretized into a finite-dimensional problem using the Fourth Order Runge Kutta (RK4) method. Then, the objective function and constraints for both the nominal and fail-safe trajectory are established. Finally, the supervisory system used to manage the activation of the fail-safe action is formulated. Chapter 4 is dedicated to the verification of the Multi-Trajectory Risk-Aware Model Predictive Control scheme (sub-question 3) by comparing the performance against the Single-Trajectory (ST) MPC scheme (see [1]), which has no fail-safe capabilities, in a simulation environment. The performances are compared across varying scenarios using Key Performance Indicators (KPIs). Finally, chapter 5 focuses on answering the main research question and provides directions for future work. The expected output of this work is a fully formulated and verified Multi-Trajectory Risk-Aware Model Predictive Control scheme.

Chapter 2

Steering and Propulsion Fault Scenarios

The ability to maintain control over the vessel is crucial to maintain safety. This capability is typically achieved via a set of rudders and propellers. A fault in the steering and propulsion system is a critical event, but the fault itself may not necessarily lead to an accident. The true danger arises when the fault happens in conjunction with other factors such as external environmental forces, like wind [1, 2]. Based on literature review a clear understanding of steering and propulsion fault scenarios and how they are handled will be established. This chapter addresses the first research sub-question: *What steering and propulsion fault scenarios create hazardous situations that require a fail-safe action?*

2.1 Steering and Propulsion Faults

Steering and propulsion faults can be classified into four categories, each having distinct impacts on the control capabilities of the vessel [10, 11]. The first category is loss of effectiveness (LoE), where the propulsion system delivers only a fraction of the commanded input [8, 12, 13]. In this scenario, for example, the propeller might generate only 50% of the expected thrust for a given revolutions per minute due to, for instance, propeller blade damage. The second type is outage, which represents the complete failure of the steering and propulsion system, rendering them unresponsive [1, 13]. The third type is bias, in which the steering and propulsion system produces an unknown, additive force or moment even when zero output is commanded [13]. The fourth category is a stuck fault, where the steering system becomes jammed regardless of the control input [1]. An example is a rudder mechanically locked at a 15 degree angle or an azimuth thruster losing the ability to rotate.

2.2 Fault Scenarios and Fault Handling

In [13], the loss of effectiveness, bias, and outage faults are considered within the dynamic model of the vessel, which is governed by $\boldsymbol{M}\dot{\boldsymbol{\nu}} + \boldsymbol{D}\boldsymbol{\nu} = \boldsymbol{\tau} + \boldsymbol{\tau}_d$. Here, $\boldsymbol{M}$ and $\boldsymbol{D}$ denote the inertia and damping matrices, respectively. The term $\boldsymbol{\nu}$ represents the velocities of the vessel, and $\boldsymbol{\tau}_d$ represents the unknown ocean disturbances. Mathematically, the actual control forces and moments ($\boldsymbol{\tau}$) generated by the steering and propulsion system are modeled as $\boldsymbol{\tau} = \boldsymbol{T}(\boldsymbol{K}\boldsymbol{u}_c + \boldsymbol{d})$. In this formulation, $\boldsymbol{T}$ denotes the thruster configuration matrix and $\boldsymbol{u}_c$ represents the commanded control inputs. The matrix $\boldsymbol{K} = \text{diag}[k_1, k_2, \dots, k_r]$ is an effectiveness matrix where each diagonal element $k_i \in [0, 1]$ quantifies the health of the i-th thruster. Specifically, $k_i = 1$ indicates a healthy thruster, $0 < k_i < 1$ denotes a loss of effectiveness, and $k_i = 0$ represents a complete outage. Furthermore, $\boldsymbol{d}$ represents the bias fault (an unknown additive force) for each thruster. These faults affect the total force generated by the vessel. To address these faults, a passive fault-tolerant control strategy is proposed (see figure 2.1). The fault scenario considers an over-actuated vessel equipped with multiple thrusters operating under unknown ocean disturbances. Although some thrusters may suffer severe degradation or total failure, only scenarios are considered where the vessel retains a sufficient number of functional thrusters to safely continue the operation. Mathematically, this capability is maintained by ensuring the controllability condition $rank(\boldsymbol{T}\boldsymbol{K}) = 3$ always holds. This constraint guarantees that, regardless of which specific thrusters fail, the surviving active thrusters are positioned and sufficiently healthy to collectively generate the control forces and moments required across all three degrees of freedom (surge, sway, and yaw).

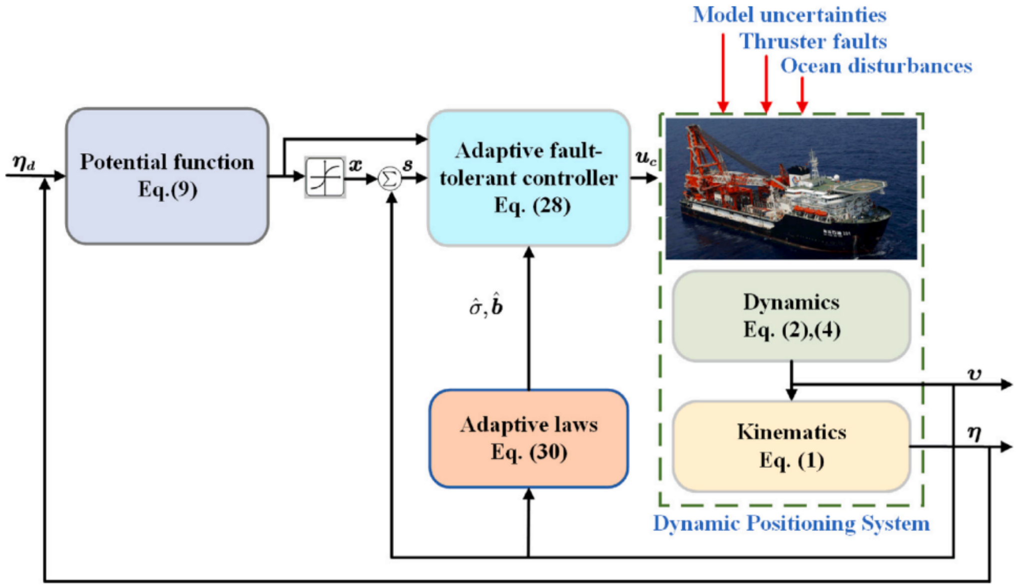


Figure 2.1: Structure diagram of the control system proposed in [13]

The control system handles these faults and drifting forces by grouping them together as a single, unknown disturbance. Instead of diagnosing exactly in which thruster the faults occur, the system utilizes adaptive laws that continuously monitor the vessel's deviation from the desired path. As errors accumulate, the control signals are adjusted until the faults and drifting forces are compensated.

In [12], the loss of effectiveness fault is considered. The fault is modeled by scaling the desired thruster output T_i by a degradation parameter λ (where $0 \leq \lambda \leq 1$), resulting in an actual output of $T_{if} = \lambda T_i$. Secondly, the maximum thruster output is capped at a value T_{ifmax} that is less than the healthy capacity (T_{hmax}), thereby restricting the overall control input bounds. The modified control input vector is defined as $u_{kf} = [T_h \quad \lambda T_f]$, combining the control input for the healthy thruster (T_h) and the faulty thruster (λT_f). This control input vector then affects the system dynamics, which is expressed as $x_{k+1} = f_{nom}(x_k, u_{kf}) + Q \cdot nn(q_k)$.

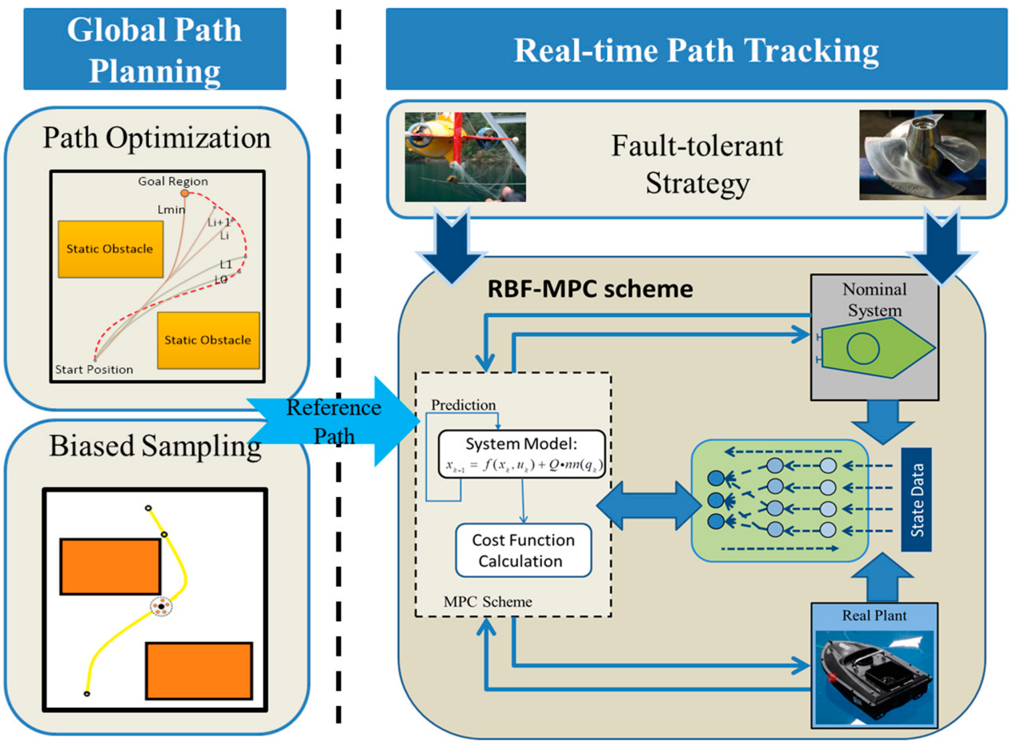


Figure 2.2: Structure diagram of the control system proposed in [12]

To manage the loss of effectiveness fault and unknown external disturbances, a Radial Basis Function Neural Network-based (RBF-NN) Model Predictive Control scheme is utilized (see figure 2.2). This scheme uses the RBF-NN to continuously learn and approximate unmodeled dynamics online, represented by the $Q \cdot nn(q_k)$ term in the x_{k+1} equation, by calculating the error between the predictions of the nominal model and the actual state and adjusting the control input accordingly. The scenarios include an ASV equipped with two (left and right) thrusters. Even when a fault occurs in one thruster, the vessel retains sufficient control from the remaining functional thrusters to safely continue the operation.

In [1], a stuck fault (constant azimuth thruster angle) and outage is considered. The scenarios include a stuck fault in both azimuth thrusters, outage in one thruster and outage in both thrusters while being exposed to drifting forces due to wind. These faults are modeled as constraints to the control input, which limit the forces and rotational rates of the thrusters. Mathematically, these faults are modeled as constraints on the control input $\boldsymbol{u}$. Specifically, an outage in one of the thrusters is modeled by changing the control input bound of the affected thruster into an equality constraint equal to zero. This results in $u_1 = 0$ or $u_2 = 0$, where u_1 and u_2 are the control input for the thruster force for both the stern and bow thruster, respectively. A stuck fault is modeled by restricting the rotational rate constraints to zero and locking the azimuth thruster angle at the exact position where the fault occurred. This results in $u_3 = 0$ or $u_4 = 0$, where u_3 and u_4 are the control input for the rotational turning rate of the azimuth thruster for both the stern and bow thruster, respectively. To account for the risk of grounding, an exponential risk term in the MPC objective function is proposed. This work considers scenarios where the vessel always retains sufficient control to counter the drifting forces and actively control its position.

In [8], the loss of effectiveness fault is considered. The fault is modeled using fault parameter vector $\boldsymbol{\theta}$, where $[\boldsymbol{\theta}]_i \in [0, 1]$. A value of 1 signifies a healthy component, while a value less than 1 indicates a fault. Mathematically, this fault acts as an affine parameter within the nonlinear system dynamics, expressed as $\boldsymbol{x}_{t+1} = \boldsymbol{f}(\boldsymbol{x}_t) + \boldsymbol{G}(\boldsymbol{u}_t)\boldsymbol{\theta} + \boldsymbol{d}_t$. Here, $\boldsymbol{d}_t$ represents the unknown but bounded environmental disturbances and $\boldsymbol{G}(\boldsymbol{u}_t)$ denotes as the input map. The vector $\boldsymbol{\theta}$ serves as a multiplicative scaling factor on the input map. The proposed scheme employs an active fault-tolerant control strategy known as Contingency-, Robust-Adaptive-Model Predictive Control (CRAMPC). This scheme requires a fault diagnosis module to enabling accurate detection, isolation, and quantification of faults (see figure 2.3).

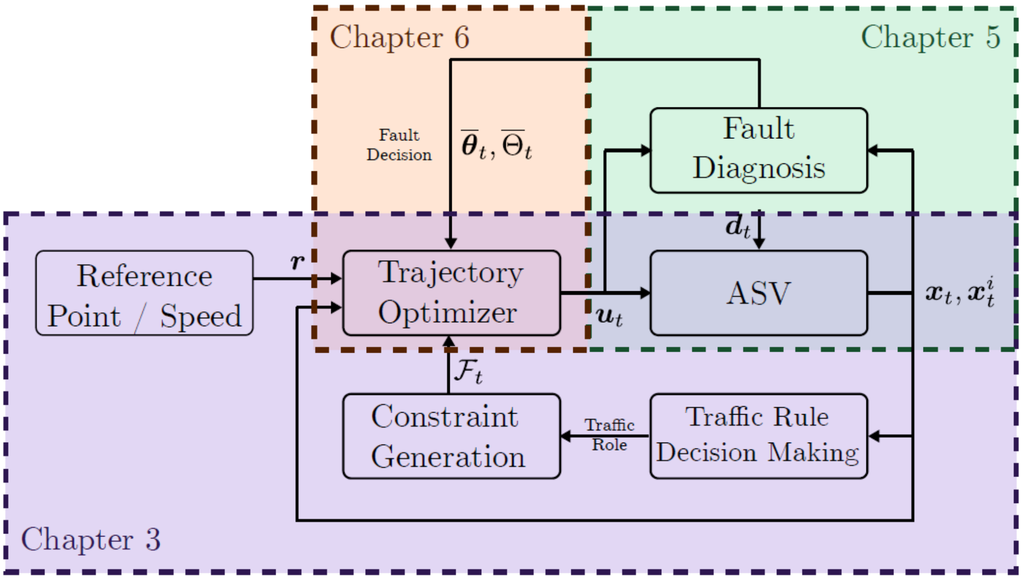


Figure 2.3: Block diagram scheme proposed in [8]

The operational scenario involves an ASV navigating a maritime environment where the ASV must avoid collisions with other vessels strictly according to maritime traffic rules (COLREGs). The scheme handles faults by constantly optimizing two trajectories simultaneously. The primary plan acts as the optimal trajectory, using the real-time fault diagnosis. Simultaneously, the contingency plan provides a mission-driven trajectory for the worst-case single fault realization for an over-actuated vessel. By forcing both of these plans to share the exact same initial control input, the ship never takes a step that would leave it trapped.

In [14], the loss of effectiveness fault is considered. This fault is modeled using fault parameter η_j, where each $\eta_j \in [0, 1]$ represents the remaining capacity of a specific thruster. A value of 1 denotes a fully healthy thruster, while values below 1 indicate a loss of effectiveness. This fault impacts the vessel's dynamic behavior, which is described by the state evolution $\dot{\boldsymbol{z}} = \boldsymbol{f}(\boldsymbol{z}) + \boldsymbol{g}(\boldsymbol{u}) + \boldsymbol{d}$, where $\boldsymbol{z} = [x, y, \psi, u, v, r]^\top$ is the state vector and $\boldsymbol{d}$ represents environmental disturbances due to wind and waves. The control input map $\boldsymbol{g}(\boldsymbol{u})$ and the generalized force vector $\boldsymbol{\tau}(\boldsymbol{u})$, which incorporates the fault parameters for the left (η_l), right (η_r), and bow (η_b) thrusters, are defined as:

$$\boldsymbol{g}(\boldsymbol{u}) = \begin{bmatrix} \boldsymbol{0}_{3\times 3} \\ \boldsymbol{M}^{-1} \end{bmatrix} \boldsymbol{\tau}(\boldsymbol{u}), \quad \boldsymbol{\tau}(\boldsymbol{u}) = \begin{pmatrix} \eta_l \tau_l \cos\alpha_l + \eta_r \tau_r \cos\alpha_r \\ \eta_l \tau_l \sin\alpha_l + \eta_r \tau_r \sin\alpha_r + \eta_b \tau_b \\ w_{lr}(\eta_r \tau_r \cos\alpha_r - \eta_l \tau_l \cos\alpha_l) - l_{lr}(\eta_l \tau_l \sin\alpha_l + \eta_r \tau_r \sin\alpha_r) + l_b \eta_b \tau_b \end{pmatrix}.$$

This work focuses on the Fault Diagnosis system. The Fault Diagnosis system utilizes nonlinear observers to continuously calculate residual errors and adaptive thresholds to enhance fault detection accuracy. The scenario considers an over-actuated ASV equipped with azimuth thrusters. Once the system isolates the faulty thruster, it reconfigures the thrusters accordingly. By rotating the remaining healthy azimuth thrusters to new angles, the system cleverly reallocates the workload to balance out the faulty thruster (see figure 2.4). This allows the vessel to safely continue the operation.

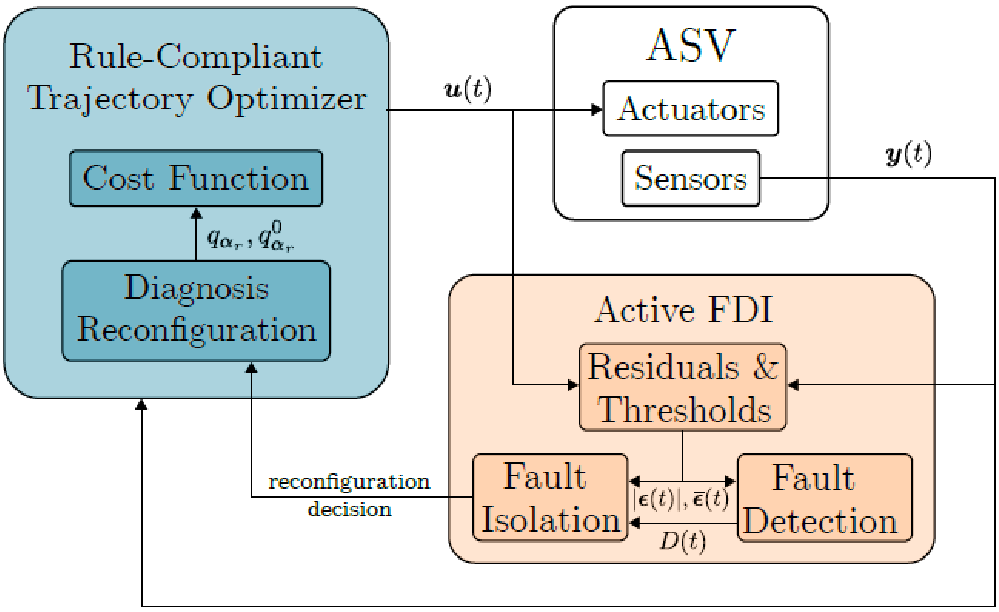


Figure 2.4: Schematic method overview proposed in [14]

The current steering and propulsion fault handling strategies broadly fall into either passive or active fault-tolerant control strategies. Passive strategies, such as the adaptive laws in [13] or the neural-network-based approximations in [12], treat faults as disturbances. These strategies continuously adjust the control inputs to compensate for errors without isolating the specific fault. Conversely, active fault-tolerant control strategies rely on fault diagnosis modules to detect, isolate and quantify faults. In [8], a multi-trajectory CRAMPC scheme is leveraged to prepare for worst-case single fault scenarios, while in [14] nonlinear observers are used to isolate faulty thrusters and reconfigure the remaining thrusters to maintain control.

However, a critical limitation is the extent of the fault scenarios, particularly regarding the severity of the fault in conjunction with the drifting forces. In the literature, only the fault scenarios are considered where the vessel is over-actuated or will retain sufficient control to counter the drifting forces and safely continue the mission. When drifting forces overpower the vessel's remaining control, it results in the drift of the vessel. When drifting towards grounding obstacles, it can result in the grounding of the vessel. In these catastrophic scenarios, continuing the mission might be impossible, yet existing control schemes lack the capability to focus on guiding the vessel towards a safe state and cease the operation when these scenarios occur. The steering and propulsion fault scenarios and operational modes considered in the literature are summarized in table 2.1.

Ref	LoE	Stuck	Bias	Outage	# of Actuators	Env. Forces	Op. Mode
[13]	x		x	x	Over-actuated	Unknown	Fail-operational
[12]	x				2	Unknown	Fail-operational
[1]		x		x	2	Wind	Fail-operational
[8]	x				Over-actuated	Unknown	Fail-operational
[14]	x				3	Wind & Waves	Fail-operational
Research Gap							
					1		Fail-Safe

Table 2.1: Summary of fault scenarios and operational mode

2.3 Fail-Safe Fault Scenarios

To address the gap in handling catastrophic scenarios where the drifting forces overpower the remaining control, this work specifically investigates *loss of effectiveness* and *stuck faults*, rather than bias or outage faults. This selection is driven by how these faults alter the control capabilities of the vessel. A total outage, which is mathematically modeled by setting the thrust constraints to zero (e.g. $u_1 = u_2 = 0$), leaves a vessel with no control, assuming there is no backup [1]. In such an event, the vessel is entirely at the mercy of the drifting forces. Conversely, bias faults represent an unknown additive force (e.g. $\boldsymbol{u}_c + \boldsymbol{d}$) [13]. Because the bias fault merely shifts the baseline operating point, this fault can be negated by adding an equal and opposite command. Therefore, bias faults will have limited impact on the control capabilities of the vessel.

However, a loss of effectiveness or a stuck fault can significantly limit the control capabilities, such that mission continuation becomes impossible. The loss of effectiveness fault is modeled by scaling the desired thruster output by a degradation parameter, resulting in an actual output which is a fraction of the desired output (e.g. $T_{if} = \lambda T_i$) [12]. This degradation can be managed by scaling the commanded control input accordingly to compensate for the loss. Secondly, the maximum thruster output is capped at a value that is less than the healthy capacity (e.g. $T_{ifmax} < T_{hmax}$) [12]. With the reduced maximum available thruster output, the vessel potentially cannot generate the force required to push back against strong drifting forces, leading to the drift of the vessel. During the stuck fault, the vessel loses steering capabilities. A stuck fault is modeled by restricting the rotational rate constraints to zero and locking the azimuth thruster angle at the exact position where the fault occurred (e.g. $u_3 = u_4 = 0$) [1]. When subjected to drifting forces, the stuck fault might make it impossible for the vessel to generate the opposing lateral forces required to counter the drifting forces. Thus, both the loss of effectiveness and stuck fault preserve some controllability, distinguishing it from a complete outage, but leave the vessel vulnerable to the drifting forces.

As demonstrated in [8, 14], fault diagnosis systems are highly capable of detecting, isolating, and quantifying the exact location and magnitude of faults in real time. However, the design and integration of a fault diagnosis module fall outside the scope of this work. Instead, this work assumes an ideal fault diagnosis. Consequently, upon the occurrence of a fault, the exact location, type, and magnitude are assumed to be instantaneously and perfectly known.

This work will focus on a vessel equipped with a single rudder and propeller at the stern of the vessel. The environment in which the vessel is operating is directly inspired by the scenarios described in [1]. The ASV will be operating in a challenging maritime environment defined as a confined channel or strait with grounding obstacles on both sides. Simultaneously, the vessel will be subjected to drifting force due to wind. The wind is set to push the vessel directly towards the grounding obstacles. In the stuck rudder fault scenario, continuing the mission might be impossible as the vessel potentially loses the ability generate the lateral force components required to counteract the drift induced by the wind. In the loss of effectiveness fault scenario, the maximum available propeller force might not be sufficient to counter the drifting forces. The scenarios are summarized in table 2.2.

In these extreme cases where mission continuation becomes physically impossible, current risk-aware control schemes, which have no fail-safe capabilities, persistently attempt to follow the reference path even if it results in grounding. Instead, the proposed scheme uses a multi-trajectory approach governed by a supervisory system. The supervisory system should continuously evaluate the optimized nominal and fail-safe trajectory to detect hazardous situations that require a fail-safe action. If the supervisory system detects that the vessel loses its stopping capability while the risk of grounding increases along the nominal trajectory, the supervisory system should trigger the fail-safe action. Evaluating the stopping capability is essential because, under steering and propulsion faults in conjunction with strong environmental drifting forces, the vessel may physically lack the capability to bring itself to a stop and counter the drifting forces due to wind. However, a degraded stopping capability alone does not immediately necessitate abandoning the mission if the vessel is operating in waters without an immediate risk of grounding or if continuing the mission results in moving towards a lower risk area. Therefore, the supervisory system simultaneously checks the risk along the nominal trajectory. If the risk is predicted to increase along the nominal trajectory, the nominal mission is actively driving the compromised vessel into a high-risk area. By triggering the fail-safe action, the focus shifts from mission progression to vessel survival. The vessel should utilize all remaining propulsion and steering capacity to guide the vessel away from the obstacles and towards a state of minimum risk of grounding.

1. Stuck-Rudder Fault with Crosswind
In this scenario, the rudder becomes mechanically locked while navigating a narrow passage. The vessel is simultaneously subjected to drifting forces pushing it towards grounding obstacles. a) The vessel retains control over the propeller, but loses steering control. b) With the fixed rudder angle the vessel potentially loses the ability to generate the lateral force component required to counter the drift induced by the wind.
2. Loss of Effectiveness Fault with Crosswind
In this scenario, the vessel suffers a fault in the propulsion system resulting in a reduced maximum available propeller force while navigating a narrow passage. This occurs while the vessel is subjected to drifting forces due to wind. a) The vessel continues to have steering control, but the maximum propeller force that the vessel can generate is only a fraction of the maximum force when there is no fault. b) When subjected to strong drifting forces, the propeller force might be insufficient to counter those forces, resulting in the drift of the vessel.

Table 2.2: Steering and propulsion fault scenarios

2.4 Conclusion

In conclusion, this chapter addressed the first research sub-question: *What steering and propulsion fault scenarios create hazardous situations that require a fail-safe action?*

Through a review of the literature, steering and propulsion faults can be categorized into a loss of effectiveness, outage, bias, and stuck fault. It was established that the bias fault is manageable using fault-tolerant control strategies, which consider the bias and adjust the commands accordingly. This ensures the vessel can successfully continue the mission and avoid grounding. Outage leaves the vessel completely at the mercy of environmental forces, assuming there is no backup system.

When there is a loss of effectiveness fault, the vessel can only deliver a fraction of the commanded output. Secondly, this fault limits the maximum available propeller force to a new, lower level. With the reduced maximum available propeller force, the vessel potentially cannot generate the force required to push back against strong drifting forces, leading to the drift of the vessel. Upon the occurrence of the stuck fault, the rudder angle is stuck at the last applied control input, eliminating the steering capability of the vessel. When subjected to drifting forces, the stuck rudder fault might make it impossible for the vessel to generate the opposing lateral forces required to counter the drifting forces. A hazardous situation that requires a fail-safe action can occur when a loss of effectiveness or stuck rudder fault occurs in conjunction with strong drifting forces that overpower the remaining control and downwind obstacles. When the vessel begins to drift towards those obstacles, the system must identify the hazardous situation and execute a fail-safe action, utilizing all remaining control to move away from the obstacles. Ultimately, this chapter identifies two specific hazardous situations requiring a fail-safe action: a stuck-rudder and loss of effectiveness fault occurring within a confined channel while being subjected to drifting forces that overpower the remaining control.

Chapter 3

Design

This chapter presents the design of the proposed Multi-Trajectory Risk-Aware MPC scheme, directly addressing the second research sub-question: *How to design a risk-aware MPC in which both the nominal and fail-safe trajectories are considered?* The design of the Multi-Trajectory Risk-Aware MPC scheme is built upon several key components. First, a 3-DOF ship model is presented, which captures the physical behaviors of the vessel. This includes the effects of mass, damping, control inputs, and wind forces. Next, the path-following methodology is explained. This section introduces the decision variable used to guide the vessel along the reference path. Subsequently, the chapter describes how the continuous-time dynamics are discretized into a finite-dimensional problem using the RK4 method. Then, the parameter vector is formulated, which consists of the reference speed and obstacle data. Next, the complete problem is defined by detailing the two main components: the full set of system constraints that govern the behavior of both the nominal and fail-safe trajectory, and the objective function. This chapter concludes by defining the supervisory system, which evaluates the solution for hazardous situations and manages the activation of the fail-safe action. Throughout this work, both the nominal and fail-safe trajectory will be referred to as trajectory I and II, respectively.

3.1 Control Structure

Figure 3.1 provides the control structure of the Multi-Trajectory Risk-Aware MPC scheme, illustrating the flow of information between the core components: planning, decision-making, and actuation. An important aspect of this control structure is the distinction between the two time horizons. The outer loop, indexed by $t \in \{1, \dots, T\}$, represents the discrete optimization time-steps. Within each of these time-steps t, the MPC performs an optimization over an internal prediction horizon, indexed by $k \in \{0, \dots, N\}$. Here, N represents the look-ahead capability of the MPC. At every optimization time-step t, the MPC optimizes the nominal and fail-safe trajectories from the current state ($k = 0$) into the future up to $k = N$. This structure implements a receding horizon strategy: although a complete trajectory is computed over N steps, only the first control action is applied.

The process begins with the current state of the ASV, denoted as ($\boldsymbol{x}_t$), being fed into the MPC block. Within this block, the complete optimization problem is formulated. It takes the current state $\boldsymbol{x}_t$ as an input, along with the reference path, parameter vector ($\boldsymbol{\theta}$), and the current operational mode, defined by the binary variable $B_t \in \{0, 1\}$. The parameter vector contains data about the reference velocity and obstacles.

Within the MPC block, an iterative loop continues until convergence is reached. The output is the optimal solution sequence, denoted as $\boldsymbol{w}_t$, which contains all the decision variables for the entire prediction horizon $k \in \{0, \dots, N\}$.

The supervisory system relies on the optimal solution sequence, alongside the fault status variable G_t provided by the Fault Diagnosis (FD). The variable G_t provides information regarding the health of the vessel, where $G_t = 1$ indicates a healthy, fault-free state, and $G_t = 0$ signals that there is an active steering and propulsion fault. The supervisory system first checks if the MPC is able to find the optimal solution. If no feasible solution is found, the supervisory system evaluates whether a backup strategy can be safely deployed. The fault status variable G_t is used in this decision process. If the MPC was able to find the solution $\boldsymbol{w}_t$, the supervisory system will analyze the safety of the optimal solution. If the solution is safe, the control input $\boldsymbol{u}_t$ is passed on to the actuator. Simultaneously, if the supervisory system detects a hazardous situation that requires the activation of the fail-safe action, it switching the binary variable B_t to zero. This modifies the objective function for subsequent optimization time-steps by deactivating the part of the objective function for the nominal trajectory. Finally, in the actuation phase, the actuator block translates the command $\boldsymbol{u}_t$ into physical forces and moments ($\boldsymbol{\tau}_t$), which act on the ASV to produce the new state $\boldsymbol{x}_{t+1}$, completing the loop. It is assumed that only a single fault can occur during a given operation and that the occurrence coincides precisely with the discrete time-steps.

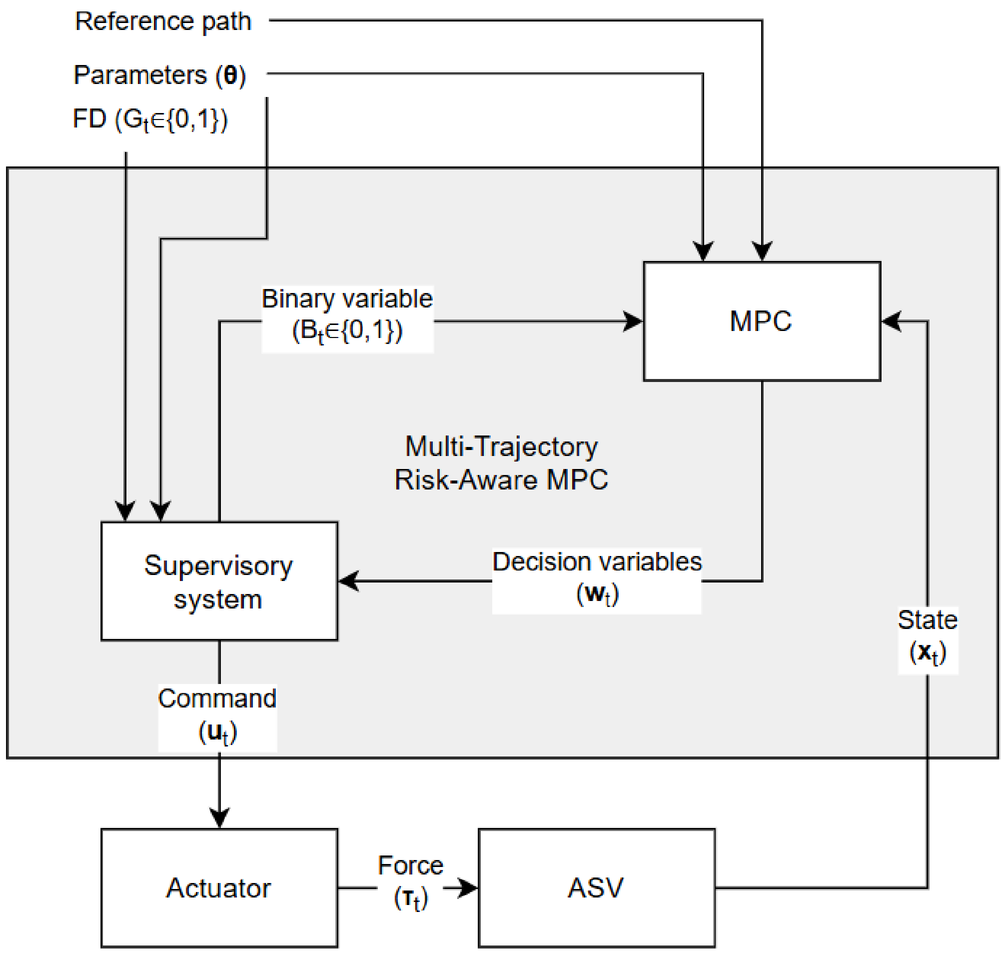


Figure 3.1: Overview of the control structure

3.2 3-DOF Ship Model

For the optimization of both trajectories, a dynamic model of the vessel is required. This model must capture the key dynamic behaviors of a vessel, such as acceleration, deceleration, and turning. The model used in this work is a 3-DOF model based on the ship models described in [1, 6, 15, 16].

3.2.1 Kinematic and Kinetic Equation

The motion of the vessel in the horizontal plane can be described by a kinematic and kinetic equation. The dynamics of the vessel in the horizontal plane are described by a state vector $\boldsymbol{x}$, which combines the earth-fixed position and body-fixed velocities. The kinematic equation relates the velocities of the vessel in the body-fixed frame, $\boldsymbol{\nu} = [u, v, r]^\top$ (surge, sway, and yaw rate), to the rate of change of the absolute position and orientation in the earth-fixed frame, $\boldsymbol{\eta} = [x, y, \psi]^\top$. This transformation is performed by the rotation matrix $\boldsymbol{J}(\psi)$, which rotates the body-fixed velocities according to the current heading of the vessel, ψ. The state vector and rotation matrix are defined as follows:

$$\boldsymbol{x} = \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\nu} \end{bmatrix} = \begin{bmatrix} x \\ y \\ \psi \\ u \\ v \\ r \end{bmatrix}, \qquad \boldsymbol{J}(\psi) = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The kinetic equation, determines the body-fixed accelerations of the vessel, $\dot{\boldsymbol{\nu}}$. These accelerations are driven by the forces and moments acting on the vessel. These include the control force vector $\boldsymbol{\tau}$, external disturbance vector $\boldsymbol{d}$ (consisting of the wind force vector), and the damping matrix $\boldsymbol{D}$. The resistance of the vessel to acceleration is captured by the inertia matrix $\boldsymbol{M}$, which includes both the rigid body mass of the vessel and the added mass. The kinematic and kinetic equations are:

$$\dot{\boldsymbol{\eta}} = \boldsymbol{J}(\psi)\,\boldsymbol{\nu} \qquad \text{(Kinematic equation)}$$
$$\dot{\boldsymbol{\nu}} = \boldsymbol{M}^{-1}\big(\boldsymbol{\tau} + \boldsymbol{d} - \boldsymbol{D}\,\boldsymbol{\nu}\big) \qquad \text{(Kinetic equation)}$$

3.2.2 Wind Force

$\boldsymbol{d}$ represents the external disturbances in the body-fixed frame and consists of the wind force vector $\boldsymbol{\tau}_{\text{wind}}$, which includes the forces and moment due to wind acting on the vessel in the body-fixed frame and are given by:

$$\boldsymbol{\tau}_{\text{wind}} = \begin{bmatrix} -c_x\cos(\gamma_w)A_{Fw} \\ c_y\sin(\gamma_w)A_{Lw} \\ c_n\sin(2\gamma_w)A_{Lw}L_{oa} \end{bmatrix} \frac{1}{2}\rho_a V_w^2.$$

The forces and moment are proportional to the pressure of the wind, $\frac{1}{2}\rho_a V_w^2$, with ρ_a being the air density and V_w represents the relative wind velocity. $\gamma_w = \psi - \psi_w - \pi$ is the angle of attack, where ψ is the heading of the vessel and ψ_w is the true wind direction. The longitudinal force (surge) is proportional to the frontal wind area of the vessel, A_{Fw}, while the lateral force (sway) and the yaw moment are proportional to the lateral wind area, A_{Lw}. The overall length of the vessel, L_{oa}, acts as a lever in the calculation of the yaw moment. The forces and moment are scaled by a set of dimensionless wind force coefficients, c_x, c_y, and c_n, which values are based on the values provided in [1]. The values of the wind force coefficients are the following:

$$c_x = 0.7, \qquad c_y = 0.8, \qquad c_n = 0.1.$$

3.2.3 Inertia Matrix

When a vessel accelerates in water, it must also accelerate some of the surrounding fluid. This is captured by the added mass $X_{\dot{u}}$, $Y_{\dot{v}}$, $N_{\dot{r}}$. Physically, each coefficient represents the equivalent mass of water that must be "carried along" by the hull when it accelerates in the surge, sway and yaw direction. Consequently, the total mass in each degree of freedom is the sum of the rigid body mass of the vessel (or moment of inertia) and the corresponding added mass. These coefficients depend on hull shape and fluid density, and are typically obtained from potential-flow theory or model-scale experiments [17]. When including the added mass in the matrix, the Inertia matrix $\boldsymbol{M}$ becomes:

$$\boldsymbol{M} = \begin{bmatrix} m + X_{\dot{u}} & 0 & 0 \\ 0 & m + Y_{\dot{v}} & 0 \\ 0 & 0 & I_z + N_{\dot{r}} \end{bmatrix}.$$

The term m corresponds to the rigid body mass of the vessel, while I_z is the moment of inertia about the vertical axis, governing the resistance to the yaw motion. These inertial properties are augmented by the added mass coefficients, $X_{\dot{u}}$, $Y_{\dot{v}}$, and $N_{\dot{r}}$, which account for the additional inertia of the water that must be accelerated with the vessel in the surge, sway, and yaw direction, respectively.

3.2.4 Damping Matrix

When the vessel moves at steady velocity or changes the velocity, it experiences resistive (damping) forces and moments. These are included in the damping matrix $\boldsymbol{D}$:

$$\boldsymbol{D} = \begin{bmatrix} X_u & 0 & 0 \\ 0 & Y_v & 0 \\ 0 & 0 & N_r \end{bmatrix}.$$

The diagonal elements X_u, Y_v, and N_r are the linear damping coefficients. Specifically, X_u quantifies the resistance to surge motion, Y_v quantifies the resistance to sway motion, and N_r quantifies the rotational resistance to the yaw rate.

3.2.5 Control Forces

The vessel is equipped with a single propeller and rudder located at the stern. This propeller-rudder configuration is modeled as a single point force and the control forces are approximated using a trigonometric decomposition. In this model, the rudder is assumed to deflect the total propeller force f by an angle a.

The control input vector is defined as $\boldsymbol{u} = [f, a]^\top$, where f represents the magnitude of the propeller force and a represents the rudder angle. Based on f and a shown in figure 3.2, the control force vector $\boldsymbol{\tau}$ can be described as:

$$\boldsymbol{\tau} = \begin{bmatrix} fcos(a) \\ fsin(a) \\ -l_x fsin(a) \end{bmatrix}.$$

The vector components describe the decomposition of the total propeller force into specific body-fixed forces and moments. The term $f\cos(a)$ represents the surge force directed along the longitudinal axis, while $f\sin(a)$ represents the sway force acting along the lateral axis. The yaw moment is given by $-l_x f\sin(a)$, where l_x denotes the lever arm distance between the center of gravity of the vessel and the rudder post. The negative sign is included to reflect the physical configuration, indicating that a positive lateral force applied at the stern generates a moment opposing the positive yaw direction.

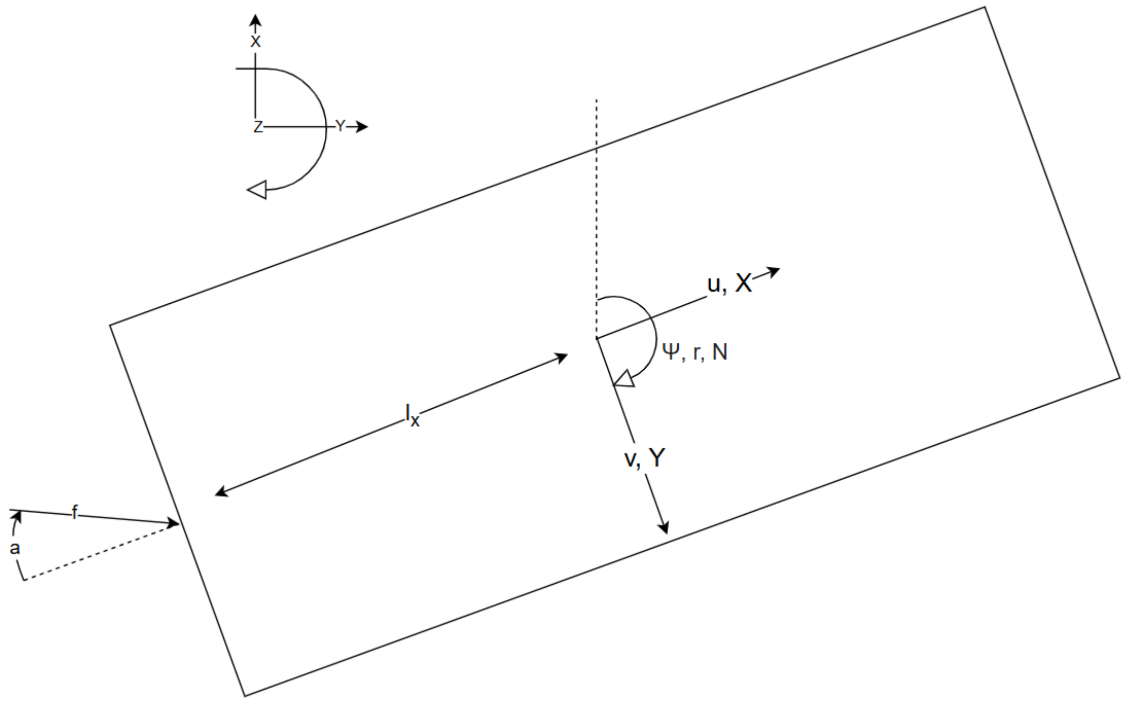


Figure 3.2: Model variables and coordinate frames used in this work

3.3 Reference Path

The reference path is parameterized using a two-dimensional vector function $\boldsymbol{r}(\alpha)$, which maps an advancement variable α to a specific coordinate in the global reference frame (see figure 3.3). α allows the MPC to determine the optimal reference point $\boldsymbol{r}(\alpha)$ for each step in the prediction horizon. This degree of freedom enables the vessel to accelerate or decelerate along the path as needed. This formulation is based on [1] and [7] and is defined as:

$$\boldsymbol{r}(\alpha) = \begin{bmatrix} r_x(\alpha) \\ r_y(\alpha) \end{bmatrix}, \quad \text{for } \alpha \in [0, L_{total}],$$

where α is the advancement variable representing the distance traveled along the path (distance in meters), and L_{total} is the total length of the path. The reference point on the path depends on the advancement variable α and $r_x(\alpha)$ and $r_y(\alpha)$ are the x and y coordinates of the reference point. $\alpha = 0$ corresponds to the start of the mission, and $\alpha = L_{total}$ corresponds to the end. The reference path is constructed by interpolating a sequence of waypoints $\mathcal{Q} = \{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_M\}$, where M is the total number of waypoints and each $\boldsymbol{q}_i = [x_i, y_i]^\top$

represents a fixed geographical coordinate. This work utilizes piecewise linear interpolation. The path consists of straight line segments connecting consecutive waypoints. For a segment between two waypoints $\boldsymbol{q}_i$ and $\boldsymbol{q}_{i+1}$, the reference function is defined as [18]:

$$\boldsymbol{r}(\alpha) = \boldsymbol{q}_i + \left(\frac{\alpha - \alpha_i}{\|\boldsymbol{q}_{i+1} - \boldsymbol{q}_i\|}\right)(\boldsymbol{q}_{i+1} - \boldsymbol{q}_i), \quad \text{for } \alpha_i \leq \alpha \leq \alpha_{i+1}$$

Here, α_i represents the total distance traveled along the path at waypoint i. The cumulative distance α_i at each waypoint is calculated as the sum of Euclidean distances between preceding points:

$$\alpha_1 = 0, \quad \alpha_i = \sum_{i=1}^{i-1} \|\boldsymbol{q}_{i+1} - \boldsymbol{q}_i\|, \quad i > 1$$

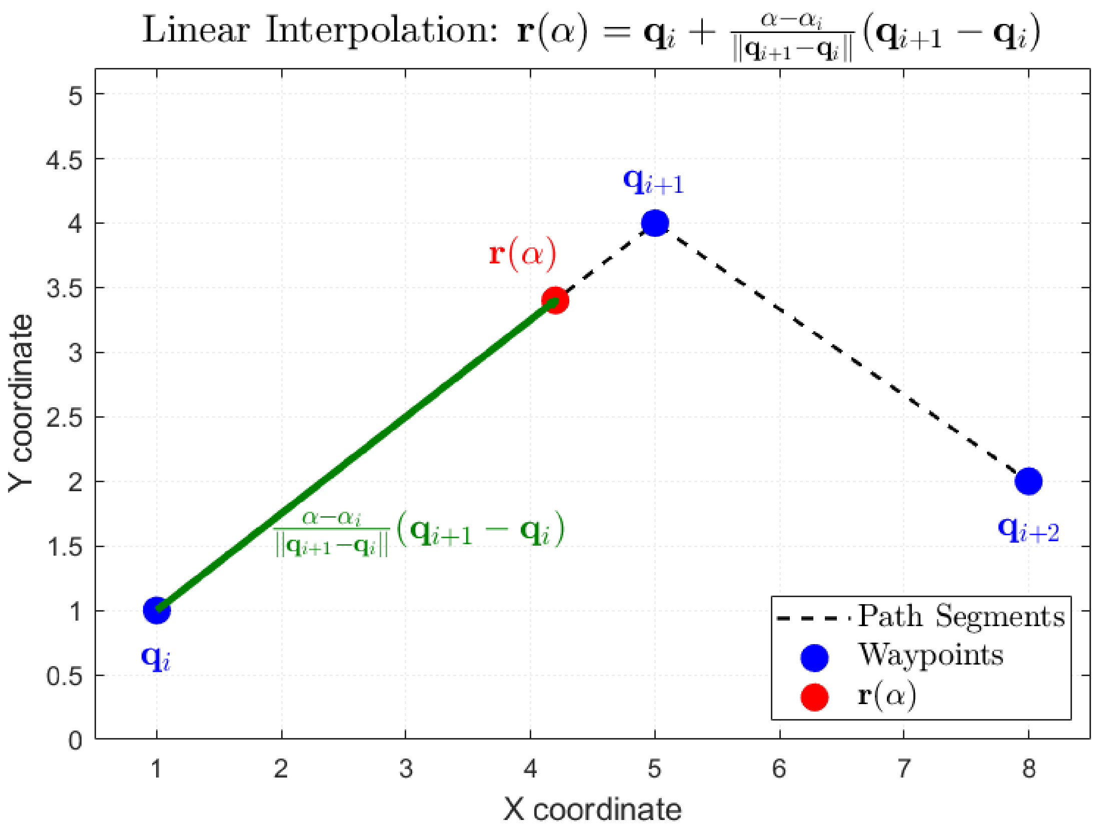


Figure 3.3: Visualization of reference path planning using piecewise linear interpolation

3.4 Ordinary Differential Equation

The ship dynamics are modeled using the three-dimensional kinematic and kinetic equations in the horizontal plane, which are the time derivatives of the state vector [1]. $\dot{\boldsymbol{x}}$ is assembled by combining both $\dot{\boldsymbol{\eta}}$ and $\dot{\boldsymbol{\nu}}$. This results in the following first-order ODE:

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}) = \begin{bmatrix} \dot{\boldsymbol{\eta}} \\ \dot{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{J}(\boldsymbol{\eta})\boldsymbol{\nu} \\ \boldsymbol{M}^{-1}(\boldsymbol{\tau} + \boldsymbol{\tau}_{\text{wind}} - \boldsymbol{D}\boldsymbol{\nu}) \end{bmatrix}$$

To optimize the nominal and fail-safe trajectory, the continuous-time ODE is discretized into a finite-dimensional problem by dividing the prediction horizon into N intervals with a duration of Δt. The control inputs are the same for the entire time interval. The dynamic behavior of the vessel is approximated using the RK4 integrator [1, 7]. The RK4 method uses intermediate evaluations (denoted $\boldsymbol{k}_1$, $\boldsymbol{k}_2$, $\boldsymbol{k}_3$ and $\boldsymbol{k}_4$) of the derivative function to approximate the next state (see figure 3.4). The process begins by calculating the initial slope, denoted as $\boldsymbol{k}_1$, which is the derivative of the current state. Using this initial slope, the integrator projects a half-step forward to estimate the state at the midpoint of the interval (Δt). A second slope, $\boldsymbol{k}_2$, is then calculated at this preliminary midpoint. Then this estimate is refined by disregarding the first half-step and taking a new half-step using the slope $\boldsymbol{k}_2$ to determine a third slope, $\boldsymbol{k}_3$. Finally, this improved midpoint slope is used to project a full step across the entire interval duration Δt, yielding a fourth slope estimate, $\boldsymbol{k}_4$, which estimates the behavior of the system at the end of the time-step. The next state, $\boldsymbol{x}_{\text{next}}$, is determined by advancing the current state along a weighted average of these four derivatives. In the final equation for the next state, the two midpoint slopes ($\boldsymbol{k}_2$ and $\boldsymbol{k}_3$) are given double the weight of the start and end slopes ($\boldsymbol{k}_1$ and $\boldsymbol{k}_4$). This weighting reflects the fact that, for many dynamic systems, the derivative at the center of the interval provides the most accurate representation of the average rate of change over the full duration [15]. The steps involve:

$$\boldsymbol{k}_1 = f(\boldsymbol{x}, \boldsymbol{u}), \qquad \boldsymbol{k}_2 = f(\boldsymbol{x} + \frac{\Delta t}{2}\boldsymbol{k}_1, \boldsymbol{u}), \qquad \boldsymbol{k}_3 = f(\boldsymbol{x} + \frac{\Delta t}{2}\boldsymbol{k}_2, \boldsymbol{u}), \qquad \boldsymbol{k}_4 = f(\boldsymbol{x} + \Delta t \boldsymbol{k}_3, \boldsymbol{u}).$$

$$\boldsymbol{x}_{\text{next}} = \boldsymbol{x} + \frac{\Delta t}{6}(\boldsymbol{k}_1 + 2\boldsymbol{k}_2 + 2\boldsymbol{k}_3 + \boldsymbol{k}_4).$$

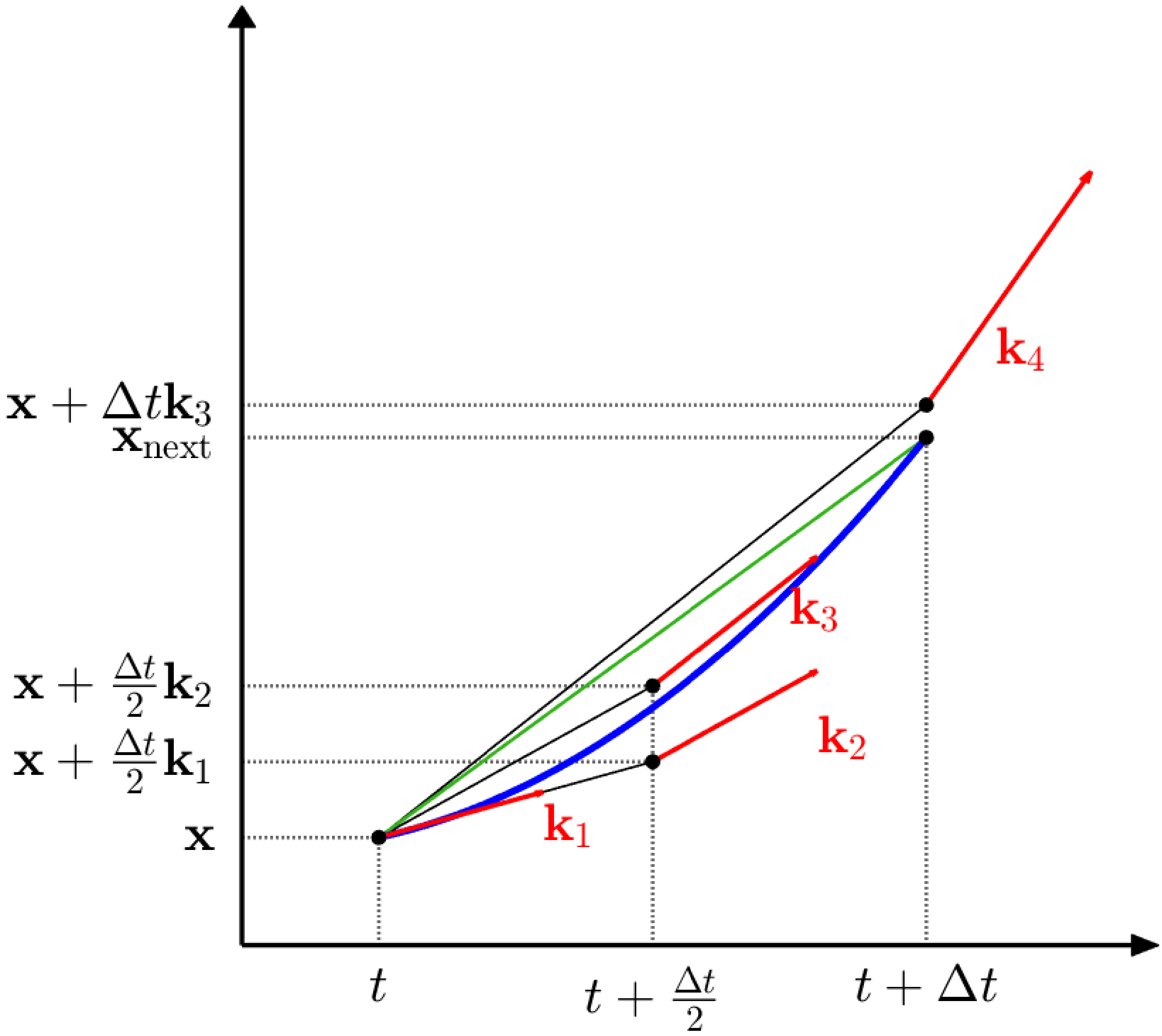


Figure 3.4: Visualization of the Fourth Order Runge Kutta method

3.5 Decision Variable Vector

The decision variable vector $\boldsymbol{w}$ represents the complete set of optimization variables that the MPC manipulates to determine the optimal solution. In this work, the MPC simultaneously optimizes the nominal and fail-safe trajectory. x_k^I and x_k^{II}, which are the state vectors of the nominal (denoted by superscript I) and fail-safe trajectory (denoted by superscript II), respectively, constitute the first component of this decision variable vector. These variables represent the predicted physical state of the vessel at each interval k within the prediction horizon. Typically, each state vector consists of the six degrees of freedom required to describe the state of the vessel in the horizontal plane, including position, heading, and the corresponding surge, sway, and yaw velocities. The MPC adjusts these values to satisfy the dynamic constraints of the vessel, ensuring that both the nominal and fail-safe trajectories are physically feasible according to the model of the vessel (see section 3.7). $\boldsymbol{u}_k^I$ and $\boldsymbol{u}_k^{II}$, which are the control input vectors of the nominal (denoted by superscript I) and fail-safe trajectory (denoted by superscript II), respectively, represent the actuator commands required to drive the vessel from one state to the next. Finally, the vector includes the advancement variable, α_k, which functions as a variable tracking the progress along the reference path. The vector $\boldsymbol{w}$ stacks these state vectors, control input vectors, and advancement variables for every step from $k = 0$ to the end of the horizon N, forming the complete search space for the optimization problem. The decision variables can be combined in the following vector:

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{x}_0^{I\top} & \boldsymbol{x}_0^{II\top} & \alpha_0 & \boldsymbol{u}_0^{I\top} & \boldsymbol{u}_0^{II\top} & \dots & \boldsymbol{x}_{N-1}^{I\top} & \boldsymbol{x}_{N-1}^{II\top} & \alpha_{N-1} & \boldsymbol{u}_{N-1}^{I\top} \\ \boldsymbol{u}_{N-1}^{II\top} & \boldsymbol{x}_N^{I\top} & \boldsymbol{x}_N^{II\top} & \alpha_N \end{bmatrix}^\top.$$

It is important to note the difference in the time horizons between the state vector and advancement variable and the control input vector within the decision vector $\boldsymbol{w}$. The state vectors $\boldsymbol{x}_k^I$ and $\boldsymbol{x}_k^{II}$ and path progression term α_k are defined for the entire horizon up to the final step $k = N$, whereas the control input vectors $\boldsymbol{u}_k^I$ and $\boldsymbol{u}_k^{II}$ are only defined up to $k = N - 1$. This distinction arises because the control inputs $\boldsymbol{u}_{N-1}^I$ and $\boldsymbol{u}_{N-1}^{II}$ represents the final action required to drive the vessel from the penultimate state $\boldsymbol{x}_{N-1}^I$ and $\boldsymbol{x}_{N-1}^{II}$ to the terminal state $\boldsymbol{x}_N^I$ and $\boldsymbol{x}_N^{II}$. Since the optimization horizon ends at step N, the MPC does not compute any further control actions beyond this point, as there is no subsequent state $\boldsymbol{x}_{N+1}^I$ and $\boldsymbol{x}_{N+1}^{II}$ to transition towards. Consequently, the final elements of the decision variable vector consist solely of the terminal states and advancement variable.

3.6 Parameter Vector

The solution, provided by the MPC, is governed by a set of parameters that define the mission objectives and operational constraints. These parameters are combined into a single parameter vector, $\boldsymbol{\theta}_k$ and consists of the reference transit speed s_{ref}, the advancement rate α_{step}, and the obstacle vector $\boldsymbol{\sigma}$ [1]. Together the parameter vector is defined as:

$$\boldsymbol{\theta}_k = \begin{bmatrix} s_{ref} \\ \alpha_{step} \\ \boldsymbol{\sigma} \end{bmatrix} \in R^{2+3J},$$

The first two components define the desired motion along the path. The parameter s_{ref} represents the reference transit speed, which sets the target velocity for the vessel. The parameter α_{step} represents the advancement rate. It defines the desired progress of the advancement variable α along the path over the prediction horizon. This parameter is used in the objective function (see section 3.8) to penalize deviating from the desired progression rate. The third component, $\boldsymbol{\sigma}$, encapsulates the environmental constraints, specifically the set of grounding obstacles. To maintain computational efficiency, complex obstacle geometries are modeled as a set of J circular exclusion zones, which are assumed to account for the dimensions of the vessel. The obstacle parameter vector $\boldsymbol{\sigma}$ is a vector defining the characteristics for all J obstacles:

$$\boldsymbol{\sigma} = \begin{bmatrix} \boldsymbol{\sigma}_1^\top & \boldsymbol{\sigma}_2^\top & \dots & \boldsymbol{\sigma}_J^\top \end{bmatrix}^\top$$

Each individual obstacle $\boldsymbol{\sigma}_j$ is defined by three scalar values, which include the coordinates of the center and the radius. This is expressed as:

$$\boldsymbol{\sigma}_j = \begin{bmatrix} x_j \\ y_j \\ r_j \end{bmatrix}, \qquad j \in \{1, ..., J\}.$$

Here, x_j and y_j denote the global coordinates of the center of obstacle j, while r_j represents the radius. An example of such a circular obstacle is shown in figure 3.5.

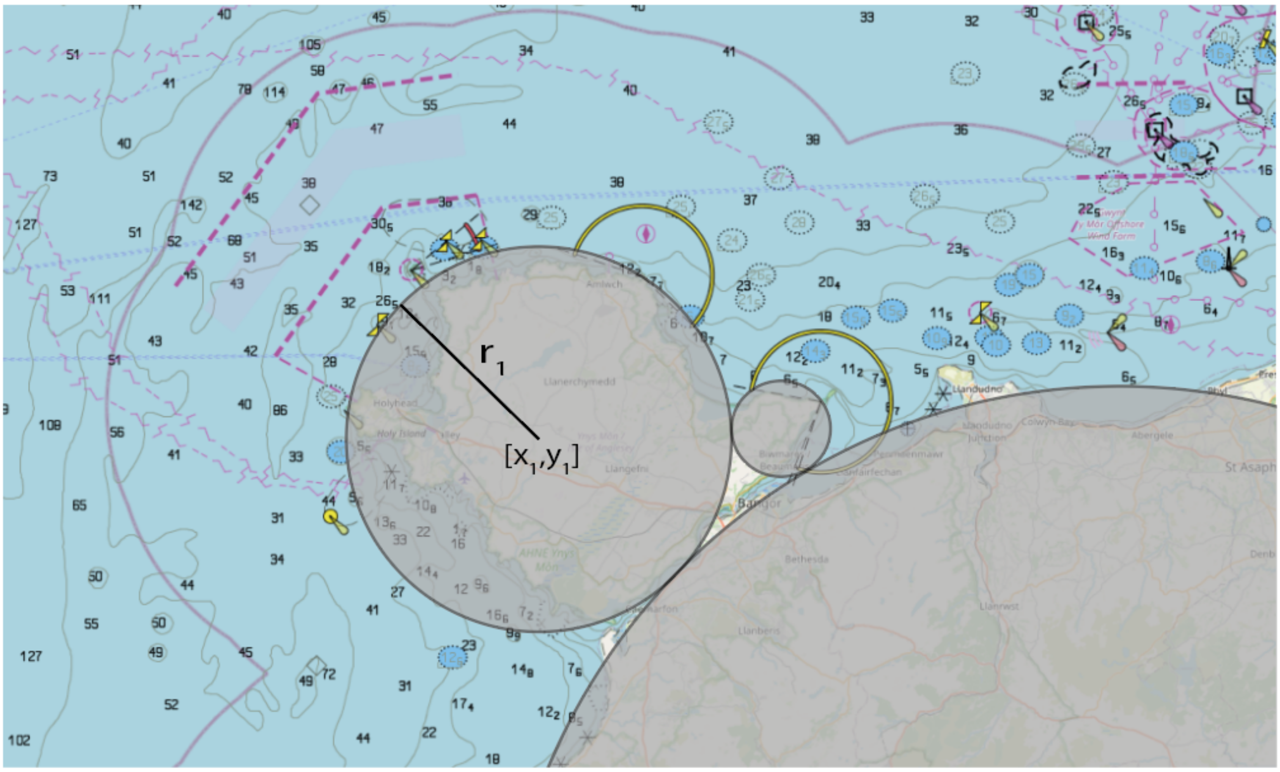


Figure 3.5: Example of circular exclusion zones [19]

3.7 System Constraints

To ensure that the solution is physically feasible and adheres to all operational rules, the solution must satisfy a set of constraints. These constraints govern the behavior of both the nominal and fail-safe trajectories and include the following:

$$\boldsymbol{x}_{k+1}^{I} = f(\boldsymbol{x}_k^{I}, \boldsymbol{u}_k^{I}) \quad \text{(System dynamics nominal trajectory)} \tag{3.1a}$$

$$\boldsymbol{x}_{k+1}^{II} = f(\boldsymbol{x}_k^{II}, \boldsymbol{u}_k^{II}) \quad \text{(System dynamics fail-safe trajectory)} \tag{3.1b}$$

$$\alpha_k \geq \alpha_{k-1} \quad \text{(Path progress)} \tag{3.1c}$$

$$\boldsymbol{u}_0^{I} = \boldsymbol{u}_0^{II} \quad \text{(Shared initial control input)} \tag{3.1d}$$

$$-\boldsymbol{u}_{limit} \leq \boldsymbol{u}_k^{I} \leq \boldsymbol{u}_{limit} \quad \text{(Nominal control bounds)} \tag{3.1e}$$

$$-\boldsymbol{u}_{limit} \leq \boldsymbol{u}_k^{II} \leq \boldsymbol{u}_{limit} \quad \text{(Fail-safe control bounds)} \tag{3.1f}$$

$$\alpha_{\text{step}} = s_{\text{ref}} \cdot \Delta t \quad \text{(Path progression step)} \tag{3.1g}$$

To ensure the optimization problem yields a solution that is physically realizable, the MPC must satisfy a set of constraints that define the boundaries of the search space. Among these are the system dynamics constraints, represented by equations (3.1a) and (3.1b). These equality constraints enforce that the predicted state of the vessel at any future step $k+1$ must be the precise physical result of applying the control input $\boldsymbol{u}_k$ to the current state $\boldsymbol{x}_k$.

The advancement of the vessel along the reference path is governed by the constraint defined in equation (3.1c). This inequality ensures that the advancement variable α_k, which determines the location of the reference point, can only move forward along the path or remain stationary, but never move backwards. This will encourage mission progression. Complementing this is equation (3.1g), which establishes the desired advancement step size. This definition effectively calibrates the desired speed of the operation, providing the target rate of change that the objective function uses to guide the vessel forward along the path.

A critical feature of this multi-trajectory MPC is the shared initial control input constraint, $\boldsymbol{u}_0^{I} = \boldsymbol{u}_0^{II}$, as stated in equation (3.1d). By forcing the initial control inputs of both trajectories to be identical, the MPC must find a specific action that is optimal for continuing the mission but simultaneously valid as the first step of a fail-safe action. This reduces the possibility that the nominal trajectory leads to a state from which an escape is impossible, since the previous fail-safe action is executed if no new, valid solution can be found.

Finally, the physical limitations of the steering and propulsion system are enforced through the control input bound constraints in equations (3.1e) and (3.1f). These inequalities bound the control inputs, specifically the propeller force and rudder angle, within the operating limits $\boldsymbol{u}_{limit} = [f_{limit}, a_{limit}]^\top$. These limits apply to both the nominal and fail-safe trajectories. This ensures feasibility by preventing the MPC from generating solutions that require impossible control inputs, such as propeller forces or rudder angles that exceed the limits. This ensures that the vessel is able to execute the control input.

3.8 Objective Function

At every optimization time-step t, the MPC solves a receding horizon problem to determine the best course of action. This process involves finding the optimal decision variable vector $\boldsymbol{w}$ that minimizes the total cost over the prediction horizon while satisfying all constraints [1]. The mathematical formulation of this problem is defined as follows:

$$\begin{aligned} &\text{min} \quad \varphi(\boldsymbol{w}, \boldsymbol{\theta}) \\ &\text{s.t.} \qquad\qquad \text{eq. [3.1a-3.1g]}, \end{aligned}$$

This represents the formulation of the problem solved by the MPC at each optimization time-step t. The expression min $\varphi(\boldsymbol{w}, \boldsymbol{\theta})$ commands the MPC to search the feasible solution space for the specific decision variable vector $\boldsymbol{w}$, consisting of the predicted states, control inputs, and path advancement variables, that yields the lowest possible cost. This minimization is dependent on the parameter vector, denoted by $\boldsymbol{\theta}$. The constraints impose the necessary restrictions on this search, ensuring that the optimal solution adheres to the constraints defined in equations [3.1a-3.1g]. This section details the objective function ($\varphi(\boldsymbol{w}, \boldsymbol{\theta})$) that enables the MPC to balance competing objectives and switch between operational modes. It begins by presenting the primary objective function, which consists of several distinct cost terms designed to guide the behavior of the vessel (see equation 3.2). An important element of this formulation is the binary variable, B, which allows the system to switch between the nominal mode and the fail-safe mode (see section 3.9).

The primary objective of the nominal trajectory is to guide the vessel along the reference path using the progression term, ξ. During the nominal operation, energy efficiency and stable control should be considered. This is enabled by the nominal control input term, ϵ^I. Finally, the optimization of the nominal trajectory must consider the environment. This is enabled by the nominal risk term, ρ^I. This term accounts for the risk of grounding. It ensures that even when the vessel is actively trying to complete the mission, it will deviate from the reference path and speed, rather than blindly pushing forward. Crucially, all these terms are multiplied by the binary variable, B. If the supervisory system detects a hazardous situation that requires a fail-safe action, these objectives can be deactivated by switching to the fail-safe mode ($B = 0$).

In contrast, the fail-safe trajectory serves as a contingency plan with the objective to guide the vessel to a safe state. The fail-safe trajectory is computed with a risk cost term in the MPC objective function to guide the vessel towards minimum-risk conditions. The risk cost term is divided into the running (ρ^{II}) and terminal (ρ^T) risk cost. As the MPC solves an optimization problem over a finite horizon (N), the MPC only considers what happens within that window. Without a terminal cost, the optimizer might sacrifice the state at the end of the horizon to minimize costs earlier, potentially leaving the vessel in an unsafe state. Separating these terms allow the running risk cost to be tuned independently from the terminal risk cost. In addition, the terminal velocity cost term, ζ, ensures that the velocity of the vessel can be reduced at each optimization step, to bring it to a near-stop at the end of the fail-safe trajectory. Finally, to ensure stable control, the objective function includes the fail-safe control input term, ϵ^{II}. The objective function is derived based on the objective function proposed in [1] and is defined as:

$$\begin{aligned} \varphi(\boldsymbol{w}, \boldsymbol{\theta}) = &\sum_{k=0}^{N} \Bigg[\underbrace{B\xi(\boldsymbol{x}_k^I, \boldsymbol{\theta}_k, \alpha_k, \alpha_{k-1})}_{\text{Progression}} + \underbrace{B\rho^I(\boldsymbol{x}_k^I, \boldsymbol{\theta}_k)}_{\text{Nominal Risk}} \Bigg] \\ &+ \sum_{k=0}^{N-1} \Bigg[\underbrace{B\epsilon^I(\boldsymbol{u}_k^I, \boldsymbol{u}_{k-1}^I)}_{\text{Nominal Control Input}} + \underbrace{\epsilon^{II}(\boldsymbol{u}_k^{II}, \boldsymbol{u}_{k-1}^{II})}_{\text{Fail-Safe Control Input}} + \underbrace{\rho^{II}(\boldsymbol{x}_k^{II}, \boldsymbol{\theta}_k)}_{\text{Fail-Safe Risk}} \Bigg] \\ &+ \underbrace{\rho^T(\boldsymbol{x}_N^{II}, \boldsymbol{\theta}_N)}_{\text{Terminal Risk}} + \underbrace{\zeta(\boldsymbol{x}_N^{II})}_{\text{Terminal Velocity}} \quad . \end{aligned} \tag{3.2}$$

3.8.1 Progression Cost Term

The first component of the objective function, $\xi(\boldsymbol{x}_k^I, \boldsymbol{\theta}_k, \alpha_k, \alpha_{k-1})$, encodes the objective for the nominal, goal-directed trajectory (I). It is given by:

$$\xi(\boldsymbol{x}_k^I, \boldsymbol{\theta}_k, \alpha_k, \alpha_{k-1}) = \boldsymbol{\kappa} \begin{bmatrix} \|\boldsymbol{r}(\alpha_k) - \boldsymbol{p}_k^I\|^2 \\ \|\alpha_k - \alpha_{k-1} - \alpha_{\text{step}}\|^2 \end{bmatrix}, \tag{3.3}$$

The first component of the objective function, denoted as $\xi(\boldsymbol{x}_k^I, \boldsymbol{\theta}_k, \alpha_k, \alpha_{k-1})$, is designed to guide the ship along the reference path. This is achieved by penalizing deviations from the desired trajectory and encouraging consistent progress. Specifically, the term $\|\boldsymbol{r}(\alpha_k) - \boldsymbol{p}_k^I\|^2$ quantifies the squared error between the predicted position of the vessel $\boldsymbol{p}_k^I = [x_k^I, y_k^I]^\top$ and the reference position on the path $\boldsymbol{r}(\alpha_k)$. Minimizing this term keeps the ASV close to the reference path. The second term, $\|\alpha_k - \alpha_{k-1} - \alpha_{\text{step}}\|^2$, encourages the ship to advance along the reference path at the desired rate α_{step}, by penalizing deviating from this rate. The influence of these terms is scaled by a weight vector, $\boldsymbol{\kappa}$, which allows for the tuning of their relative importance.

3.8.2 Nominal Control Input Cost Term

To ensure smooth and efficient control actions, the objective function penalizes the magnitude and rate of change of the propeller force and rudder angle for the nominal trajectory. This is encoded in the term $\epsilon^I(\boldsymbol{u}_k^I, \boldsymbol{u}_{k-1}^I)$:

$$\epsilon^I(\boldsymbol{u}_k^I, \boldsymbol{u}_{k-1}^I) = \boldsymbol{u}_k^{I\top} \boldsymbol{\Lambda}^I \boldsymbol{u}_k^I + (\boldsymbol{u}_k^I - \boldsymbol{u}_{k-1}^I)^\top \boldsymbol{\Delta}^I (\boldsymbol{u}_k^I - \boldsymbol{u}_{k-1}^I) \tag{3.4}$$

The first component of control input cost term, weighted by the diagonal matrix $\boldsymbol{\Lambda}^I \in R^{2\times2}$, penalizes the magnitude of the control inputs. This discourages large power commands and extreme rudder angles, thereby reducing energy consumption. The second component, weighted by the diagonal matrix $\boldsymbol{\Delta}^I \in R^{2\times2}$, penalizes the difference between consecutive control inputs. This improves smoothness by encouraging that control inputs do not change abruptly over time, leading to more stable vessel behavior. [1]

3.8.3 Fail-Safe Control Input Cost Term

To ensure stable control when executing a fail-safe action, the objective function penalizes the rate of change of the control inputs for the fail-safe trajectory. This is encoded in the term $\epsilon^{II}(\boldsymbol{u}_k^{II}, \boldsymbol{u}_{k-1}^{II})$:

$$\epsilon^{II}(\boldsymbol{u}_k^{II}, \boldsymbol{u}_{k-1}^{II}) = (\boldsymbol{u}_k^{II} - \boldsymbol{u}_{k-1}^{II})^\top \boldsymbol{\Delta}^{II} (\boldsymbol{u}_k^{II} - \boldsymbol{u}_{k-1}^{II}) \tag{3.5}$$

The fail-safe control input cost term, weighted by the diagonal matrix $\boldsymbol{\Delta}^{II} \in R^{2\times2}$, penalizes the difference between consecutive control inputs. This improves smoothness by encouraging that control inputs do not change abruptly over time, leading to more stable vessel behavior.

The distinction in cost terms between the nominal and fail-safe trajectories reflects the fundamentally different operational objectives of each trajectory. The nominal trajectory is designed for long-term, mission-driven operations, where both energy efficiency and smoothness are considered. In contrast, the fail-safe trajectory is optimized solely to prevent accidents. In a hazardous situation that requires a fail-safe action, energy efficiency is irrelevant. The only priority is preventing accidents. Therefore, the magnitude cost is removed to allow the vessel to use maximum thrust if necessary to avoid grounding. However, the rate of change penalty is retained to ensure that the fail-safe trajectory remains stable.

3.8.4 Risk Cost Terms

To quantify the vulnerability to grounding obstacles and wind disturbances, risk cost terms are applied. These terms calculate a risk cost, denoted by ρ, which penalizes proximity to the set of obstacles defined in the parameter vector $\boldsymbol{\theta}_k$. The risk cost is computed for the nominal trajectory (ρ^I), the fail-safe trajectory (ρ^{II}), and the terminal state of the fail-safe trajectory (ρ^T). The generalized risk cost term, with $s \in \{I, II, T\}$, is defined as:

$$\rho^s(\boldsymbol{x}_k^s, \boldsymbol{\theta}_k) = \sum_{j=1}^{J} \left(\mu_1^s + \mu_2^s\, V_w\, \max(0, \hat{\boldsymbol{i}}_{j,k}^s \cdot \hat{\boldsymbol{\omega}}_{\text{wind}})\right) e^{-\Xi^s(\|\boldsymbol{p}_k^s - \boldsymbol{c}_j\| - r_j)} \tag{3.6}$$

The risk cost term acts as a repulsive potential field around obstacles. This formulation is designed to quantify the risk of grounding posed by obstacles as a continuous cost that scales with proximity and wind conditions. The core of each equation is the exponential term, $e^{-\Xi^s(\|\boldsymbol{p}_k^s - \boldsymbol{c}_j\| - r_j)}$. This component calculates the Euclidean distance between the current position of the vessel $\boldsymbol{p}_k$ and the center of the j-th obstacle $\boldsymbol{c}_j$. By subtracting the radius of the obstacle r_j, the distance is measured relative to the boundary of the obstacle rather than the center. The tuning parameter Ξ governs the steepness of the potential field. A higher value creates a localized barrier that allows closer operation before penalties are significant, while a lower value generates a wider, more gradual gradient that influences the trajectory from a greater distance.

The magnitude of the risk penalty is scaled dynamically by the term, $(\mu_1^s + \mu_2^s\, V_w\, \max(0, \hat{\boldsymbol{i}}_{j,k}^s \cdot \hat{\boldsymbol{\omega}}_{\text{wind}}))$, which accounts for both static risk and additional risk introduced by the wind. The parameter μ_1 establishes a baseline cost for approaching an obstacle, independent of wind disturbances. The second component, weighted by μ_2, introduces a dependency on the wind speed and direction. To evaluate this, two unit vectors are defined: $\hat{\boldsymbol{i}}_{j,k}$ represents the direction from the current position of the vessel $\boldsymbol{p}_k$ to the center of the obstacle $\boldsymbol{c}_j$, calculated as $(\boldsymbol{c}_j - \boldsymbol{p}_k)/\|\boldsymbol{c}_j - \boldsymbol{p}_k\|$, while $\hat{\boldsymbol{\omega}}_{\text{wind}} = [\cos(\psi_w), \sin(\psi_w)]^\top$ is the unit vector representing the global wind direction ψ_w. The dot product $\hat{\boldsymbol{i}}_{j,k} \cdot \hat{\boldsymbol{\omega}}_{\text{wind}}$ evaluates the alignment between the direction to the obstacle and the wind vector. The inclusion of the $\max(0, \cdot)$ function ensures that this environmental penalty is asymmetric. It increases the cost only when the vessel is positioned upwind of the grounding obstacle (where the drifting force due to wind pushes the vessel towards the obstacle). Conversely, if the vessel is downwind of an obstacle, this term is equal to zero, reflecting the reduced risk of being downwind of an obstacle.

While the mathematical structure of the risk cost term is identical across the nominal and fail-safe trajectory and the terminal state of the fail-safe trajectory, the specific weighting parameters allow for independent tuning.

3.8.5 Terminal Velocity Cost Term

The terminal velocity cost term, $\zeta(\boldsymbol{x}_N^{II})$, penalizes non-zero velocity at the end of the fail-safe trajectory:

$$\zeta(\boldsymbol{x}_N^{II}) = \Omega((u_N^{II})^2 + (v_N^{II})^2) \tag{3.7}$$

$\zeta(\boldsymbol{x}_N^{II})$ serves as a terminal velocity penalty to encourage to bring the vessel to a near-stop at the end of the fail-safe trajectory. This is achieved by penalizing any non-zero velocity at the end of the prediction horizon ($k = N$). The term is defined by the squared velocity of the ship at the end of the fail-safe trajectory, $\sqrt{(u_N^{II})^2 + (v_N^{II})^2}$, calculated from the body-fixed frame velocity components u_N^{II} and v_N^{II}. The magnitude of this penalty is controlled by the weighting scalar Ω.

The quadratic nature of this term ensures that higher terminal velocities incur a disproportionately larger cost compared to lower velocities. This characteristic drives the MPC to prioritize aggressive deceleration when the terminal speed is high, while gradually relaxing the penalty as the velocity approaches zero. Furthermore, by formulating this as a cost term rather than a strict hard constraint (i.e., forcing $V_N^{II} = 0$), it provides the MPC with greater flexibility. It allows the MPC to find valid solutions where the terminal velocity is small but non-zero, particularly in scenarios where a complete stop is physically impossible with an active fault. This use of a "soft" constraint is crucial for computational efficiency and robustness, as it avoids the potential for the MPC to fail in finding a feasible solution under overly rigid constraints.

3.9 The Supervisory System

The supervisory system acts as a critical safety check. Positioned as shown in figure 3.1, the primary responsibility is to check solver convergence and the safety of the predicted trajectories before any control input is executed.

3.9.1 Solver Convergence

First, the supervisory system verifies if the MPC is able to find a solution. The system monitors the convergence status of the MPC, denoted as E_t, where $E_t = 1$ indicates successful convergence and $E_t = 0$ indicates a failure (e.g., infeasibility). To enhance the robustness of the system against temporary infeasibility, a backup strategy is implemented. If the MPC fails to find a solution ($E_t = 0$), the supervisory system does not immediately activate the fail-safe action. Instead, it evaluates the reliability of utilizing the solution from the previous optimization time-step. The decision of using the previous solution is governed by two conditions, which are:

1. **No Recent Fault:** A steering and propulsion fault has not occurred in the current time-step. This condition is mathematically verified by evaluating the inequality $G_t \geq G_{t-1}$. Relying on a previous plan derived for a healthy vessel would be unsafe if the capabilities of the vessel have just drastically changed due to a fault ($G_t = 0$ while $G_{t-1} = 1$).

2. **No Consecutive Solver Failures:** No consecutive solver failures are allowed, ensuring that the system does not operate based on the same solution for an extended duration. Since the current optimization has failed ($E_t = 0$), this condition requires that the MPC was able to find a solution in the previous time-step, mathematically expressed as $E_{t-1} = 1$.

Based on these conditions, the supervisory system determines the appropriate course of action when the MPC fails to find a solution ($E_t = 0$). If the conditions are met, the system bypasses the failed optimization by retrieving the solution from the previous time-step ($\boldsymbol{w}_{t-1}$) and shifts it forward in time. In the nominal mode ($B = 1$), the system executes the next control input derived from the shifted *nominal* trajectory ($\boldsymbol{u}_1^I$ from $\boldsymbol{w}_{t-1}$). In the fail-safe mode ($B = 0$), the system executes the next control input derived from the shifted *fail-safe* trajectory ($\boldsymbol{u}_1^{II}$ from $\boldsymbol{w}_{t-1}$). Conversely, if the backup is deemed unsafe, either due to a fresh fault or because the consecutive solver failure limit has been exceeded, the supervisory system immediate activates the fail-safe action ($B \to 0$) when the vessel is operating in the nominal mode ($B = 1$). In the fail-safe mode ($B = 0$), since no further fallback is available, the system terminates the optimization. The backup strategy logic is summarized in figure 3.6.

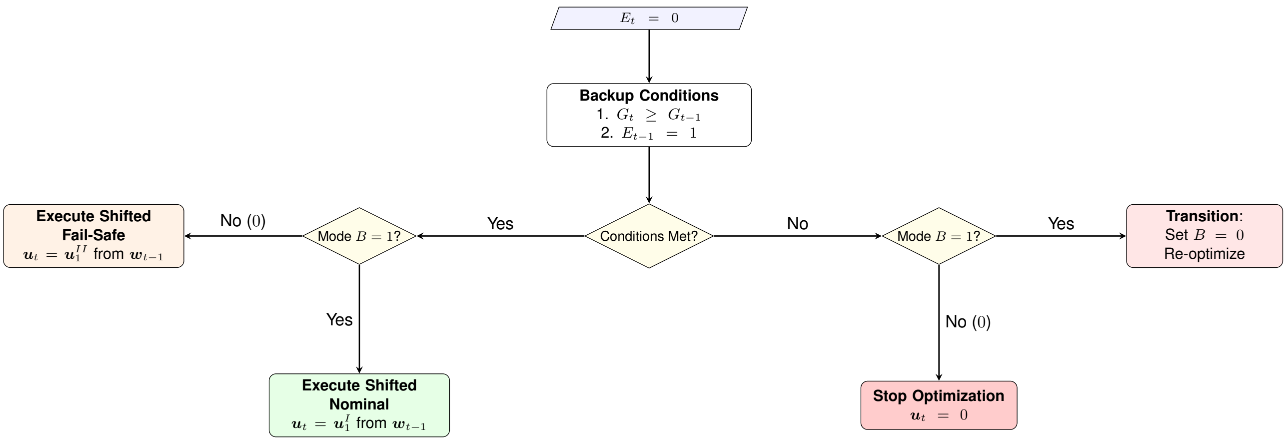


Figure 3.6: The backup strategy logic.

3.9.2 Safety Check

If the MPC successfully finds a solution ($E_t = 1$), the system ignores the backup logic and proceeds to the next stage: the safety check.

Grounding Safety Check

If the MPC is able to find a solution, the MPC returns the decision variable vector, denoted as $\boldsymbol{w}$. This vector contains the complete set of predicted states, path progression terms and control inputs for the entire prediction horizon. The supervisory system decomposes $\boldsymbol{w}$ into two matrices: the nominal position matrix $\boldsymbol{P}^I = [\boldsymbol{p}_0^I, ..., \boldsymbol{p}_N^I] \in R^{2\times(N+1)}$ and the fail-safe position matrix $\boldsymbol{P}^{II} = [\boldsymbol{p}_0^{II}, ..., \boldsymbol{p}_N^{II}] \in R^{2\times(N+1)}$. Each column k of these matrices corresponds strictly to the predicted position vector $\boldsymbol{p}_k = [x_k, y_k]^\top$ at time step k for both the nominal (superscript I) and fail-safe (superscript II) trajectory. The supervisory system evaluates these extracted matrices for grounding. The grounding evaluation function, $\mathcal{S}(\boldsymbol{P})$, uses the position matrix $\boldsymbol{P}$ and computes the minimum distance to any obstacle over the entire horizon:

$$\mathcal{S}(\boldsymbol{P}) = \min_{k\in\{0,...,N\}} \left(\min_{j\in\{1,...,J\}} (\|\boldsymbol{p}_k - \boldsymbol{c}_j\| - r_j) \right),$$

where $\boldsymbol{p}_k$ is the vector at index k of the matrix $\boldsymbol{P}$. The terms $\boldsymbol{c}_j$ and r_j denote the center coordinates and radius of obstacle j, respectively, which are extracted from the parameter vector $\boldsymbol{\theta}$. A trajectory is deemed physically safe if and only if $\mathcal{S}(\boldsymbol{P}) > 0$. By applying this function individually to $\boldsymbol{P}^I$ and $\boldsymbol{P}^{II}$, the supervisor makes sure the optimized trajectories do not predict grounding, ensuring that only grounding-free solutions are passed to the steering and propulsion system.

Drifting Safety Check

The second evaluation assesses the stopping capability of the vessel. During steering and propulsion faults where the remaining control is severely limited, drifting forces due to wind might overpower the remaining control. In these scenarios, bringing the vessel to a stop is impossible. The second safety check ensures that if the capabilities of the vessel are limited such that it cannot bring itself to a stop, the vessel cannot continue operating in the nominal mode when the risk is predicted to increase along the nominal trajectory. The supervisory system extracts the terminal velocity of the fail-safe trajectory, $v_{terminal} = \sqrt{(u_N^{II})^2 + (v_N^{II})^2}$, and computes the risk gradient along the nominal trajectory, $\Delta\rho = \rho^I(\boldsymbol{x}_N^I, \boldsymbol{\theta}_N) - \rho^I(\boldsymbol{x}_0^I, \boldsymbol{\theta}_0)$. If the terminal velocity exceeds the safe threshold (v_{th}) and the risk is predicted to increase along the nominal trajectory ($\Delta\rho > \Delta\rho_{\text{th}}$), the mission must be aborted.

Conversely, if the vessel cannot stop but the risk is decreasing along the nominal trajectory or the increase is below the threshold ($\Delta\rho \leq \Delta\rho_{\text{th}}$), the condition is not triggered, allowing the mission to continue as the risk of grounding is insignificant or is being reduced while the mission progresses. Furthermore, if the vessel is fully capable of bringing itself to a stop, the vessel is also permitted to continue in the nominal mode.

Incorporating this secondary safety check facilitates the early detection of hazardous situations that necessitate a fail-safe action. Relying solely on the grounding check means that a vessel with limited capabilities, which is actively drifting towards obstacles, would have to wait until grounding is predicted within the finite prediction horizon before activating the fail-safe action. In contrast, this drifting safety check detects the inability to counteract the drifting forces by evaluating the terminal velocity of the fail-safe trajectory. Concurrently, it uses the risk gradient of the nominal trajectory to identify whether the vessel is approaching obstacles or a high-risk area. By including this safety check, the supervisory system can preemptively abort the mission and activate the fail-safe action well before a grounding event falls within the prediction horizon.

Safety Check Logic

In the default nominal mode ($B = 1$), the MPC optimizes two trajectories: the nominal trajectory and the fail-safe trajectory. In the nominal mode, the supervisory system evaluates both trajectories for grounding and evaluates the terminal velocity of the fail-safe trajectory and the risk gradient along the nominal trajectory. The logic for the transition from $B = 1$ is defined as follows:

$$B_{next} = \begin{cases} 1 & \text{if } \mathcal{S}(\boldsymbol{P}^I) > 0 \land \mathcal{S}(\boldsymbol{P}^{II}) > 0 \land v_{\text{terminal}} \leq v_{\text{th}} \\ 0 & \text{if } \mathcal{S}(\boldsymbol{P}^I) \leq 0 \lor \mathcal{S}(\boldsymbol{P}^{II}) \leq 0 \lor (v_{\text{terminal}} > v_{\text{th}} \land \Delta\rho > \Delta\rho_{\text{th}}) \end{cases}$$

If $B_{next} = 1$, the nominal control input $\boldsymbol{u}_t = \boldsymbol{u}_0^I$ is executed. If $B_{next} = 0$, the supervisory system discards the current solution, updates the binary variable to $B = 0$, and triggers an immediate re-optimization in the current optimization time-step. Once B transitions to 0, it cannot revert to 1. In the fail-safe mode, the nominal cost terms are deactivated (see section 3.8). The MPC now optimizes only for the fail-safe trajectory. In the fail-safe mode, the supervisory system checks if the fail-safe trajectory is grounding-free. The logic in the fail-safe mode is defined as:

$$\boldsymbol{u}_t = \begin{cases} \boldsymbol{u}_0^{II} & \text{if } \mathcal{S}(\boldsymbol{P}^{II}) > 0 \\ 0 & \text{if } \mathcal{S}(\boldsymbol{P}^{II}) \leq 0 \end{cases}$$

In the scenario $\boldsymbol{u}_t = 0$, the system determines that no feasible fail-safe trajectory exists to avoid grounding from the current state and stops the optimization. The system effectively hands over control to a remote operator or drops the anchor as a last resort. This ensures that the vessel never executes a control input derived from an unsafe plan. By checking $\boldsymbol{P}^{II}$ even while in nominal mode, the system reduces the possibility that the nominal trajectory leads to a state from which grounding is inevitable, since the previous fail-safe action is executed if no new, valid solution can be found. The safety check logic is summarized in 3.7.

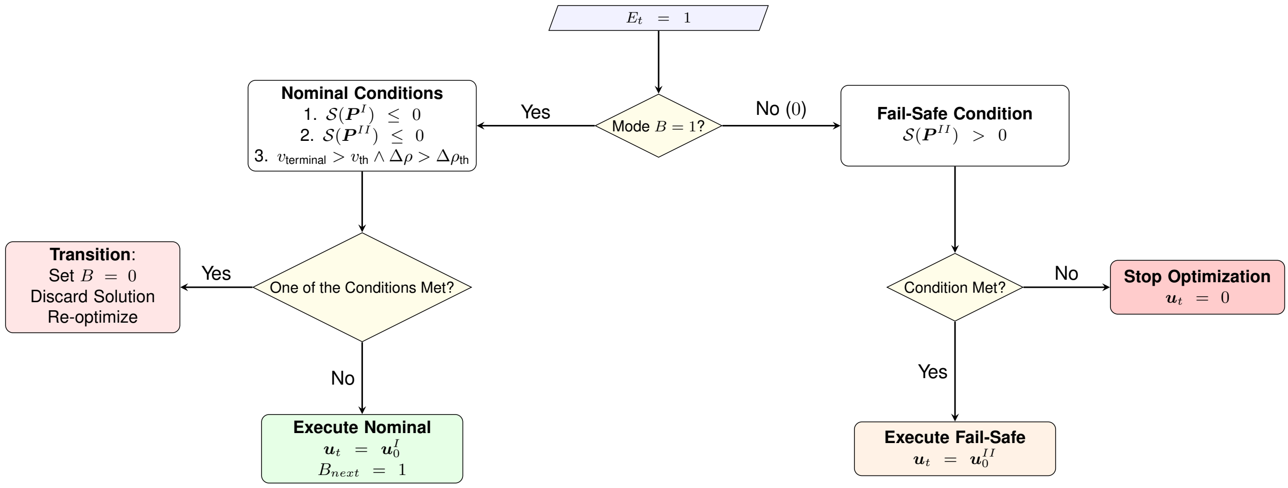


Figure 3.7: The safety check logic.

3.10 Conclusion

In conclusion, this chapter addressed the second research sub-question: *How to design a risk-aware MPC in which both the nominal and fail-safe trajectories are considered?*.

To answer this, a Multi-Trajectory Risk-Aware MPC scheme was developed, starting with a 3-DOF ship model capable of capturing essential vessel dynamics. The core of the design lies in the simultaneous optimization of two distinct trajectories: a nominal trajectory (I) responsible for mission progression, and a fail-safe trajectory (II), which has the objective to guide the vessel to a safe state. The fail-safe trajectory is computed with a risk cost term in the MPC objective function to guide the vessel towards minimum-risk conditions. In addition, a terminal velocity cost term ensures that the velocity of the vessel can be reduced at each optimization step, to bring it to a near-stop at the end of the fail-safe trajectory. The supervisory system can use the fail-safe trajectory to verify the stopping ability of the vessel. The objective function includes a binary variable (B), which makes it possible to deactivate the nominal cost terms of the objective function when triggering the fail-safe action. A critical design feature is the shared initial control input constraint. This ensures that every action taken by the vessel in the nominal mode is simultaneously the first step of the fail-safe action. Finally, the supervisory system continuously evaluates the safety of the predicted trajectories. The supervisory system monitors for predicted grounding by verifying the distances to all obstacles over the entire prediction horizon, and also evaluates the terminal velocity of the fail-safe trajectory alongside the risk gradient of the nominal trajectory. The supervisory system activates the fail-safe action if an imminent grounding is predicted or if the risk is predicted to increase over the nominal trajectory without the ability to safely come to a stop in the fail-safe trajectory.

Chapter 4

Verification

This chapter addresses the research sub-question: *How can the performance of the proposed Multi-Trajectory Risk-Aware MPC be verified through simulations and key performance indicators?* To answer this, the proposed Multi-Trajectory Risk-Aware MPC scheme is compared in a simulation environment against a single-trajectory risk-aware MPC scheme, which was proposed in previous work (see [1]). The verification relies on comparing both schemes in a simulation environment subjected to varying wind velocities and fault scenarios. By simulating normal fault-free operations alongside the stuck rudder and loss of effectiveness fault scenarios (based on chapter 2), the performances and the fail-safe capabilities of the MT scheme are verified.

The performance of the MT scheme is quantified using a set of KPIs. These metrics are selected to provide an evaluation across four categories: safety, path tracking performance, operational efficiency, and computational cost (see section 4.1). By comparing the performances between the MT and ST scheme, this chapter verifies whether the MT scheme enhances vessel safety and what the compromises are in terms of operating efficiency, path tracking performance and computational cost.

The remainder of this chapter is structured as follows. Section 4.1 details the ST scheme, the simulation scenarios, and the KPIs used to quantify the performances. Section 4.2 outlines the simulation environment. Section 4.3 defines the initial states, physical vessel properties, and the precise tuning of the MPC parameters. Finally, section 4.4 presents and discusses the results from the simulations.

4.1 Performance Evaluation and Comparative Analysis

To verify the effectiveness of the proposed Multi-Trajectory Risk-Aware MPC scheme. The performance is evaluated and compared against a ST risk-aware MPC scheme, which will be discussed in this section. Furthermore, this section describes the simulation scenarios selected to test both schemes, and defines the KPIs used for a quantitative comparison of their performances.

4.1.1 Single-Trajectory Risk-Aware Model Predictive Control Scheme

To provide a verification of the MT scheme, this work uses a reference for comparison. The performance of the MT scheme is therefore evaluated against the ST scheme, the logic of which is detailed in algorithm 1. The ST scheme represents the approach found in the literature (see [1]), where the MPC optimizes a single trajectory, which balances mission progression against risk at all times. Unlike the MT scheme, the ST scheme has no fail-safe capabilities. The comparison against the ST scheme is critical as it highlights the safety benefits of the MT scheme. By utilizing the exact same cost function weights, prediction horizon, and vessel model as the nominal trajectory of the MT scheme, the ST scheme acts as a reference for comparison. Improvements in safety can be attributed to the fail-safe capabilities of the MT scheme. Furthermore, it reveals the cost of this enhanced safety in terms of degradation in operational efficiency, path tracking accuracy and computational cost.

Algorithm 1 Pseudo-code for the single-trajectory risk-aware MPC scheme in a MATLAB simulation environment.

```
while not at destination AND not stop_optimization do
    Set initial guess for decision variables from w_{t-1}

    try
        Solve optimization problem for Trajectory I (Nominal)
        Save current solution to w_t
        u_t = u_0;
    catch solver_failure
        Check backup conditions

        if conditions are met
            u_t = u_1 from w_{t-1}
        else
            stop_optimization = true; break; // Stop simulation
        end
    end

    x_next = rk4step(x_t, u_t, dt);

    x_t = x_next;
end
```

4.1.2 Simulation Scenarios

Three types of simulation scenarios have been selected to evaluate the MT scheme against the ST scheme. These types consist of a no-fault, stuck-rudder fault, and loss of effectiveness fault scenario.

The no-fault scenario is used to compare the performance of both schemes during normal no-fault operations. In this scenario, the vessel operates with a fully functional steering and propulsion system. The primary expectation is that both the ST and MT scheme will successfully counteract the wind-induced drift, follow the reference path, and safely execute the mission across different wind velocities. This scenario will also be used to compare the performances quantitatively using the KPIs (see section 4.1.3) and identify the costs of optimizing the fail-safe and nominal trajectory simultaneously. In the stuck-rudder fault scenario, the rudder becomes locked in a fixed angle. The loss of effectiveness fault scenario is characterized by a tightened propeller force bound. Both faults severely limit the ability to successfully execute the mission. By testing both schemes for these faults across multiple wind velocities and triggering them at different stages of the mission, the simulations provide insight into how each scheme behaves under varying situations. While the ST scheme is expected to fail during some simulations in the stuck-rudder and loss of effectiveness fault scenario, the proposed MT scheme is designed to recognize hazardous situations that require a fail-safe action and execute a fail-safe action.

To analyze the behavior of both control schemes under degraded functionality, the schemes are tested in scenarios where the fault is resolved after a significant duration to test post-fault behavior, where the fault persists indefinitely, and where the wind velocity changes mid-operation.

4.1.3 Key Performance Indicators

A set of KPIs will provide a quantitative basis for comparison. These metrics are selected to evaluate the schemes across four categories: safety, path tracking, operational efficiency, and computational costs.

Safety Metrics

The primary indicator of safety is the **Distance to Grounding**. This metric quantifies the physical safety margin remaining between the vessel and the grounding obstacles. At each simulation step t, the Euclidean distance from the current position of the vessel $\boldsymbol{p}_t$ to the center $\boldsymbol{c}_j$ of each obstacle j is calculated. To obtain the true clearance, the radius of the obstacle, r_j, is subtracted from this value. The distance to grounding, which is equal to the minimum distance to an obstacle at time-step t, is denoted as $d_{\text{min}}(t)$, is defined as:

$$d_{\text{min}}(t) = \min_{j \in \{1 \ldots J\}} (\|\boldsymbol{p}_t - \boldsymbol{c}_j\| - r_j)$$

A positive value indicates a buffer until grounding, while a value of $d_{\text{min}} \leq 0$ signifies a grounding event. The minimum distance to grounding over the entire mission,

$$D_{min} = \min_{t \in \{1 \ldots T\}} (d_{\text{min}}(t)),$$

serves as a clear indicator if the operation resulted in the grounding of the vessel.

Path Tracking Performance

To measure the ability of the vessel to follow the reference trajectory, the **Cross-Track Error (CTE)** is analyzed. In this simulation environment, the continuous reference path is discretized into a finite set of coordinate points. The CTE is approximated by calculating the Euclidean distance from the vessel to each of these generated points and selecting the absolute minimum distance. Let the discretized reference path be defined by a set of L points, $\mathcal{Q}_{CTE} = \{\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_L\}$, where each point is given by the coordinates $\boldsymbol{q}_l = [x_l, y_l]^\top$. If the position of the vessel at time-step t is denoted as $\boldsymbol{p}_t = [x_t, y_t]^\top$, the CTE at that specific time-step is mathematically defined as:

$$\text{CTE}(t) = \min_{l \in 1, \ldots, L} \|\boldsymbol{p}_t - \boldsymbol{q}_l\|.$$

The path-tracking accuracy can be summarized using the average CTE, which reflects general tracking precision. The average CTE is calculated as:

$$\text{CTE}_{\text{avg}} = \frac{1}{T} \sum_{t=1}^{T} \text{CTE}(t)$$

where CTE(t) is the calculated deviation at step t.

Operational Efficiency

Operational efficiency is evaluated through **Accumulative Energy Consumption** and **Travel Time**. The travel time, T_{travel}, is simply the product of the number of simulation steps, T, required to reach the destination. T_{travel} is defined as:

$$T_{\text{travel}} = T$$

Energy consumption is estimated by integrating the power usage over the duration of the voyage. The power at a specific time-step, P_{step}, is approximated as the sum of the absolute power expended in the surge, sway, and yaw direction. This is derived from the propeller force f, the rudder angle a, the lever arm l_x, and the body-fixed velocities (u, v, r):

$$P_{\text{step}}(t) = \|f \cos a \cdot u_t\| + \|f \sin a \cdot v_t\| + \| - l_x f \sin a \cdot r_t\|$$

The Accumulative Energy Consumption, E_{total}, is then obtained via discrete-time integration:

$$E_{\text{total}} = \sum_{t=1}^{T-1} P_{\text{step}}(t) \cdot \Delta t$$

The operational efficiency highlights the cost of safety. It allows for the analysis of whether the MT scheme demands significantly more energy and travel time to maintain a valid fail-safe action compared to the ST scheme.

Computational Costs

Finally, the **Solver Time** is recorded to compare the computational costs of optimizing two trajectories simultaneously. This metric measures the time required to find an optimal solution at each optimization time-step. Since the proposed MT scheme involves solving a larger optimization problem (almost doubling the decision variables for two trajectories), it is expected to measure an increase in solver time. The computational cost is summarized using the average solver time, which is defined as:

$$t_{\text{comp, avg}} = \frac{1}{T} \sum_{t=1}^{T} t_{\text{solve}}(t)$$

where $t_{\text{solve}}(t)$ is the duration of the optimization process at step t.

4.2 Simulation Environment

The MT scheme is implemented in a MATLAB environment. At each time-step, the optimization problem is constructed and solved for the prediction horizon. The `CasADi` framework is used to define the optimization problem, which is then solved using the Interior Point OPTimizer (`IPOPT`).

To improve computational efficiency and aid solver convergence, a **warm start** strategy is used. The solution from the previous optimization time-step is stored. This solution is then shifted forward by one time-step and used as the initial guess for the decision variables in the current optimization problem. By providing the MPC with an initial guess that is close to the new optimal solution, the number of iterations required to converge is significantly reduced.

If the fail-safe trajectory is predicted to result in grounding in the fail-safe mode, the vessel has entered an unrecoverable state at this point. Normally, the optimization would stop and a remote operator will be requested or an anchor is dropped. In the simulations, the vessel is allowed to continue until grounding. In the nominal mode ($B = 1$) and for the ST scheme, the simulation stops if the distance between the vessel and the last waypoint is below a threshold (D_{th}). This is considered a mission success. Conversely, in the fail-safe mode ($B = 0$), the objective shifts from reaching the last waypoint to moving away from obstacles and reducing the velocity of the vessel. The simulation terminates in this mode if the velocity of the vessel drops below the threshold (V_{th}). Finally, the simulation also terminates if there is a solver failure without a valid backup plan for the MT scheme in the fail-safe mode and the ST scheme and if the maximum number of time-steps (T_{max}) is reached.

4.2.1 Steering and Propulsion Faults

In this work, two distinct steering and propulsion faults are considered. These include a stuck-rudder and loss of effectiveness fault. Normally, a fault diagnosis module detects, isolates and quantifies the faults. The design and integration of a fault diagnosis module fall outside the scope of this work. Instead, this work assumes an ideal fault diagnosis. Consequently, upon the occurrence of a fault, the exact location, type, and magnitude are assumed to be instantaneously and perfectly known. The occurrence of these faults will be modeled as constraints imposed on the control input vector $\boldsymbol{u}_k^I$ and $\boldsymbol{u}_k^{II}$ (see chapter 2).

During the stuck-rudder fault, the following constraints for the rudder angle in the nominal and fail-safe trajectory are included:

$$a_k^I = a_{\text{stuck}} \tag{4.1}$$

$$a_k^{II} = a_{\text{stuck}} \tag{4.2}$$

In the event that a loss of effectiveness fault occurs, the maximum available propeller force is reduced to a new lower level f_{LoE}, where $f_{LoE} < f_{limit}$. During the loss of effectiveness fault, the following constraints for the propeller force in the nominal and fail-safe trajectory are included:

$$-f_{LoE} \leq f_k^I \leq f_{LoE} \tag{4.3}$$

$$-f_{LoE} \leq f_k^{II} \leq f_{LoE} \tag{4.4}$$

These constraints on the control input vector $\boldsymbol{u} = [f, a]^\top$, affect the control force vector $\boldsymbol{\tau}$. During a stuck-rudder fault, locking the rudder angle at a_{stuck} freezes the trigonometric decomposition of the propeller force. This results in surge force $f\cos(a_{\text{stuck}})$ with sway force $f\sin(a_{\text{stuck}})$ and yaw moment $-l_x f\sin(a_{\text{stuck}})$, depriving the vessel of steering control. During a loss of effectiveness fault, the propeller force is capped at f_{LoE}, directly affecting the maximum magnitude of control force vector $\boldsymbol{\tau}$. According to the kinetic equation $\dot{\boldsymbol{\nu}} = \boldsymbol{M}^{-1}(\boldsymbol{\tau} + \boldsymbol{d} - \boldsymbol{D}\boldsymbol{\nu})$, restricting $\boldsymbol{\tau}$ limits the body-fixed accelerations $\dot{\boldsymbol{\nu}}$. With diminished control, the vessel has reduced capabilities to overcome the resistance from the damping matrix $\boldsymbol{D}$ and the external disturbance vector $\boldsymbol{d}$.

4.3 Initialization and Parameter Settings

This section outlines the initial state of the vessel and the key tuning parameters, which are summarized in table 4.1.

The simulation starts at $t = 1$ with the vessel positioned at the first waypoint of the reference path, denoted as $\boldsymbol{w}_1 = [x_1, y_1]^\top$. To ensure a smooth entry into the simulated segment, the initial heading is set to strictly align with the orientation of the first path segment. This value is calculated using the four-quadrant inverse tangent of the vector connecting the first two waypoints, ensuring the vessel points directly towards $\boldsymbol{w}_2$. The initial surge velocity, u_1, is initialized to the reference speed (s_{ref}), operating under the assumption that the vessel has already attained the desired transit speed prior to entering the simulated segment. The initial sway velocity (v_1) and yaw rate (r_1) are both set to zero, representing a state with no lateral drift or turning motion at the start of the simulation. Furthermore, the control input from the preceding time-step ($t = 0$) is initialized to zero, i.e. $\boldsymbol{u}_0^I = [0, 0]^T$.

The MPC is configured with a prediction horizon of $N = 10$ steps. The duration of each time step is set to $\Delta t = 30\,\text{s}$, resulting in a total look-ahead time of 5 minutes. This configuration was empirically determined to provide a sufficient planning horizon for the MPC, while maintaining a computationally feasible optimization problem. The vessel parameters and maximum propeller force used in this work are based on the cargo vessel specifications presented in [1]. Based on the principles of ship maneuverability outlined in [20], the efficiency of a rudder is significantly compromised at high deflection angles due to flow separation. As the rudder angle increases beyond approximately $35°$, the flow pattern transitions from laminar to turbulent, leading to stall. To prevent operating in this stalled condition, rudder limits are typically implemented. Consistent with these limitations, the simulation constraints in this work restrict the maximum rudder angle to 0.2π rad ($\approx 36°$).

The tuning of the MPC objective function weights was performed empirically and based on priority. The highest priority is assigned to avoiding grounding, followed by bringing the vessel to a stop at the end of the fail-safe trajectory, mission progression, path tracking accuracy, and finally, control effort. Consequently, the risk-related weights, denoted by μ, are set higher than the velocity-related weight (Ω). The velocity-related weight is set to magnitudes larger than the path progression weights (κ), which in turn are orders of magnitude larger than the control input weights ($\boldsymbol{\Lambda}, \boldsymbol{\Delta}$). This scaling ensures that the MPC focuses primarily on maintaining safety and being able to bring the vessel to a stop.

Parameter	Symbol	Value	Unit
MPC Tuning Parameters			
Prediction Horizon	N	10	steps
Time Step Duration	Δt	30	s
Reference Path Speed	s_{ref}	2.0	m/s
Initial Conditions			
Initial Surge Velocity	u_0	2	m/s
Initial Sway Velocity	v_0	0.0	m/s
Initial Yaw Rate	r_0	0.0	rad/s
Initial Thruster Force	f_0	0.0	kN
Initial Thruster Angle	a_0	0.0	rad
Vessel Physical Parameters			
Rigid-body ship mass	m	1.5×10^6	kg
Rotational inertia (yaw)	I_z	9.8×10^8	kgm^2
Overall ship length	L_{oa}	75	m
Thruster arm length	l_x	33	m
Frontal projected area	A_{Fw}	110	m^2
Lateral projected area	A_{Lw}	624	m^2
Viscous damping (surge)	X_u	50	kNs/m
Viscous damping (sway)	Y_v	200	kNs/m
Viscous damping (yaw)	N_r	3.0×10^4	kNs/rad
Added mass (surge)	$X_{\dot{u}}$	4.4×10^4	kg
Added mass (sway)	$Y_{\dot{v}}$	8.6×10^5	kg
Added mass (yaw)	$N_{\dot{r}}$	4.9×10^7	kgm^2
Cost Function Weights			
Path Following (Position Error)	κ_1	1	-
Path Following (Progress Error)	κ_2	4.0×10^2	-
Nominal Control Effort Force (Magnitude)	Λ_{11}^I	1.0×10^{-5}	-
Nominal Control Effort Angle (Magnitude)	Λ_{22}^I	1.0×10^{-9}	-
Nominal Control Effort Force (Rate)	Δ_{11}^I	1.0×10^{-6}	-
Nominal Control Effort Angle (Rate)	Δ_{22}^I	1.0×10^{-9}	-
Fail-Safe Control Effort Force (Rate)	Δ_{11}^{II}	1.0×10^{-7}	-
Fail-Safe Control Effort Angle (Rate)	Δ_{22}^{II}	1.0×10^{-10}	-
Nominal Base Risk	μ_1^I	1.0×10^6	-
Nominal Wind-related Risk	μ_2^I	1.0×10^3	-
Fail-Safe Base Risk	μ_1^{II}	1.0×10^6	-
Fail-Safe Wind-related Risk	μ_2^{II}	1.0×10^4	-
Terminal Base Risk	μ_1^T	1.0×10^7	-
Terminal Wind-related Risk	μ_2^T	1.0×10^5	-
Risk Sensitivity (Nominal)	Ξ^I	$1/30$	-
Risk Sensitivity (Fail-Safe)	Ξ^{II}	$1/15$	-
Risk Sensitivity (Terminal)	Ξ^T	$1/20$	-
Terminal Velocity (Fail-Safe)	Ω	1.0×10^6	-
System Constraints			
Maximum Propeller Force	f_{limit}	400	kN
Maximum Rudder Angle	a_{limit}	0.2π	rad
Thresholds			
Nominal Mission Success Distance Threshold	D_{th}	100	m
Fail-Safe Mission Success Velocity Threshold	V_{th}	0.1	m/s
Supervisory Safety Velocity Threshold	v_{th}	0.2	m/s
Supervisory Safety Risk Threshold	$\Delta\rho_{th}$	1000	-
Simulation Time-Step Threshold	T_{max}	150	Steps

Table 4.1: Simulation and MT scheme parameter values

4.4 Results

This section presents the results. The results for each scenario are consistently visualized using a trajectory map and plots for the control input, velocities and KPIs. The trajectory map (e.g., figure 4.1) provides a two-dimensional view of the trajectory of the vessel within the simulated environment. The map includes elements, such as the start location (black dot), the goal (red cross), the reference path (black dashed line), and grounding obstacles (gray areas). A magenta arrow indicates the direction of the wind disturbance. The actual paths executed by the ST scheme (solid blue line) and the MT scheme (dotted orange line) are included alongside the reference path. In scenarios involving steering and propulsion faults, specific event markers will appear along these paths to indicate critical moments. Yellow triangles denote the start and end of a fault window, a red star marks a catastrophic grounding event, and a green cross indicates a successful safe stop.

The plots (e.g., figure 4.2) show the data over the duration of the simulation. The top four plots detail the surge and sway velocity and the control input. The "Propeller Force" and "Rudder Angle" plots show the control input. The "Surge Velocity" and "Sway Velocity" plots illustrate the effect of those control inputs on the vessel.

The lower four subplots detail the safety, path-tracking accuracy, operational efficiency and computational costs of the mission. The "Distance to Grounding" plot measures the exact clearance between the vessel and the nearest obstacle at any given timestep. A trajectory that intersects the solid red line at zero indicates a grounding event. The "Cross-Track Error" evaluates tracking precision by measuring the distance from the vessel to the reference path. Finally, the "Cumulative Energy Consumption" plot provides a metric for operational efficiency, where a steeper slope indicates higher power draw. As mission progression and energy efficiency are not considered when the MT scheme is operating in the fail-safe mode, the data for the energy consumption and cross-track error are not included in the plots when the MT scheme is operating in the fail-safe mode. Finally, the "Solver Time" plot shows the time required to find the optimal solution throughout the simulation.

4.4.1 No Fault Scenario and Varying Wind Velocities

The initial test scenario compares the MT and the ST scheme when the vessel is fully functional with no faults in the propulsion or steering system. The primary variable introduced is drifting forces due to wind, which exerts a force in the positive y-direction (North). This scenario is designed to assess the path-following capabilities and performance when subjected to increasing drifting forces. Three distinct wind velocities are tested: 10 m/s, 15 m/s, and 20 m/s. By increasing the wind speed, the ability to compensate for significant drifting forces, while continuing along the reference path, is tested.

10 m/s wind velocity

Figure 4.1 illustrates the trajectories generated by both the ST and MT schemes during a normal, no-fault scenario subject to a constant 10 m/s wind in the positive y-direction. The simulation results demonstrate that the behavior of both schemes is virtually indistinguishable. Both schemes successfully navigate from the starting position to the goal, effectively counteracting the drifting forces while maintaining safe clearances from the obstacles. The vessel safely deviates from the reference path to execute the necessary turns, with the ST and MT paths perfectly overlapping throughout the entire operation.

This identical performance is further confirmed by the plots presented in figure 4.2. Throughout the 53 timestep mission, both schemes apply the exact same sequence of propeller forces and rudder angles, resulting in matching surge and sway velocities. Furthermore, the distance to grounding and CTE metrics remain consistent between the two schemes. As detailed in table 4.2, both schemes maintain a minimum safety margin of approximately 122 meters from the obstacles and exhibit an average cross-track error of roughly 17.5 meters. Finally, the cumulative energy consumption for both schemes is identical, finalizing at 2.2e8 J.

The only significant difference between the two schemes lies in their computational demands. Because the MT scheme optimizes multiple trajectories, it requires substantially more computational effort. Table 4.2 shows that the average solver time for the MT scheme is 0.4030 seconds per timestep, which is nearly double the 0.2363 seconds required by the ST scheme. While figure 4.2 reveals a singular, massive spike in the ST solver time near timestep 30, the MT scheme consistently has a higher solver time across the entire mission.

The results from this no-fault scenario demonstrate that the ability to execute the mission is identical between the ST and MT schemes. Both schemes achieve the same path-tracking accuracy, maintain identical safety margins, and consume the exact same amount of energy. The only significant difference between the two schemes is the computational cost. The higher average solver time required by the MT scheme is an expected trade-off and can be attributed to the almost double the number of decision variables, which need to be solved to simultaneously optimize both the nominal and fail-safe trajectory.

Table 4.2: Results summary of the no-fault scenario with 10 m/s wind velocity

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	53	2.2×10^8	122.16	17.89	0.2363	x
MT	53	2.2×10^8	121.97	17.41	0.4030	x

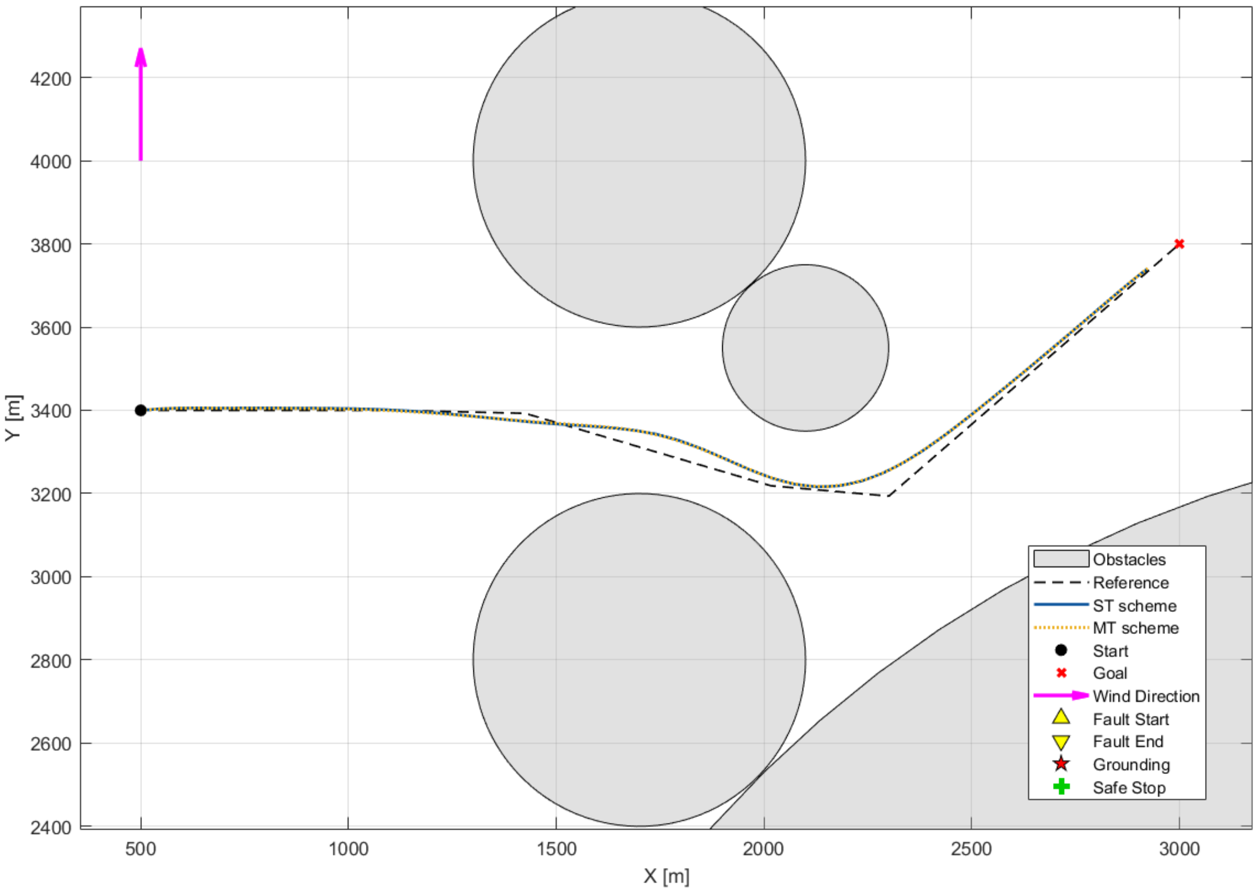


Figure 4.1: Resulting trajectories for the no-fault scenario with 10 m/s wind velocity

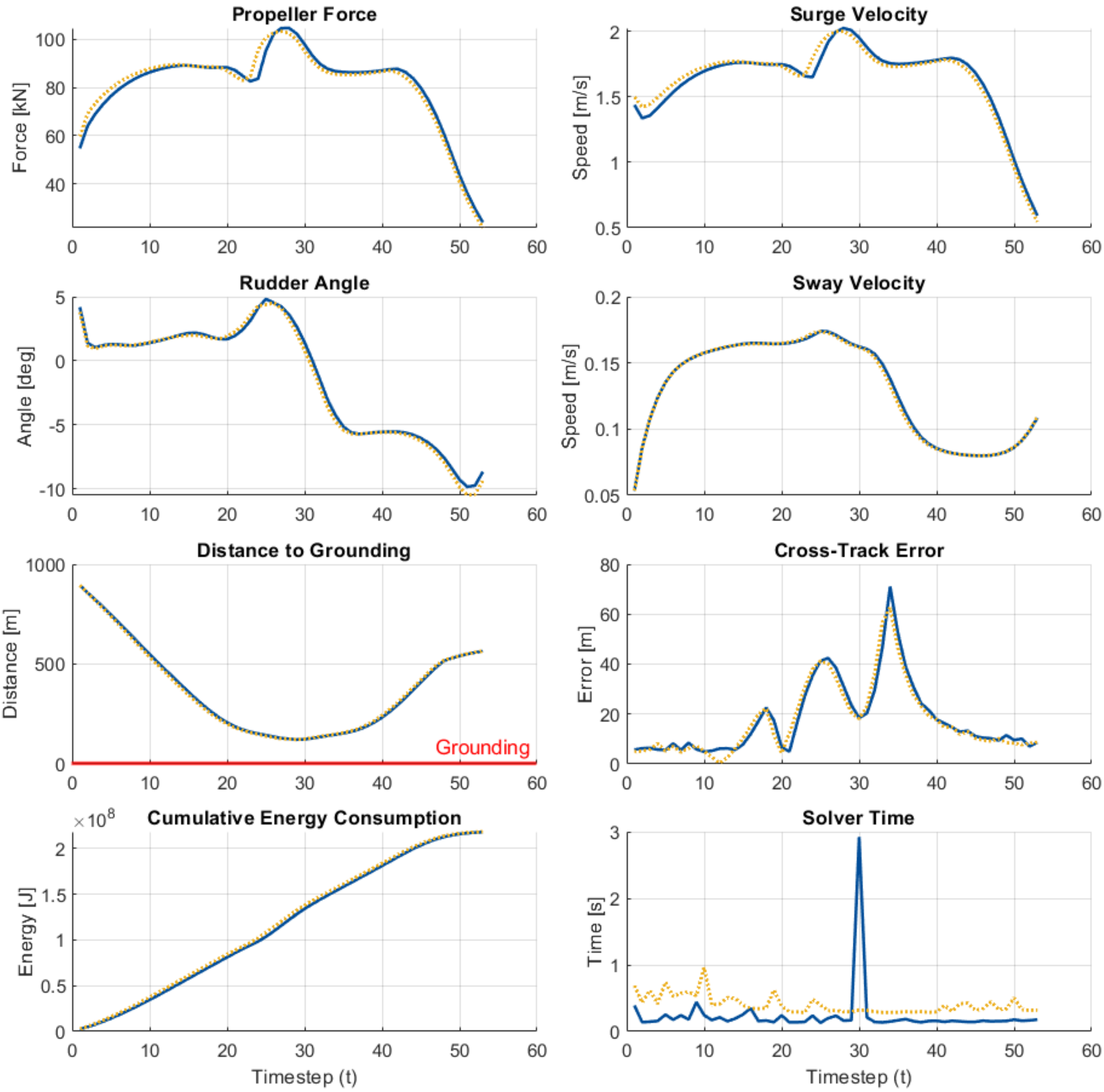


Figure 4.2: Plots for the no-fault scenario with 10 m/s wind velocity

15 m/s wind velocity

Figure 4.3 illustrates the trajectories generated by both the ST and MT schemes during a normal, no-fault scenario subject to a constant 15 m/s wind in the positive y-direction. The simulation results demonstrate that the behavior of both schemes is virtually indistinguishable. Both schemes successfully navigate from the starting position to the goal, effectively counteracting the drifting forces while maintaining safe clearances from the obstacles. The vessel safely deviates from the reference path to execute the necessary turns, with the ST and MT paths perfectly overlapping throughout the entire operation.

This identical performance is further confirmed by the plots presented in figure 4.4. Throughout the 53 step mission, both schemes apply the exact same sequence of propeller forces and rudder angles, resulting in matching surge and sway velocities. Furthermore, the distance to grounding and CTE metrics remain consistent between the two schemes. As detailed in table 4.3, both schemes maintain a minimum safety margin of approximately 122 meters from the obstacles and exhibit an average cross-track error of roughly 19.5 meters. Finally, the cumulative energy consumption for both schemes is identical, finalizing at 2.2e8 J.

The only significant difference between the two schemes lies in their computational demands. Table 4.3 shows that the average solver time for the MT scheme is 0.4437 seconds per timestep, which is nearly triple the 0.1589 seconds required by the ST scheme. Clearly visible is the spike in the solver time in the first time-step.

The warm start strategy is not applied in the first timestep, since there is no previous solution to be provided as the initial guess. This can result in a significantly longer solver time for the first timestep compared to the rest of the simulation.

The results from this no-fault scenario demonstrate that the ability to execute the mission is identical between the ST and MT schemes. Both schemes achieve almost the same path-tracking accuracy, maintain identical safety margins, and consume the exact same amount of energy. The only significant difference between the two schemes is the computational cost.

Table 4.3: Results summary of the no-fault scenario with 15 m/s wind velocity

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	53	2.2×10^8	122.23	19.84	0.1589	x
MT	53	2.2×10^8	121.99	19.21	0.4437	x

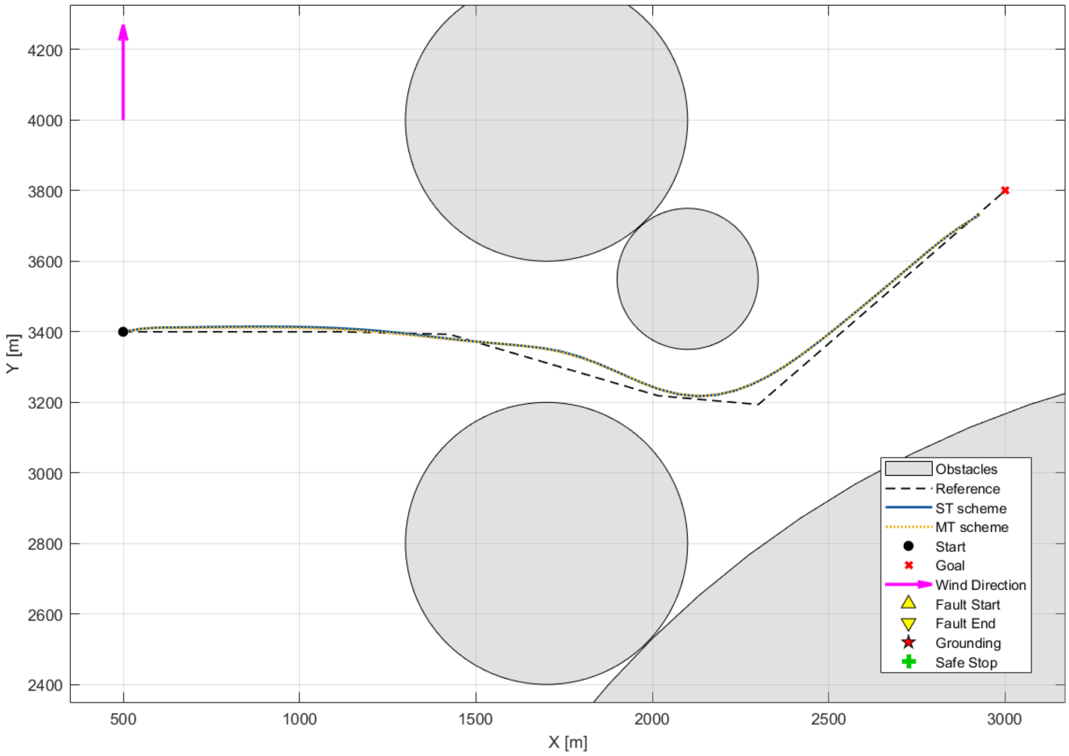


Figure 4.3: Resulting trajectories for the no-fault scenario with 15 m/s wind velocity

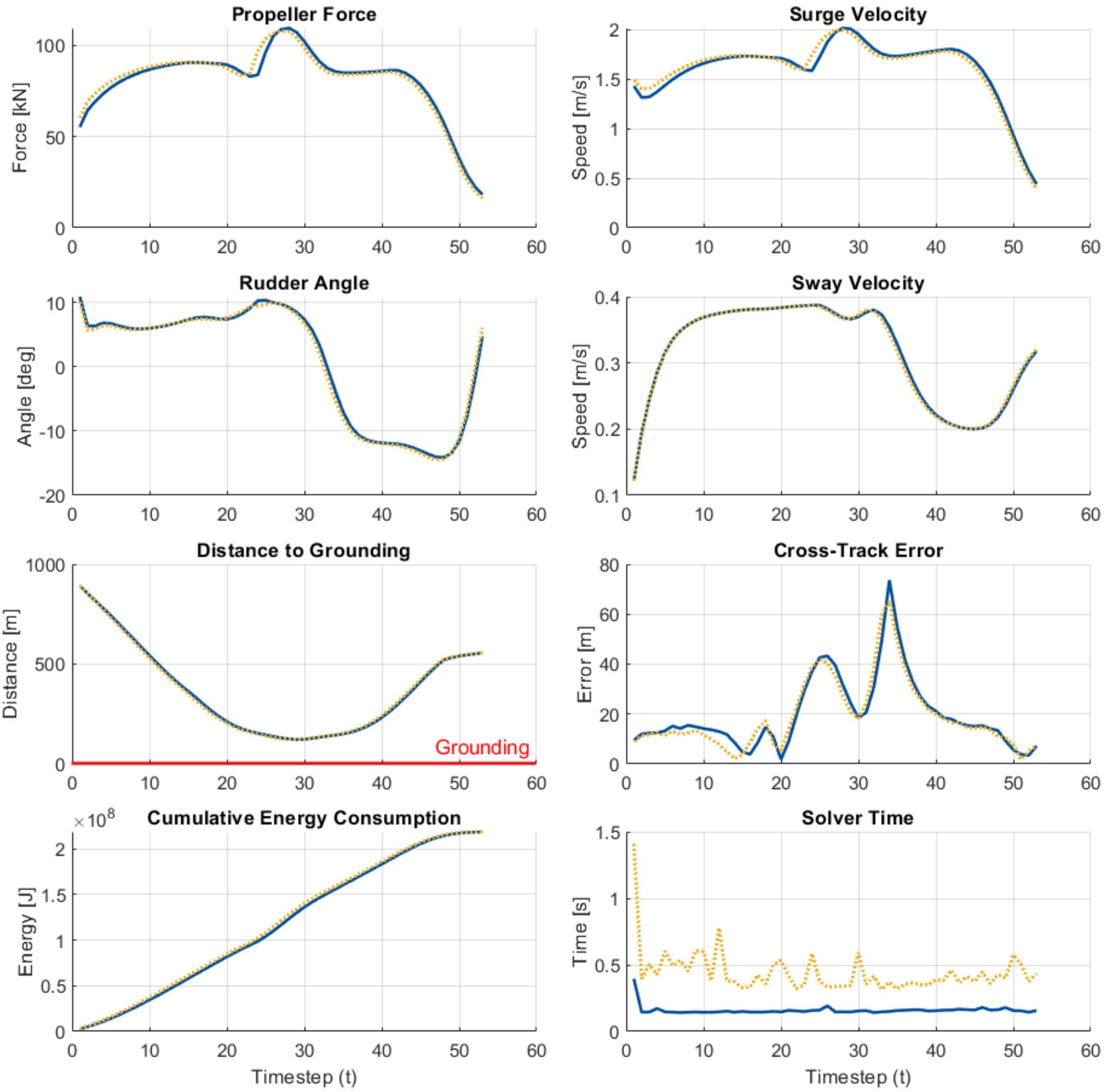


Figure 4.4: Plots for the no-fault scenario with 15 m/s wind velocity

20 m/s wind velocity

Figure 4.5 illustrates the trajectories generated by both the ST and MT schemes during a normal, no-fault scenario subject to a constant 20 m/s wind in the positive y-direction. The simulation results demonstrate that the behavior of both schemes is virtually indistinguishable. Both schemes successfully navigate from the starting position to the goal, effectively counteracting the drifting forces while maintaining safe clearances from the obstacles. The vessel safely deviates from the reference path to execute the necessary turns, with the ST and MT paths perfectly overlapping throughout the entire operation.

This identical performance is further confirmed by the plots presented in figure 4.6. Throughout the mission, both schemes apply the exact same sequence of propeller forces and rudder angles, resulting in matching surge and sway velocities. Furthermore, the distance to grounding and CTE metrics remain consistent between the two schemes. As detailed in table 4.4, both schemes maintain a minimum safety margin of approximately 122 meters from the obstacles. The average CTE for the MT scheme is lower than the average CTE for the ST scheme (23.53 m vs 26.01 m). Finally, the cumulative energy consumption for both schemes is identical, finalizing at 2.2e8 J.

The only significant difference between the two schemes lies in their computational demands. Table 4.4 shows that the average solver time for the MT scheme is 0.4531 seconds per timestep, which is nearly triple the 0.1548 seconds required by the ST scheme.

The results from this no-fault scenario demonstrate that the ability to execute the mission is identical between the ST and MT schemes. Both schemes maintain almost identical safety margins, and consume the exact same amount of energy. The only significant difference between the two schemes is the computational cost.

Table 4.4: Results summary of the no-fault scenario with 20 m/s wind velocity

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	54	2.2×10^8	121.36	26.01	0.1548	x
MT	53	2.2×10^8	122.22	23.52	0.4531	x

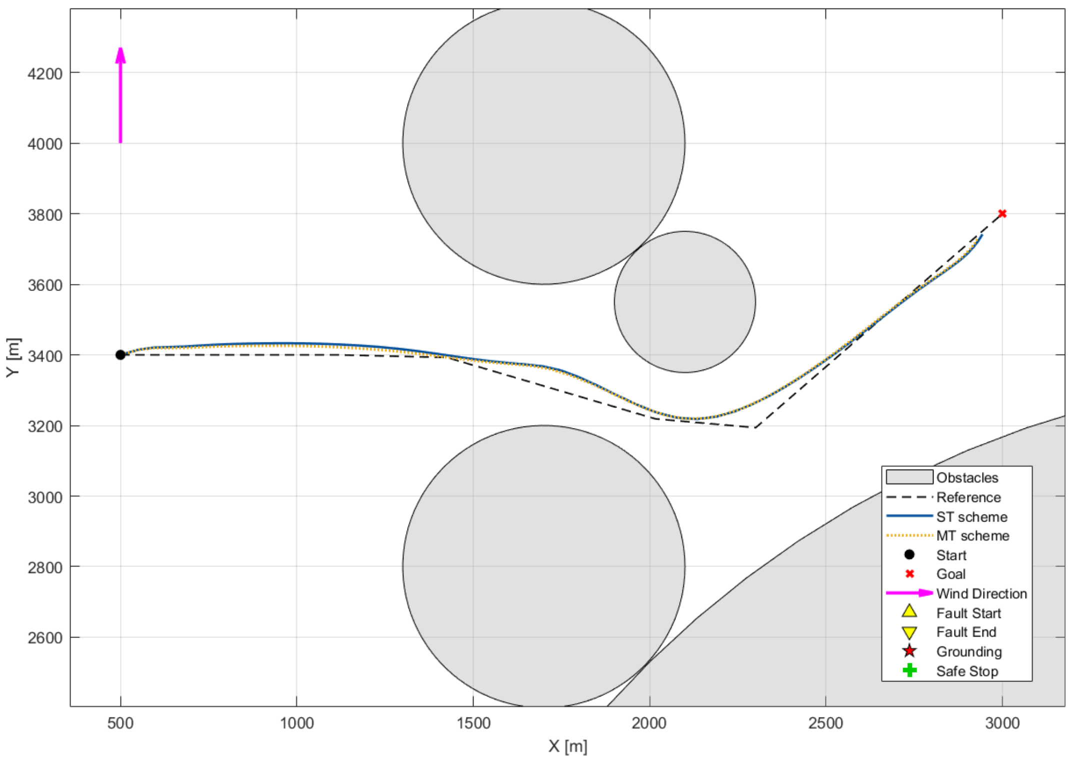


Figure 4.5: Resulting trajectories for the no-fault scenario with 20 m/s wind velocity

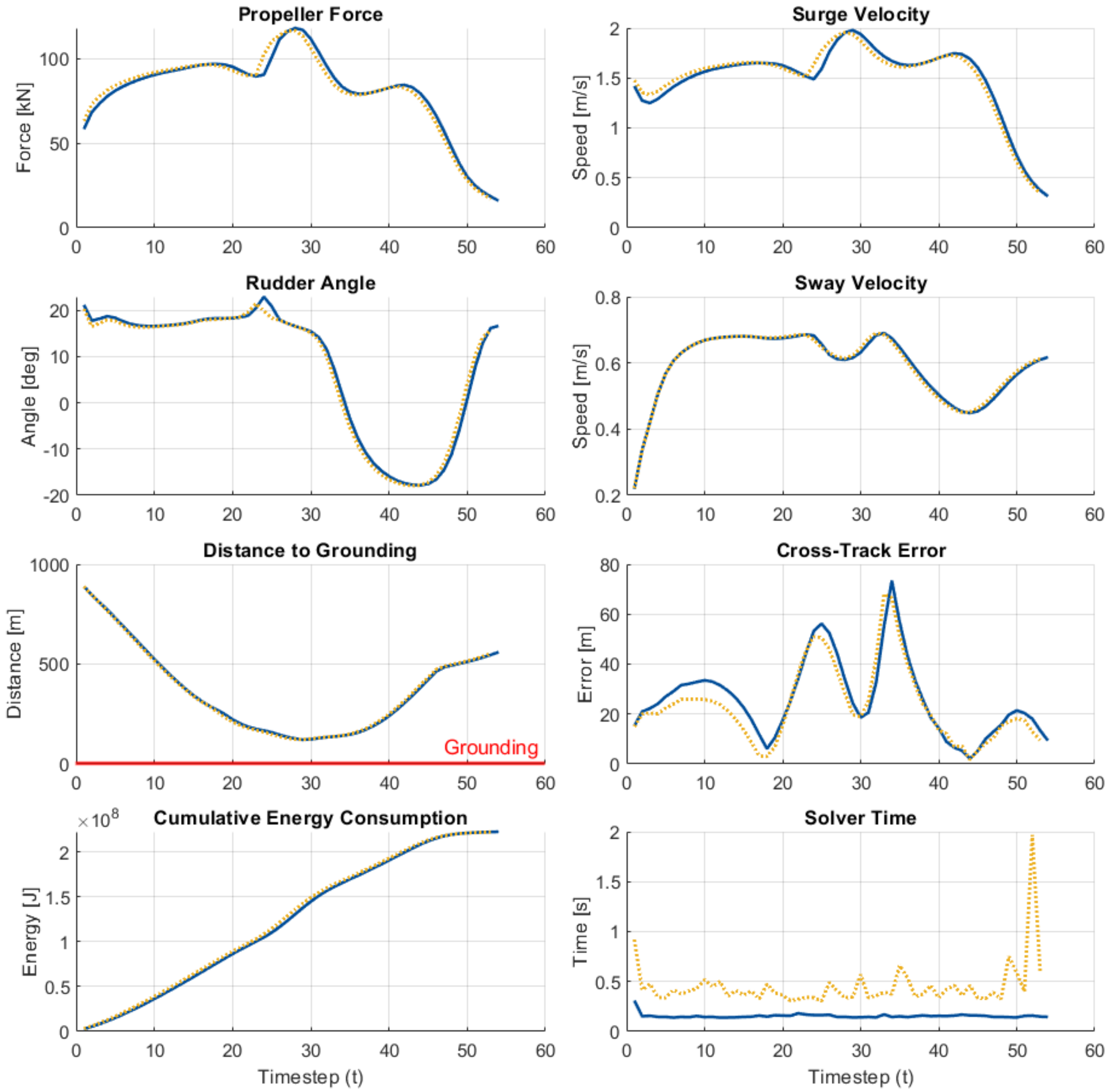


Figure 4.6: Plots for the no-fault scenario with 20 m/s wind velocity

4.4.2 Stuck Fault Scenario and Varying Wind Velocities

This section evaluates the enhanced safety of the MT scheme compared to the ST scheme during a stuck rudder fault scenario (fault occurring between t=1 and t=50) for varying wind velocities (0 and 10 m/s).

No wind

Figure 4.7 illustrates the trajectories for the scenario with no wind, where a stuck rudder fault occurs immediately. The fault begins at $t = 1$ and continues until $t = 50$, during which the rudder is completely stuck at a neutral $0°$ angle. As observed in the trajectory map, the vessel is unable to execute the required turn to follow the reference path, forcing the vessel to temporarily hold position until the fault is resolved. The plots in figure 4.8 reveal how both the ST and MT scheme handle this loss of steering capability. Recognizing the inability to turn, both schemes reduce the propeller force to zero. Consequently, the surge velocity steadily comes down from roughly 1.5 m/s to near zero over the duration of the fault. By holding position, both schemes successfully prevent grounding.

At timestep 50, full functionality is restored. The two schemes behave different after the fault is resolved. The ST scheme aggressively reverses the vessel, characterized by a negative propeller force and surge velocity. In contrast, the MT scheme applies a positive propeller force, resulting in a positive surge velocity. Despite these entirely opposite approaches, both schemes successfully navigate around the obstacle and reach the goal.

Table 4.5 highlights the trade-offs between these two schemes. The aggressive reversing maneuver of the ST scheme yields a slightly better safety margin, maintaining a minimum distance to grounding (D_{min}) of 53.29 meters, compared to the 34.04 meters for the MT scheme. However, the MT scheme is slightly more energy-efficient, consuming 3.1×10^8 J versus the 3.3×10^8 J with the ST scheme. Because the vessel spent a significant duration of the mission holding position away from the reference path during the fault, both schemes naturally share a high average cross-track error compared to the no-fault scenarios of approximately 67 meters. Total travel time is virtually identical, taking 68 steps for the ST scheme and 69 steps for the MT scheme.

Finally, the computational costs of the two schemes differ significantly. As shown in the solver time plot in figure 4.8, the ST scheme maintains a low solver time throughout the entire simulation when compared to the MT scheme. Overall, the average solver time for the MT scheme (0.1994 s) is more than double that of the ST scheme (0.0810 s).

These results demonstrate the ability of the MT scheme to limit the activation of the fail-safe action. This behavior can be directly attributed to the inclusion of the risk cost term within the optimization of the nominal trajectory itself. Rather than blindly pushing forward into the grounding obstacles, the nominal trajectory reduces velocity and holds position as the vessel approaches the obstacles. Furthermore, because there are no wind-induced drifting forces in this scenario, the vessel maintains the capability to come to a complete stop in the fail-safe trajectory. Since this stopping capability is never compromised, the supervisory system correctly allows the vessel to remain in the nominal mode throughout the fault window without triggering the fail-safe action. Ultimately, both schemes successfully completed the mission under these conditions. The primary distinctions between the two are the higher safety margin achieved by the ST scheme, due to the aggressive reversing maneuver, and the increased computational cost of the MT scheme, which requires a higher average solver time to optimize nearly double the number of decision variables.

Table 4.5: Results summary of the stuck fault scenario with no wind and the fault between t=1 and t=50

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	68	3.3×10^8	53.29	67.15	0.0810	x
MT	69	3.1×10^8	34.04	67.69	0.1994	x

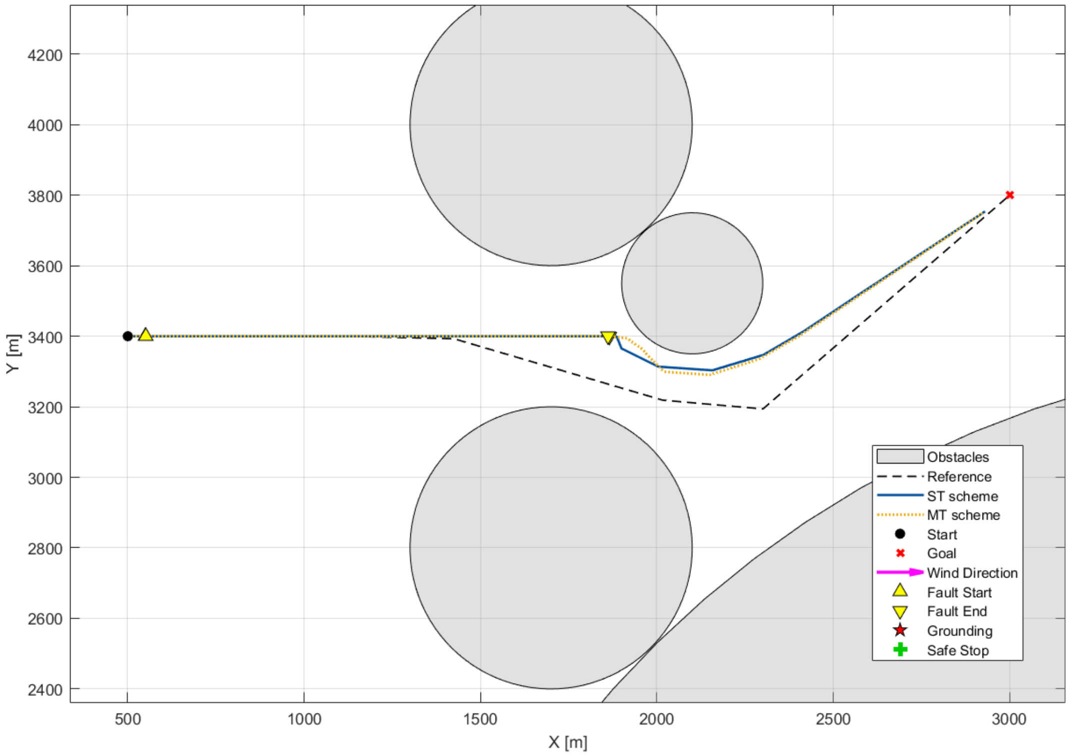


Figure 4.7: Resulting trajectories for the stuck fault scenario with no wind and the fault between t=1 and t=50

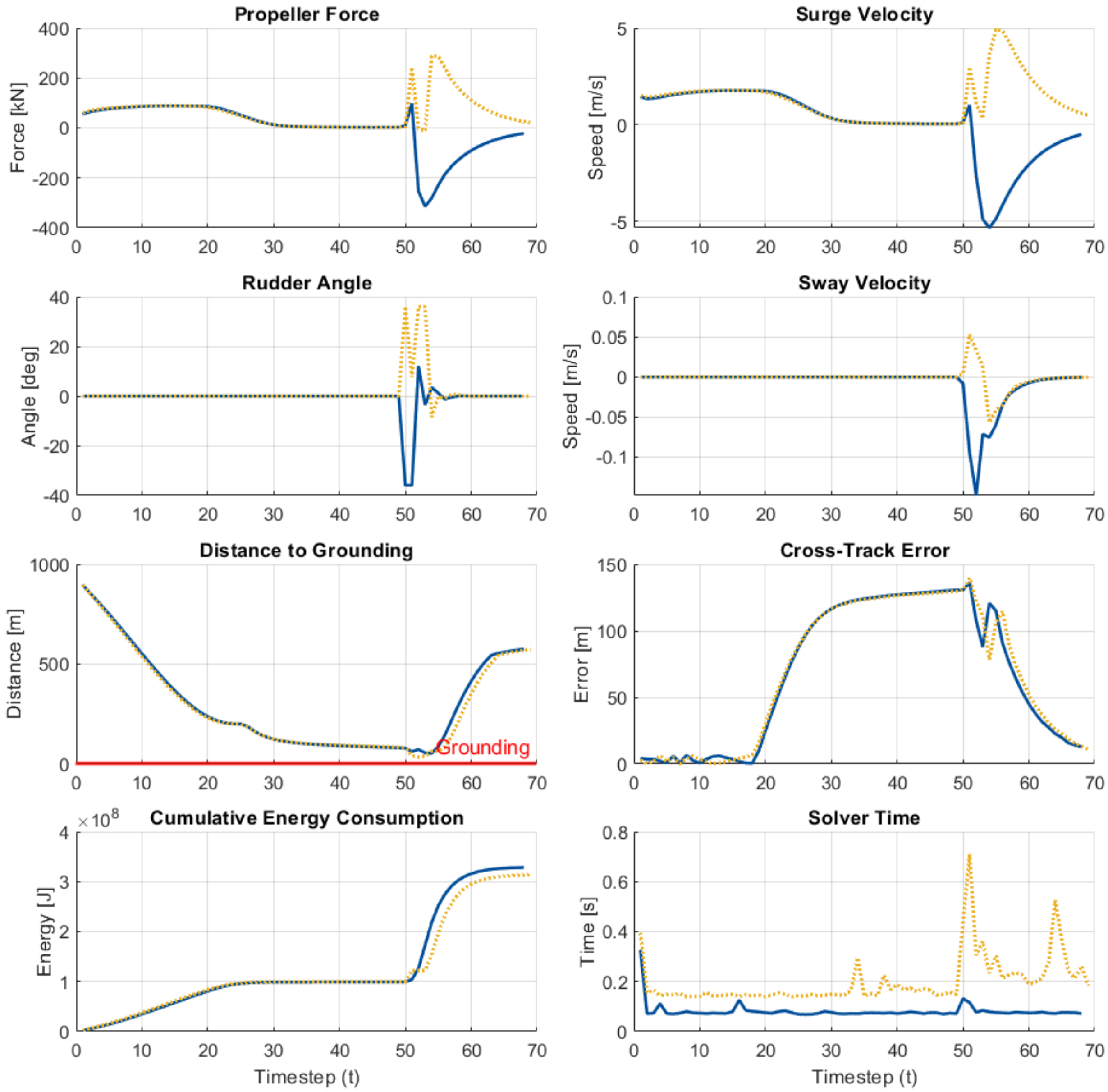


Figure 4.8: Plots for the stuck fault scenario with no wind and the fault between t=1 and t=50

10 m/s wind velocity

Figure 4.9 presents a scenario where the vessel experiences a fault (rudder stuck at a neutral $0°$) from $t = 1$ to $t = 50$, while simultaneously being exposed to 10 m/s crosswind. As indicated by the magenta wind vector, the environmental disturbance steadily pushes the vessel in the positive y direction, directly toward the obstacles. The ST scheme completely fails to manage this situation. As illustrated in the plots (figure 4.10), the ST scheme commands a propeller force, which exceeds 300 kN, at timestep 20, causing the surge velocity to aggressively spike above 5 m/s. Unable to steer, this acceleration drives the vessel directly into the grounding obstacle, resulting in grounding at $t = 24$. Table 4.6 confirms this grounding event, recording a negative minimum distance to grounding of -101.62 m.

In stark contrast, the MT scheme successfully executes a fail-safe action to prevent grounding. The MT scheme initially reduces the velocity of the vessel as it is approaching the obstacles. However, at approximately timestep 34, the fail-safe action is activated. The MT scheme commands a negative propeller force of -400 kN to violently reverse the vessel. As seen in the surge velocity plot, the vessel accelerates backwards. This aggressive reversal successfully prevents grounding, maintaining a minimum clearance of 42.23 meters from the obstacles, as noted in table 4.6. As the distance to the obstacles increases, the velocity of the vessel is reduced. When the stuck rudder fault is resolved at $t = 50$, the MT scheme uses the regained rudder control to bring the vessel to a stop.

This scenario demonstrates the ability of the MT scheme to prevent grounding in scenarios where the ST scheme fails to prevent grounding. The supervisory system successfully intervenes and activates the fail-safe action. Additionally, these results highlight the adaptability of the fail-safe action. During the active fault, when the 10 m/s crosswind and the fixed rudder angle make a complete stop impossible, the MT scheme focuses entirely on moving the vessel away from the obstacles, making sure the vessel drifts in open waters. It is only after the fault is resolved at $t = 50$ that the focus shifts to bringing the vessel to a controlled stop. This enhanced safety does result in an increased computational cost, with the MT scheme requiring an average solver time of 0.2876 seconds compared to the 0.1387 seconds for the ST scheme.

Table 4.6: Results summary of the stuck fault scenario with 10 m/s wind velocity and the fault between t=1 and t=50

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	-	-	-101.62	-	0.1387	-
MT	-	-	42.23	-	0.2876	-

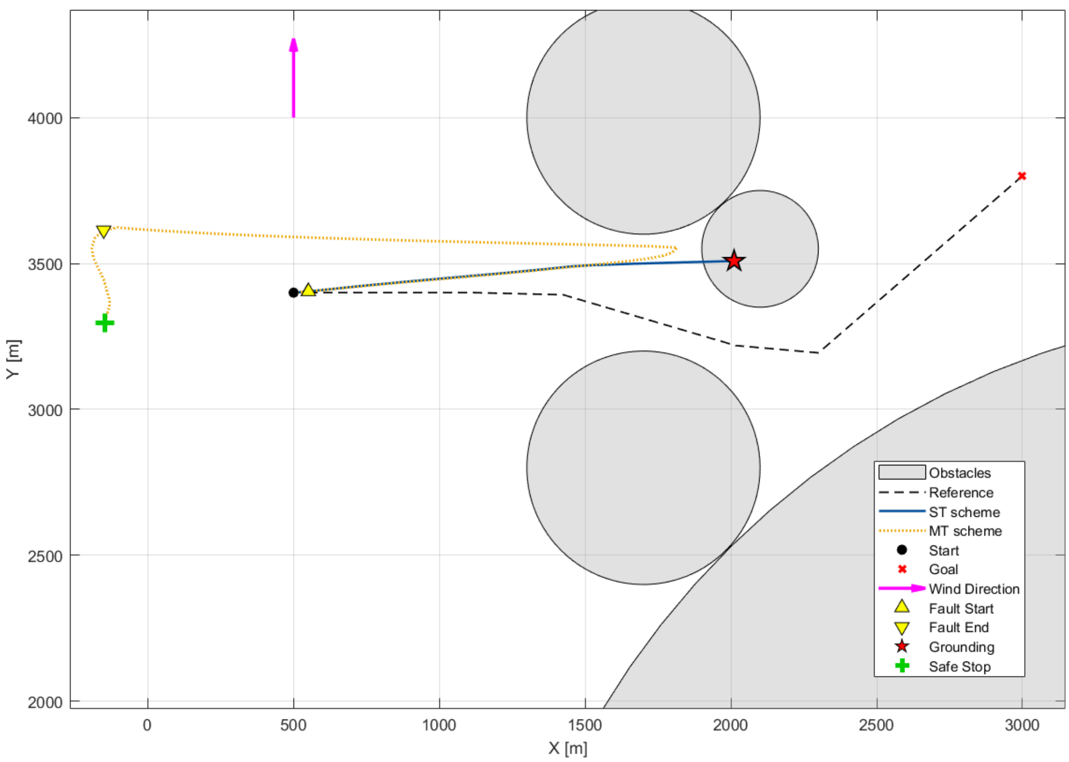


Figure 4.9: Resulting trajectories for the stuck fault scenario with 10 m/s wind velocity and the fault between t=1 and t=50

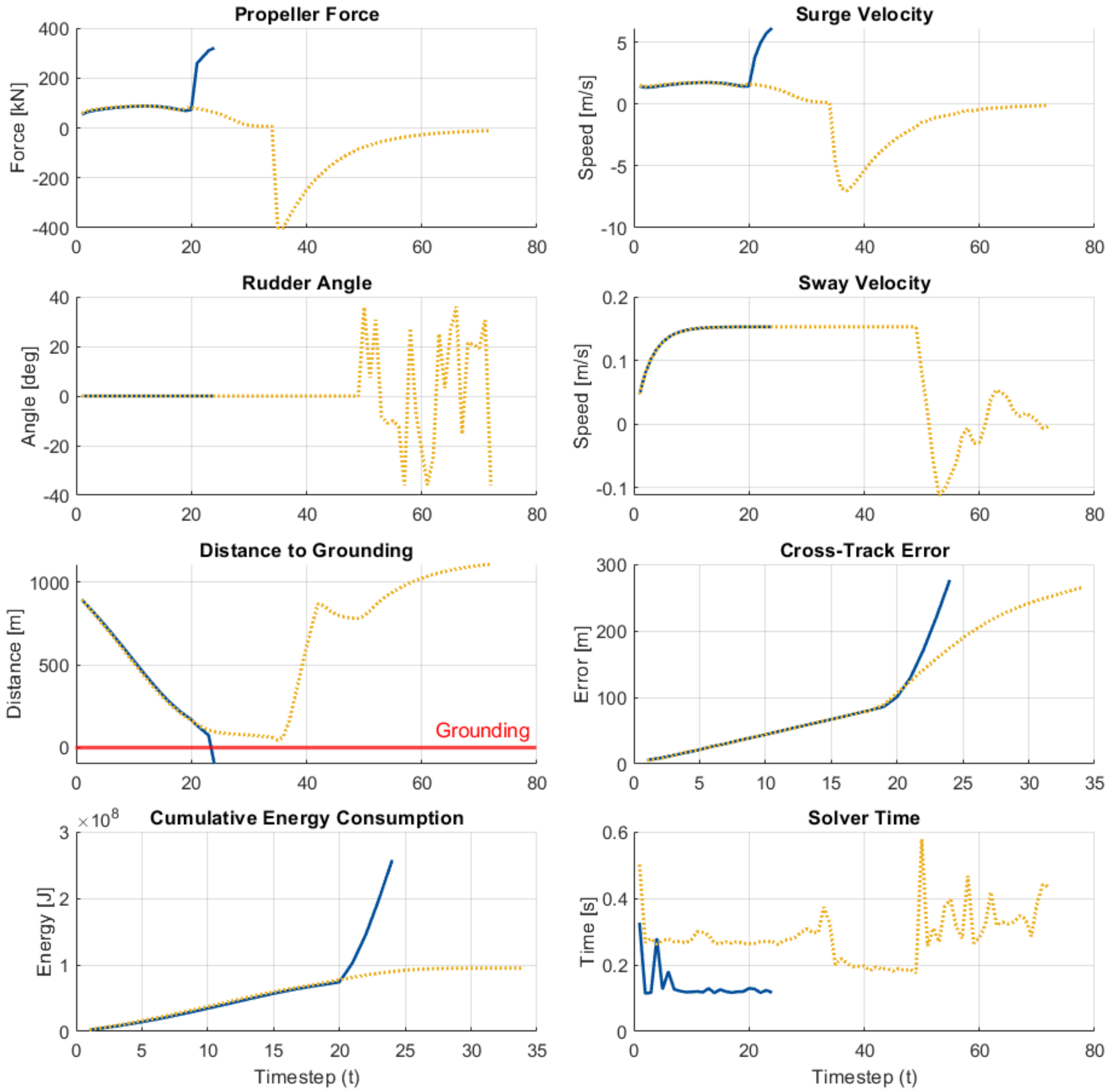


Figure 4.10: Plots for the stuck fault scenario with 10 m/s wind velocity and the fault between t=1 and t=50

4.4.3 Loss of Effectiveness Fault Scenario and Varying Moments of Fault Initiation

This section evaluates the performance of the ST and MT schemes under the loss of effectiveness fault. The loss of effectiveness fault severely restricts the capabilities of the vessel by reducing the maximum available propeller force from the normal 400 kN down to just 30 kN ($f_{LoE} = 30\ kN$). To assess the performance of both schemes, the schemes are tested under 20 m/s wind velocity with different moments of fault initiation ($t = 25$ to 70, $t = 30$, and $t = 35$ to 70).

Fault at t=25

Figure 4.11 depicts a scenario where the vessel is exposed to a 20 m/s wind in the positive y-direction, compounded by a loss of effectiveness fault that occurs at $t = 25$ and is resolved at $t = 70$. This fault occurs when the vessel is deep within the narrow passage. As demonstrated in figure 4.12, the ST scheme is completely unable to handle this situation. While the ST scheme attempts to fight the wind and starts oscillating between maximum positive and negative rudder angles, the vessel is unable to prevent drifting. This leads to a grounding event at approximately $t = 44$, marked by the red star on the map and confirmed by a negative minimum distance to grounding (-10.52 m) in table 4.7.

In stark contrast, the MT scheme activates the fail-safe action immediately after the fault occurs as the vessel is unable to bring itself to a stop and drifts towards grounding obstacles. The MT scheme commands a negative propeller force, dropping the surge velocity below zero. Instead of fighting the wind and trying to continue the mission, the MT scheme utilizes the degraded steering and propulsion system to reverse out of the narrow passage. This action prevents the vessel from grounding, maintaining a minimum distance to grounding of 57.16 meters. After the fault is resolved at $t = 70$, the MT scheme brings the vessel to a safe stop (green cross) in open water. This enhanced safety comes at a computational cost. The MT scheme requires an average solver time of 0.7008 seconds, more than triple the 0.2113 seconds of the ST scheme.

Table 4.7: Results summary of the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault between t=25 and t=70

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	-	-	-10.52	-	0.2113	-
MT	-	-	57.16	-	0.7008	-

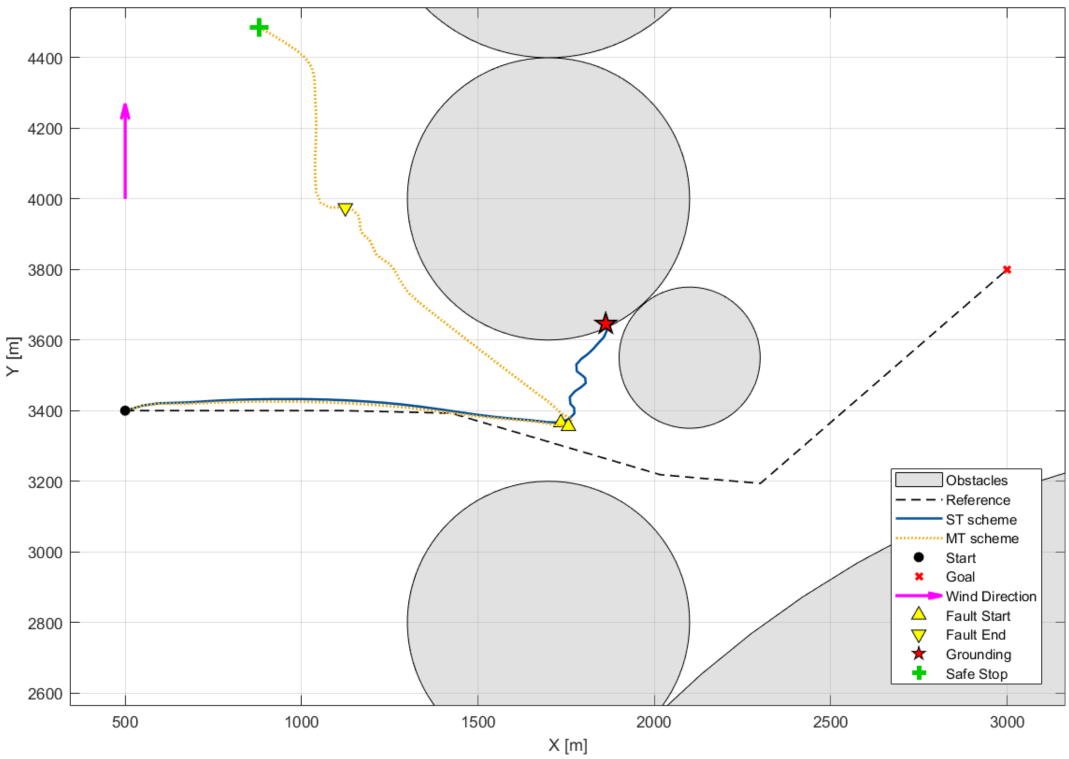


Figure 4.11: Resulting trajectories for the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault between t=25 and t=70

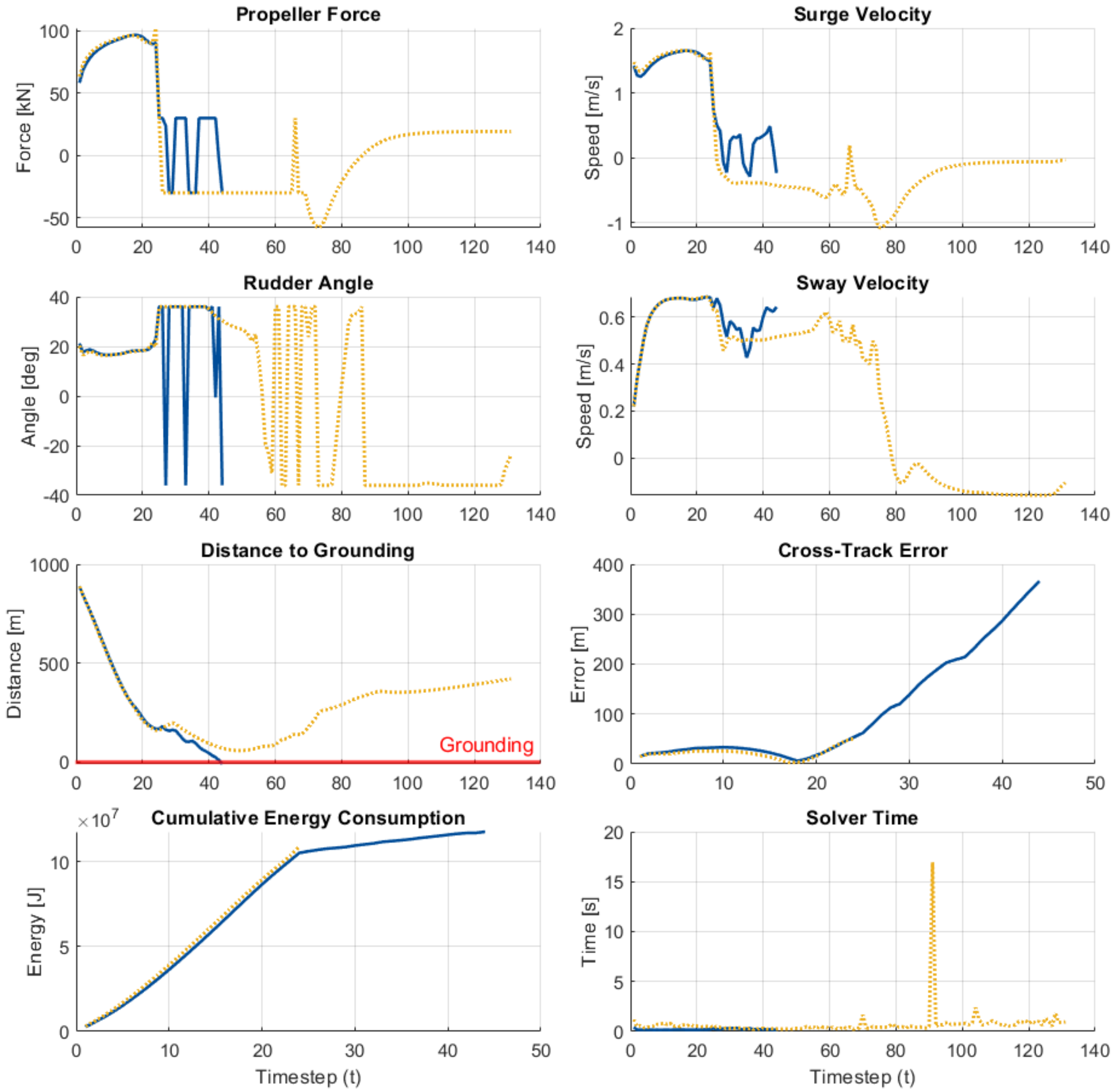


Figure 4.12: Plots for the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault between t=25 and t=70

Fault at t=30

Figure 4.13 illustrates a scenario with 20 m/s wind in the positive y-direction and a loss of effectiveness fault that occurs at $t = 30$. Unlike the previous scenario where the fault occurred at $t = 25$ (allowing the vessel just enough time to reverse out of the narrow passage), at $t = 30$ the fault occurs when the vessel is approximately at the deepest, narrowest point inside the narrow passage. The fail-safe action is immediately activated after the fault occurs. Although the MT scheme attempts to guide the vessel outside the narrow passage, it is physically unable to do so before grounding occurs.

The plots in figure 4.14 detail the responses of both schemes. Both the ST and MT scheme immediately drop their propeller force to the maximum available negative value (approximately -30 kN under degraded conditions) to reverse the vessel. After passing the first obstacle, the ST scheme stops reversing out of the narrow passage and drifts towards the dead end between two obstacles where it grounds.

This scenario highlights a limitation for the proposed MT scheme. It demonstrates that the MT scheme is bound by the limits of the vessel. If a fault degrades the vessel such that it is physically impossible to exit the narrow passage, grounding is inevitable.

Table 4.8: Results summary of the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault at t=30

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	-	-	-13.50	-	0.2065	-
MT	-	-	-3.43	-	0.3134	-

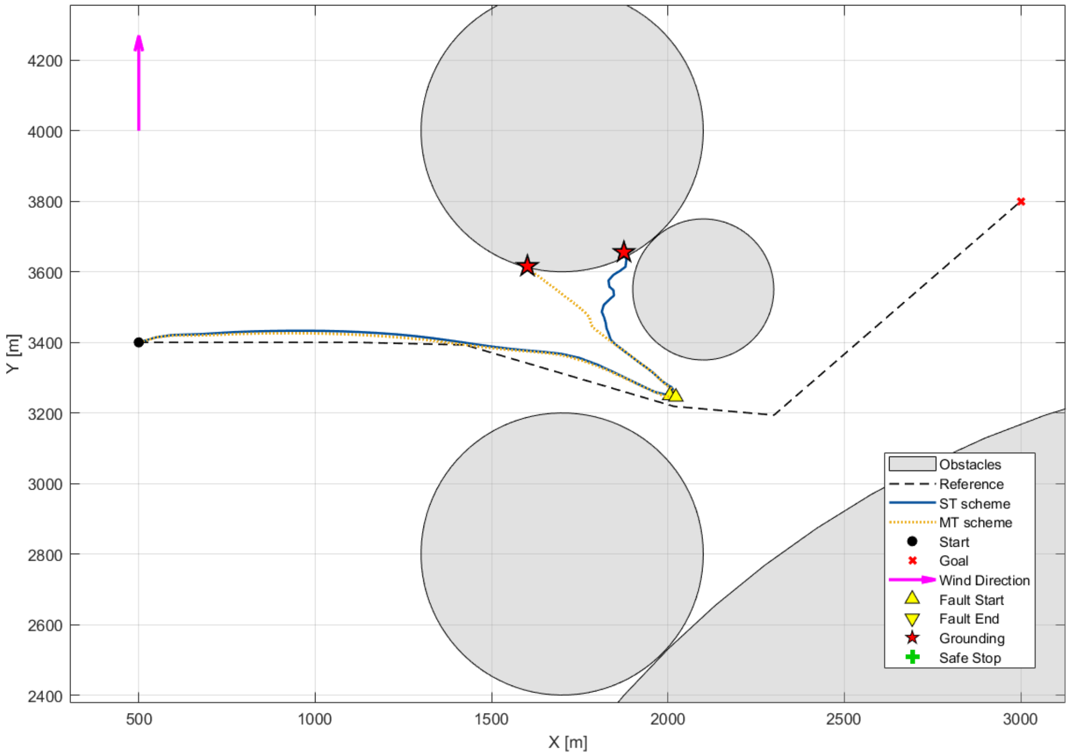


Figure 4.13: Resulting trajectories for the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault at t=30

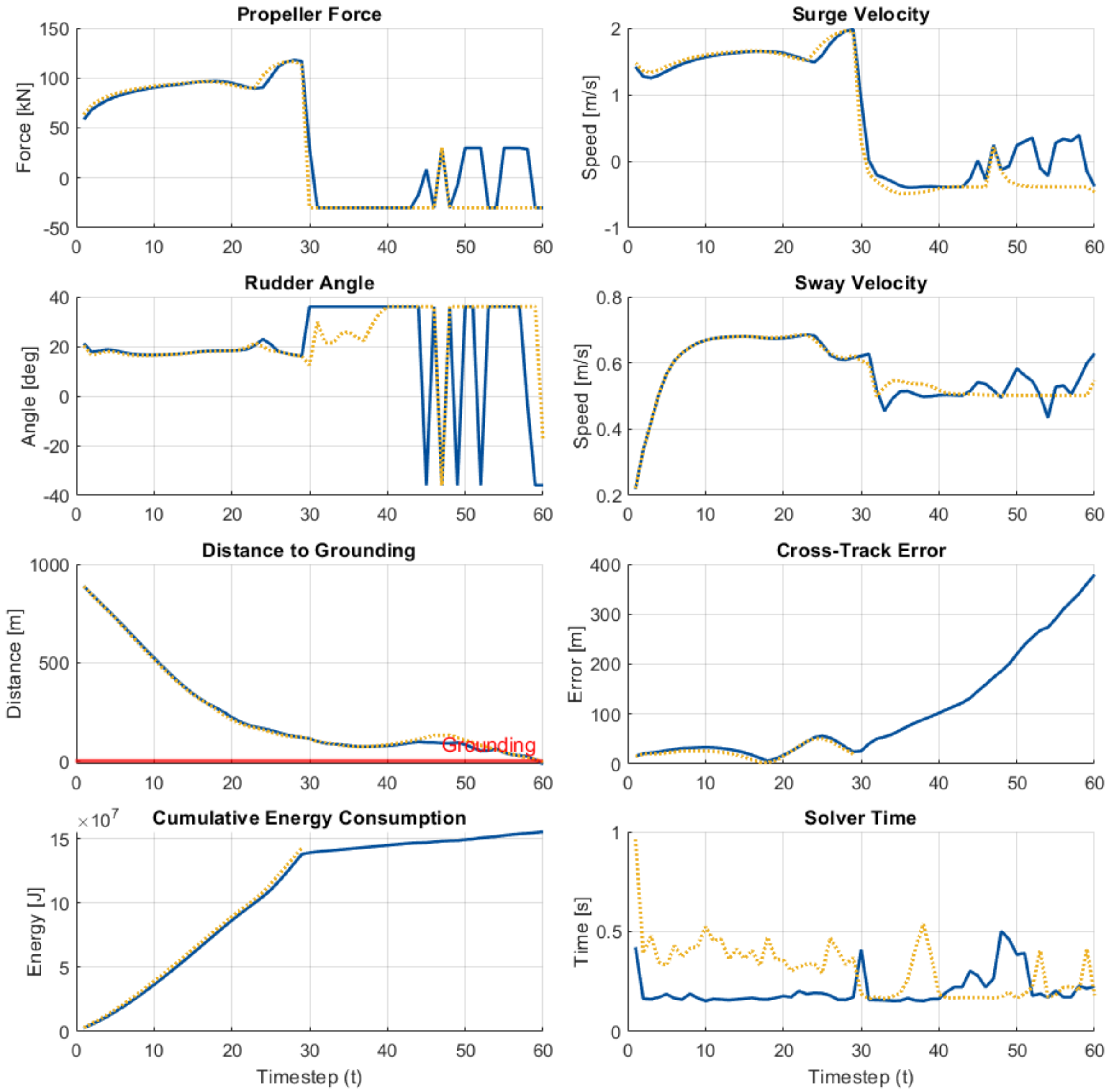


Figure 4.14: Plots for the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault at t=30

Fault at t=35

Figure 4.15 illustrates a loss of effectiveness fault combined with 20 m/s wind in the positive y-direction, with the fault occurring at $t = 35$. As seen on the map, by the time this fault occurs, the vessel is already exiting the narrow passage. The risk along the nominal trajectory is predicted to decrease as the vessel is exiting the narrow passage. Therefore, the fail-safe action is not activated with the MT scheme. Consequently, both the ST and MT scheme are able to continue the mission with the active fault. The vessel starts to drift off course with both schemes, but after the fault is resolved both schemes are able to return to the reference path and complete the mission.

The plots in figure 4.16 detail the response to this situation. At $t = 35$, the propeller force for both schemes drop to the degraded maximum propeller force of approximately 30 kN, causing the surge velocity to come down from 2 m/s to 0.5 m/s. Meanwhile, the 20 m/s crosswind causes a high lateral sway velocity of roughly 0.6 to 0.7 m/s. With the reduced maximum propeller force while being pushed hard in the positive y direction, the trajectory deviates significantly from the reference path, leading to a steady climb in the CTE that peaks near 100 meters.

Table 4.9 confirms the results for both schemes are similar. Both schemes complete the mission in exactly 70 time-steps. They have almost identical safety margins, with the ST scheme recording a minimum distance to grounding (D_{min}) of 121.36 meters, and the MT scheme maintaining 122.22 meters. Energy consumption is also close to each other. However, there is a significant difference in computational costs. The MT scheme has an average solver time of 0.5164 seconds per timestep, compared to the 0.1830 seconds for the ST scheme.

This scenario demonstrates that during manageable fault scenarios, where the vessel is exiting the narrow passage as the fault occurs, the MT scheme correctly does not activate the fail-safe action. In this scenario, the performance of the MT scheme remains virtually identical to the ST scheme with the only significant difference being the higher computational cost.

Table 4.9: Results summary of the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault between t=35 and t=70

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	70	1.9×10^8	121.36	56.69	0.1830	x
MT	70	2.0×10^8	122.22	52.98	0.5164	x

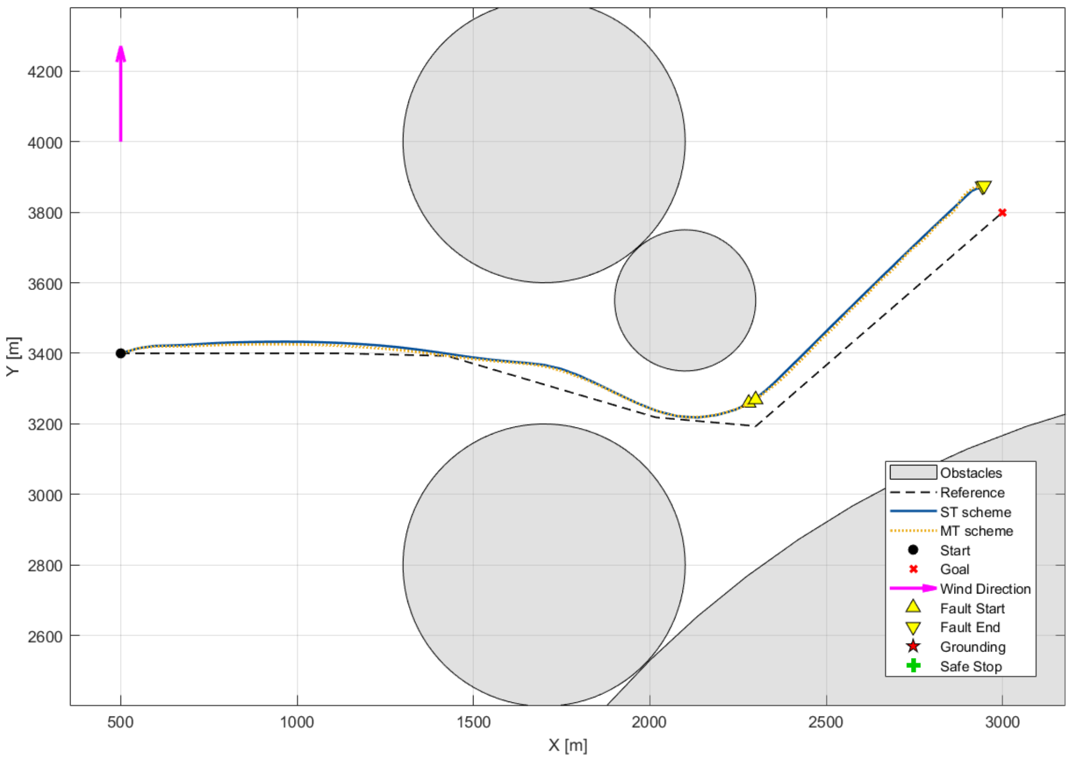


Figure 4.15: Resulting trajectories for the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault between t=35 and t=70

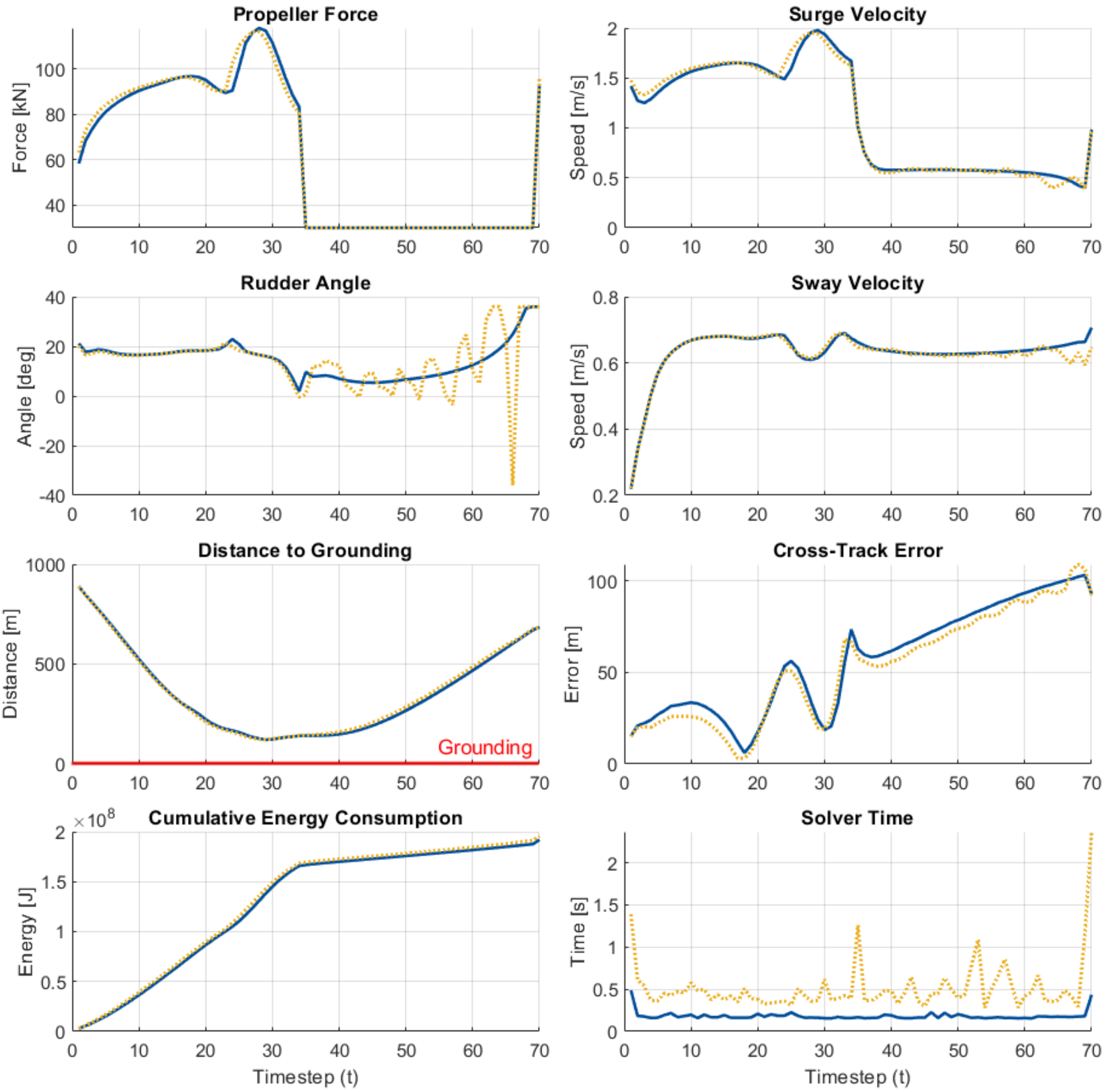


Figure 4.16: Plots for the loss of effectiveness fault scenario with 20 m/s wind velocity and the fault between t=35 and t=70

4.4.4 Loss of Effectiveness Fault Scenario, Constant vs Changing Wind Velocity

This section evaluates the performance of the ST and MT schemes during a persistent loss of effectiveness fault (from $t = 20$). The loss of effectiveness fault severely restricts the capabilities of the vessel by reducing the maximum available propeller force from the normal 400 kN down to just 30 kN ($f_{LoE} = 30\ kN$). To assess the performance of both schemes, the schemes are tested in a scenario where the wind velocity remains constant for the entire simulation duration (16 m/s wind velocity) and where the wind velocity drops mid-operation (from 16 to 5 m/s).

16 m/s wind velocity

Figure 4.17 illustrates the response to loss of effectiveness fault under a 16 m/s crosswind, initiated at $t = 20$. Unlike the previous scenario, this fault is not resolved. As shown in the trajectory map, the ST scheme fails to counteract the wind-induced drift and results in grounding. In contrast, the MT scheme recognizes the hazardous situation, and through the activation of the fail-safe action, guides the vessel away from the narrow passage into open waters. While the mission is not completed, the MT scheme successfully avoids grounding.

The performance plots in figure 4.18 provide further insight. At $t = 20$, the maximum propeller force for both schemes drops to 30 kN. This reduction causes the surge velocity to decrease from 1.7 m/s to approximately 0.5 m/s. The ST scheme attempts to continue along the reference path despite the drifting forces overpowering the remaining control, resulting in a grounding at $t = 69$. Conversely, the MT scheme prioritizes safety over mission progression. It utilizes the remaining control to move the vessel out of the narrow passage.

Table 4.10 summarizes the quantitative differences between the two approaches. The ST scheme records a negative minimum distance to grounding (D_{min}) of -5.32 meters, signifying a grounding event. The MT scheme, however, maintains a safety margin of at least 160.40 meters throughout the simulation. Because the fault is never resolved, the drifting forces continue to overpower the remaining control. Therefore, the vessel is never able to bring itself to a stop. Instead, it continues to drift until the maximum number of simulation steps of 150 steps is reached.

Consistent with previous simulations, the enhanced safety of the MT scheme comes at a higher computational cost, with an average solver time of 0.3829 seconds compared to 0.2296 seconds for the ST scheme.

Table 4.10: Results summary of the loss of effectiveness fault scenario with 16 m/s wind velocity and the fault at t=20

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	-	-	-5.32	-	0.2296	-
MT	-	-	160.40	-	0.3829	-

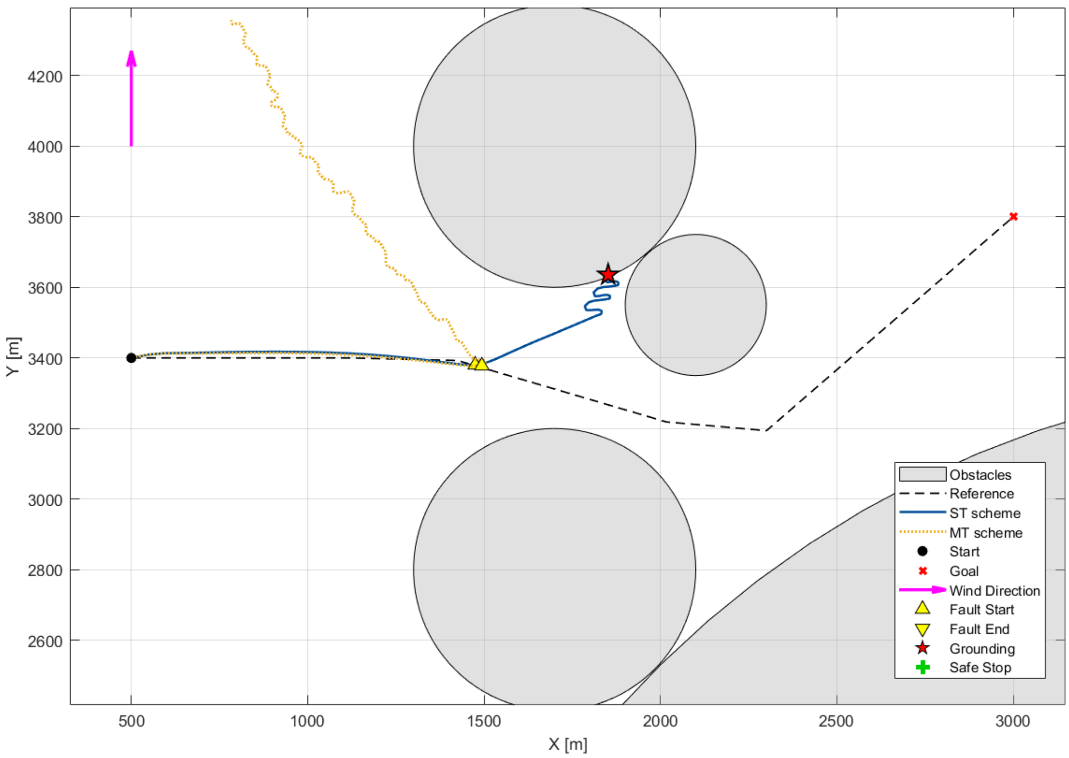


Figure 4.17: Resulting trajectories for the loss of effectiveness fault scenario with 16 m/s wind velocity and the fault at t=20

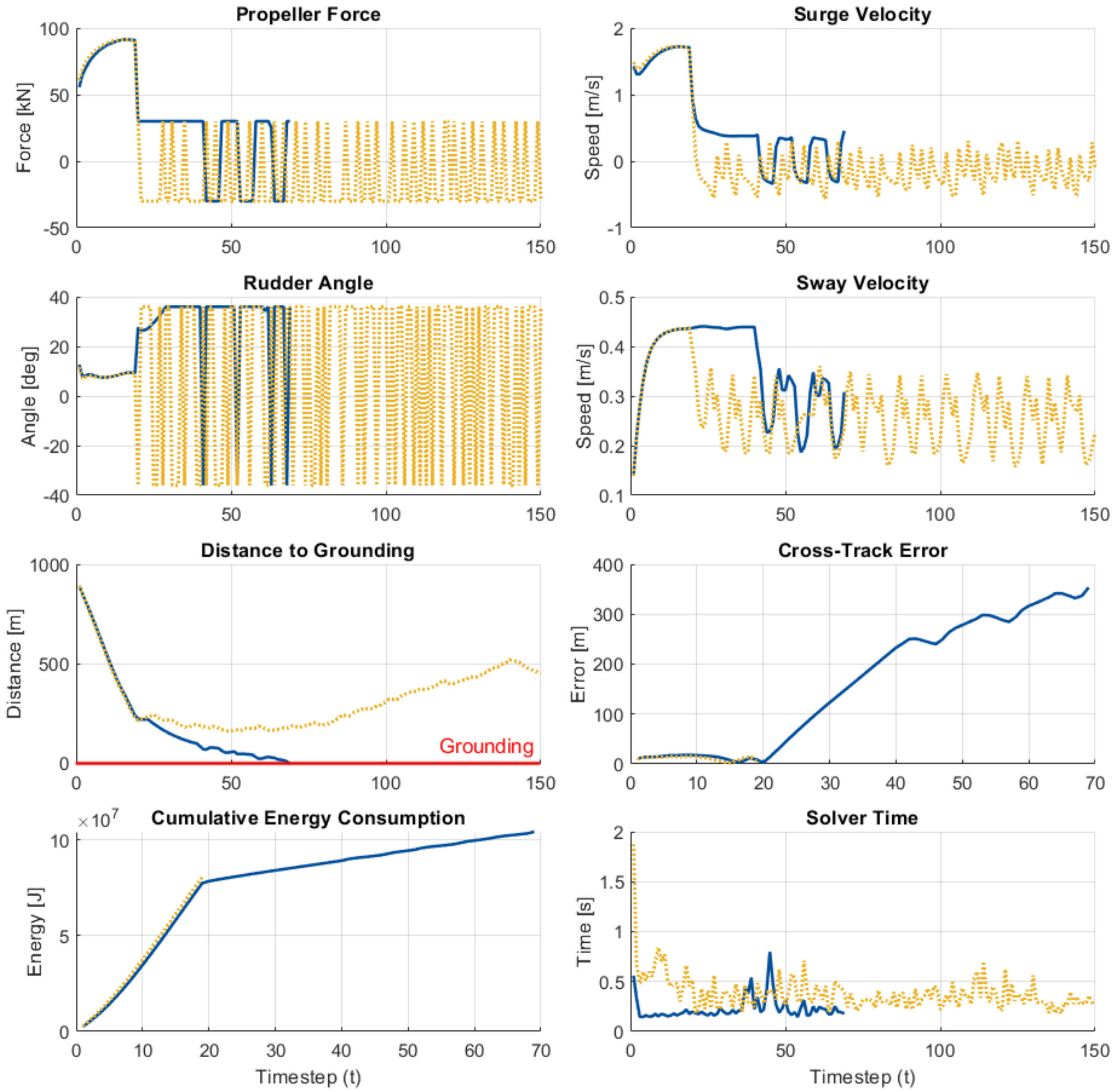


Figure 4.18: Plots for the loss of effectiveness fault scenario with 16 m/s wind velocity and the fault at t=20

Changing wind velocity, 16-5 m/s at t=80

Figure 4.19 illustrates a loss of effectiveness fault initiated at $t = 20$, where the wind velocity drops from 16 m/s to 5 m/s at $t = 80$. In this scenario, the fault is sustained for the remainder of the simulation. The ST scheme is unable to maintain position within the narrow passage under the initial 16 m/s wind, resulting in a grounding event at $t = 69$. The MT scheme, however, activates the fail-safe action at the onset of the fault, and guides the vessel out of the narrow passage into open waters.

The plots in figure 4.20 provide further insight. At $t = 80$, the reduction in wind velocity significantly decreases the drifting forces acting on the vessel. With the decreased wind velocity, the degraded propeller force (30 kN) becomes sufficient to counter the drifting forces. Consequently, the MT scheme is able to reduce the velocity of the vessel and achieve a controlled safe stop at $t = 99$, as indicated by the surge and sway velocities reaching near-zero values.

Table 4.11 provides a quantitative summary of the results. The ST scheme fails the mission with a grounding distance of -5.32 meters. The MT scheme successfully maintains a minimum distance to grounding (D_{min}) of 160.40 meters. The computational cost remains consistent with previous simulations, with the MT scheme requiring a higher average solver time of 0.3380 seconds, compared to 0.2408 seconds for the ST scheme. This scenario demonstrates the ability of the MT scheme to adapt to changing environmental conditions by transitioning from drifting in open waters to a safe stop when the remaining control becomes sufficient to counter the drifting forces.

Table 4.11: Results summary of the loss of effectiveness fault scenario with varying wind velocity and the fault at t=20

Scheme	T_{travel} (Steps)	E_{total} (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)	**Mission Success**
ST	-	-	-5.32	-	0.2408	-
MT	-	-	160.40	-	0.3380	-

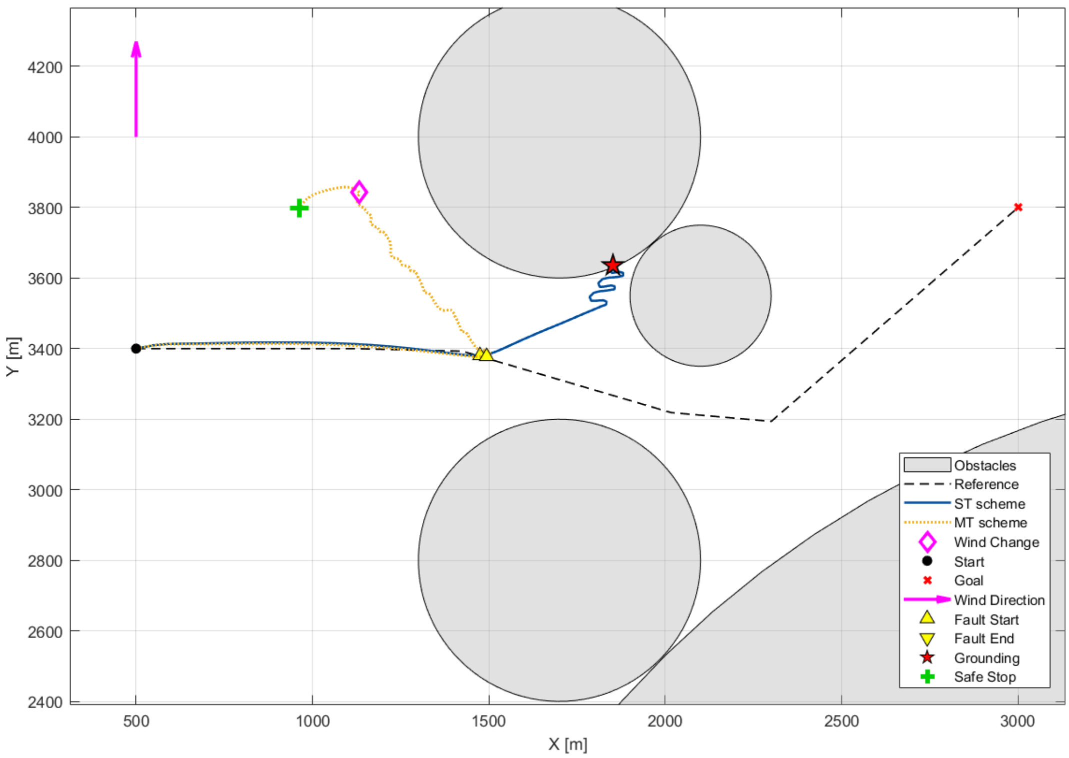


Figure 4.19: Resulting trajectories for the loss of effectiveness fault scenario with varying wind velocity and the fault at t=20

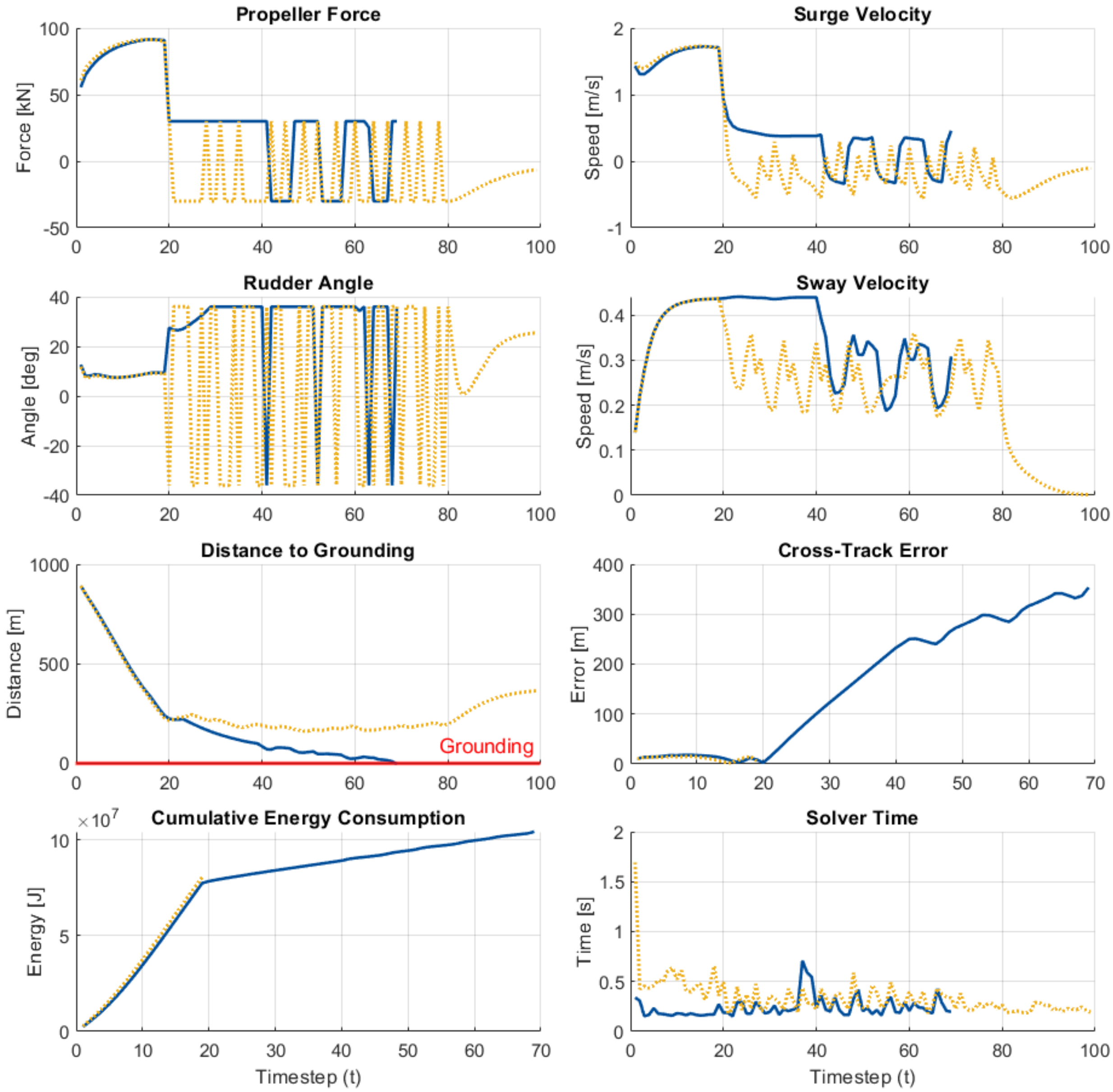


Figure 4.20: Plots for the loss of effectiveness fault scenario with with varying wind velocity and the fault at t=20

4.5 Conclusion

This chapter addressed the research sub-question: *How can the performance of the proposed Multi-Trajectory Risk-Aware MPC be verified through simulations and key performance indicators?*

Verification was achieved by comparing the proposed MT scheme against a ST scheme, which has no fail-safe capabilities. The simulation environment, built upon the CasADi framework and the IPOPT solver, allowed for testing both schemes under varying environmental conditions and steering and propulsion faults. A set of three simulation scenarios are selected: a no-fault scenario to compare the performance under normal operating conditions, alongside a stuck-rudder and loss of effectiveness fault scenario to verify the enhanced safety of the proposed MT scheme. Additionally, the schemes are tested under different moments of fault initiation, fault duration and wind velocities.

To quantitatively verify and compare the performance of the MT scheme, four specific categories of KPIs were defined. Safety was measured using the Distance to Grounding, which provides the safety margin between the vessel and a grounding obstacle. Path tracking performance was evaluated using the Cross-Track Error, while operational efficiency was quantified through Accumulative Energy Consumption and Travel Time.

Finally, the computational cost was evaluated by tracking the Solver Time. These KPIs were used to quantitatively verify the enhanced safety of the MT scheme, as well as the necessary compromises in computational cost, path tracking performance and operational efficiency.

Across all tested scenarios where both schemes were able to successfully complete the mission, specifically all no-fault scenarios, the stuck rudder fault under calm conditions (0 m/s wind, $t = 1$ to 50), and the loss of effectiveness fault during severe wind (20 m/s, $t = 35$ to 70), the data confirms that the MT scheme matches the ST scheme in path tracking accuracy (CTE_{avg}), total travel time (T_{travel}), and total energy consumption (E_{total}). These results show that the simultaneous optimization of the nominal and fail-safe trajectory does not necessarily impose a penalty on path tracking accuracy and operational efficiency compared to the ST scheme. The only significant trade-off for this enhanced safety is the computational cost. Across all scenarios, optimizing two trajectories simultaneously requires solving nearly double the number of decision variables, resulting in an average solver time which is two to three times higher than the average solver time of the ST scheme. Besides, the success of the MT scheme in these specific fault scenarios highlights the ability to limit unnecessarily activating the fail-safe action. For instance, during the stuck rudder fault in calm conditions (0 m/s wind, $t = 1$ to 50), the vessel retains the ability to bring itself to a safe stop. Consequently, the fail-safe action is not triggered, and the fault is handled entirely within the nominal mode (e.g., by safely waiting in front of an obstacle until the rudder fault is resolved before continuing). Furthermore, when the loss of effectiveness fault occurs under severe wind conditions (20 m/s, $t = 35$ to 70) but the vessel is already exiting the narrow passage, the risk decreases over the nominal trajectory as the vessel moves into open waters. As the risk is decreasing, the vessel is allowed to proceed in the nominal mode.

The safety enhancement of the MT scheme becomes evident in scenarios where the ST scheme fails to take decisive risk-mitigating action and results in grounding. During the stuck rudder fault under moderate wind (10 m/s, $t = 1$ to 50) and the loss of effectiveness fault under strong wind (20 m/s, $t = 25$ to 70), the use of the ST scheme results in grounding. The MT scheme correctly identifies the hazardous situation, activates the fail-safe action, reverses the vessel out of the narrow passage and lets the vessel drift in waters where there is no immediate risk of grounding during the fault window. Once the fault is resolved, the vessel is brought to a stop. Furthermore, in the scenario with a persistent loss of effectiveness fault under moderate wind (16 m/s, $t = 20$), the ST scheme grounds at $t = 69$. In contrast, the MT scheme activates the fail-safe action, utilizing the remaining control to steer away from the passage and lets the vessel drift where there is no immediate risk of grounding for the remainder of the simulation, thereby avoiding grounding. Additionally, when the wind velocity drops from 16 m/s to 5 m/s at $t = 80$ while the fault remains active, the degraded propeller force becomes sufficient to counter the drifting forces. This allows the MT scheme to autonomously transition the vessel from drifting to bringing the vessel to a stop.

Despite these successes, the results also exposed limitations within the MT scheme. In the 20 m/s wind velocity with a loss of effectiveness fault between $t = 30$ and 70, the fault occurs as the vessel is deep inside the narrow passage. At this moment, the vessel is surrounded by obstacles and subjected to strong drifting forces. The MT scheme reverses the vessel and attempts to guide the vessel out of the narrow passage, but grounding is inevitable. In scenarios where the fault occurs, such that the remaining control is insufficient to move out of the narrow passage, it will inevitably result in the grounding of the vessel.

In short, the results confirm that in scenarios where the fail-safe action is not activated, the MT scheme consistently achieves path-tracking accuracy and operational efficiency identical to the ST scheme, with the only significant difference being a higher computational cost. The MT scheme demonstrates the ability to continue the operation with active faults when grounding is not predicted and a safe stop remains feasible or the risk is predicted not to increase along the nominal trajectory. Most importantly, the MT scheme reduces the likelihood of grounding compared to the ST scheme. However, in scenarios where the remaining control is insufficient to safely exit the narrow passage, grounding remains inevitable. In the fail-safe mode, when the drifting forces overpower the remaining control of the vessel, the MT scheme guides the vessel out of the narrow passage and lets the vessel drift where there is no immediate risk of grounding. When the control capabilities become sufficient to counter these drifting forces, for instance, due to a reduced wind velocity or after the fault is resolved, the vessel is brought to a stop.

Chapter 5

Conclusion and Future Work

The main research question of this work is: *How can a risk-aware control scheme be designed for fail-safe actions of ASVs during steering and propulsion faults?*. To answer this question, this work proposes a Multi-Trajectory Risk-Aware Model Predictive Control scheme, which enhances safety of ASVs during steering and propulsion faults. This design addresses the main research question by moving beyond traditional schemes that do not consider risk-aware control capabilities that can enable fail-safe actions. The proposed MT scheme simultaneously optimizes two trajectories, which include a mission-driven nominal trajectory and a fail-safe contingency trajectory. These trajectories are optimized under the constraint that the initial control input is identical ($u_0^I = u_0^{II}$). This approach reduces the possibility that the nominal trajectory leads to a state from which an escape is impossible, since the previous fail-safe action is executed if no new, safe solution can be found. Simultaneously, the optimization of the fail-safe trajectory provides the ability to check if a valid fail-safe action exists.

The optimization of both the nominal and fail-safe trajectory is enabled by the objective function $\varphi(\boldsymbol{w}, \boldsymbol{\theta})$, which includes cost terms for both the nominal and fail-safe trajectory. The cost terms for the nominal trajectory are multiplied with binary variable $B \in \{0, 1\}$, providing the ability to switch between the nominal mode ($B = 1$) and fail-safe mode ($B = 0$). If $B = 1$, the cost terms for the nominal trajectory are included in the optimization. If $B = 0$, the cost terms for the nominal trajectory are not included in the optimization. The fail-safe trajectory is computed with a risk cost term in the MPC objective function to guide the vessel towards minimum-risk conditions. In addition, a terminal velocity cost term ensures that the velocity of the vessel can be reduced at each optimization step, to bring it to a near-stop at the end of the fail-safe trajectory. The objective of the fail-safe trajectory is to guide the vessel towards a safe state. Therefore, the risk cost term is divided into the running and terminal risk cost. As the MPC solves an optimization problem over a finite horizon, the MPC only considers what happens within that window. Without a terminal cost, the optimizer might sacrifice the state at the end of the horizon to minimize costs earlier, potentially leaving the vessel in an unsafe state. Separating these terms allow the running risk cost to be tuned independently from the terminal risk cost.

Hazardous situations that require a fail-safe action are identified by the supervisory system that monitors the predicted outcomes of both trajectories. The supervisory system performs two safety checks. First, it executes a grounding check using the safety function

$$\mathcal{S}(\boldsymbol{P}) = \min_{k \in \{0,\ldots,N\}} \left(\min_{j \in \{1,\ldots,J\}} \left(\|\boldsymbol{p}_k - \boldsymbol{c}_j\| - r_j \right) \right).$$

If $\mathcal{S}(\boldsymbol{P}) \leq 0$ for either trajectory, grounding is predicted, and the fail-safe action is immediately activated. Second, it evaluates the stopping capability by monitoring the terminal velocity of the fail-safe trajectory ($v_{terminal} = \sqrt{(u_N^{II})^2 + (v_N^{II})^2}$) and the risk gradient along the nominal trajectory ($\Delta\rho = \rho^I(\boldsymbol{x}_N^I, \boldsymbol{\theta}_N) - \rho^I(\boldsymbol{x}_0^I, \boldsymbol{\theta}_0)$). If $v_{terminal} > v_{th}$ (indicating the vessel cannot bring itself to a stop) and simultaneously $\Delta\rho > \Delta\rho_{th}$ (indicating the nominal path is driving the vessel towards a higher-risk area), the supervisory system activates the fail-safe action. The fail-safe action is activated by switching the binary variable ($B \to 0$), which accordingly modifies the MT MPC objective function. The supervisory system also includes a backup strategy that utilizes the shifted solution from the previous optimization time-step to handle temporary infeasibility.

To verify the performance and fail-safe capabilities of the proposed MT scheme, it was evaluated in a simulated environment against an ST risk-aware control scheme, which is identical to the nominal trajectory of the MT scheme and has no fail-safe capabilities. The verification process tested both schemes across no fault scenarios with varying wind velocities, stuck rudder fault scenarios with varying wind velocities, loss of effectiveness fault scenarios with different moments of fault initiation and loss of effectiveness fault scenarios where the wind remains constant and the wind velocity changes mid-operation. To quantify the performance, a set of KPIs were used, measuring safety (Minimum Distance to Grounding), path tracking accuracy (Average Cross-Track Error), operational efficiency (Cumulative Energy Consumption and Travel Time), and computational cost (Average Solver Time).

The simulation results show that the MT scheme enhances vessel safety by successfully reducing the likelihood of grounding compared to the ST scheme. For example, during a stuck-rudder fault with a 10 m/s crosswind, the ST scheme attempted to continue the mission despite drifting towards grounding obstacles, inevitably resulting in the grounding of the vessel. In contrast, the MT scheme accurately identified the hazardous situation, activated the fail-safe action, and successfully prevented grounding. Similarly, during a loss of effectiveness fault (maximum propeller force reduced from 400 kN to 30 kN) initiated at $t = 25$ under a 20 m/s wind, the use of the ST scheme resulted in grounding. Simultaneously, the MT scheme safely reversed the vessel out of the narrow passage, maintaining a safety margin of at least 57.16 m. Second, this enhanced safety did not compromise operational efficiency and path tracking accuracy during scenarios that do not require a fail-safe action. For instance, under a 10 m/s wind, both schemes recorded identical total energy consumption of 2.2×10^8 J and total travel time of 53 steps, and had a similar average Cross-Track Error of roughly 17.5 m.

However, the enhanced safety of the MT scheme results in an increased computational cost. Because the MT scheme solves for nearly double the decision variables to continuously optimize two trajectories, the average solver time is typically two to three times higher than that of the ST scheme (e.g., 0.4030 s versus 0.2363 s in the 10 m/s no-fault scenario, and 0.4437 s versus 0.1589 s in the 20 m/s no-fault scenario). Furthermore, the simulations demonstrated that if a fault occurs deep inside a narrow passage (e.g., at $t = 30$ under a 20 m/s wind) where the remaining control is insufficient to counter the drift and exit the channel, grounding remains inevitable for both the MT and ST scheme.

Future Work

The proposed Multi-Trajectory Risk-Aware MPC scheme establishes a foundation for risk-aware control that enable fail-safe capabilities for ASVs. The proposed scheme successfully limits the likelihood of grounding compared to the ST scheme, which has no fail-safe capabilities. However, the design and verification in this work rely on several assumptions. To make real-world deployment possible, future work should build upon this work by addressing the following areas:

1. **Integrating a Fault Diagnosis Module**
   A primary assumption in this work is that upon the occurrence of a fault, the exact location, type, and magnitude of the fault (e.g., the specific degradation factor γ_{LoE} or the stuck-rudder angle a_{stuck}) were assumed to be instantaneously and perfectly known. Future research should integrate the MT scheme with a Fault Diagnosis module. Investigating how to handle the phase where a fault has occurred but is not yet fully isolated or quantified is critical.

2. **Dynamic Obstacles and Maritime Traffic Rules**
   The operational environment and verification scenarios in this work exclusively considered static grounding obstacles within a confined channel. However, real-world maritime environments also include other moving vessels. A necessary continuation of this work is expanding the MT scheme to navigate mixed-traffic environments. This requires incorporating maritime traffic rules (COLREGs).

3. **Soft Grounding Strategies for Unrecoverable States**
   As demonstrated in the simulations, if a fault occurs while surrounded by obstacles and subjected to strong drifting forces, the remaining control may be insufficient to exit the narrow passage. Even when activating the fail-safe action immediately after the fault occurs. Under these extreme conditions, grounding becomes inevitable. To manage these unrecoverable states, future research should investigate strategies to minimize the impact of a grounding event. When the supervisory system detects that grounding is inevitable, the objective should shift from avoiding grounding to reducing damage to the vessel and the environment.

4. **Complex Drifting Forces and Sensor Data Handling**
   This work only considers drifting forces due to wind. Future work should include complex waves, irregular currents, and varying weather conditions. Additionally, this work assumed the vessel always has perfect information about its own state. Because real-world sensors are noisy and imperfect, future work should include sensor data handling.

Bibliography

[1] S. Blindheim, S. Gros, and T. A. Johansen, "Risk-based model predictive control for autonomous ship emergency management," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 14 524–14 531, 2020.

[2] SAFETY4SEA, "Tanker runs aground in the philippines; pcg monitors for oil spills," 2026. [Online]. Available: https://safety4sea.com/tanker-runs-aground-in-the-philippines-pcg-monitors-for-oil-spills/

[3] Davis Instruments, "What is navigation at sea?" 2026. [Online]. Available: https://www.davisinstruments.com/pages/what-is-navigation-at-sea

[4] A. Dhyani, Y. Wang, M. Verbeke, D. Pissoort, and V. Reppa, "A pomdp model-based online risk mitigation method for autonomous inland vessels," *IFAC-PapersOnLine*, vol. 58, no. 20, pp. 335–340, 2024.

[5] N. Lodder, C. van der Ploeg, L. Ferranti, and E. Silvas, "Optimization-based fault mitigation for safe automated driving," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 1094–1100, 2023.

[6] M. Akdağ, T. A. Pedersen, T. I. Fossen, and T. A. Johansen, "A decision support system for autonomous ship trajectory planning," *Ocean Engineering*, vol. 292, p. 116562, 2024.

[7] S. Blindheim, T. A. Johansen, and I. B. Utne, "Risk-based supervisory control for autonomous ship navigation," *Journal of Marine Science and Technology*, vol. 28, no. 3, pp. 624–648, 2023.

[8] A. Tsolakis, "Rule-compliant and fault-tolerant motion planning: With application to autonomous surface vehicles," Dissertation (TU Delft), Delft University of Technology, 2025.

[9] D. Saccani, L. Cecchin, and L. Fagiano, "Multitrajectory model predictive control for safe uav navigation in an unknown environment," *IEEE Transactions on Control Systems Technology*, vol. 31, no. 5, pp. 1982–1997, 2022.

[10] S. V. Kumar and J. Kim, "Actuator fault-tolerant mixed passivity and $H\infty$ control for autonomous marine surface vehicles: A sampled-data approach," *European Journal of Control*, vol. 81, p. 101167, 2025.

[11] X. Wang and G.-H. Yang, "Distributed fault-tolerant control for a class of cooperative uncertain systems with actuator failures and switching topologies," *Information Sciences*, vol. 370, pp. 650–666, 2016.

[12] Y. Song, Y. Chen, J. Gao, Y. Wang, and G. Pan, "Collision avoidance strategy for unmanned surface vessel considering actuator faults using kinodynamic rapidly exploring random tree-smart and radial basis function neural network-based model predictive control," *Journal of Marine Science and Engineering*, vol. 11, no. 6, p. 1107, 2023.

[13] M.-Y. Li, L.-T. Liu, W.-B. Xie, and J.-T. Li, "Collision avoidance fault-tolerant control for dynamic positioning vessels under thruster faults," *Ocean Engineering*, vol. 286, p. 115458, 2023.

[14] A. Tsolakis, L. Ferranti, and V. Reppa, "Active thruster fault diagnosis for an overactuated autonomous surface vessel," *IFAC-PapersOnLine*, vol. 58, no. 4, pp. 43–48, 2024.

[15] T. I. Fossen, *Handbook of marine craft hydrodynamics and motion control*. John Wiley & Sons, 2011.

[16] I. Prodan, E. I. Grøtli, L. Lefèvre, and N.-Q.-H. Tran, "Safe navigation in a coastal environment of multiple surface vehicles under uncertainties: A combined use of potential field constructions and nmpc," *Ocean Engineering*, vol. 216, p. 107706, 2020.

[17] P. Causin, J.-F. Gerbeau, and F. Nobile, "Added-mass effect in the design of partitioned algorithms for fluid–structure problems," *Computer methods in applied mechanics and engineering*, vol. 194, no. 42-44, pp. 4506–4527, 2005.

[18] T. A. Driscoll and R. J. Braun, “Fundamentals of numerical computation: Julia edition,” 2023, section 5.2. [Online]. Available: https://tobydriscoll.net/fnc-julia/localapprox/pwlin.html

[19] OceanWise, “ENC Web Map Service (WMS),” 2026. [Online]. Available: https://www.oceanwise-global.com/data/enc-wms/

[20] United States Naval Academy, “EN400: Principles of Ship Performance – Chapter 9: Ship Maneuverability,” 2026. [Online]. Available: https://www.usna.edu/NAOE/_files/documents/Courses/EN400/02.09%20Chapter%209.pdf

Appendix A

Scientific Paper

Risk-aware control for the fail-safe navigation of Autonomous Surface Vessels under thruster faults

Brandon Daver, Abhishek Dhyani, Vasso Reppa

Abstract — Driven by the need to improve safety, the maritime industry is undergoing a transformative change with the development of autonomous surface vessels, which can aim to reduce the likelihood of accidents by incorporating risk awareness directly into motion planning and control. Existing literature does not consider risk-aware control capabilities that can enable fail-safe actions. To address this limitation, this work proposes a Multi-Trajectory (MT) Risk-Aware Model Predictive Control (MPC) scheme that simultaneously optimizes a mission-driven nominal trajectory and a fail-safe contingency trajectory. A supervisory system monitors for hazardous situations that require a fail-safe action and manages the activation of the fail-safe action. Verification is performed in a simulation environment by comparing the MT scheme against a Single-Trajectory (ST) MPC scheme during no-fault and stuck rudder and loss of effectiveness fault scenarios in conjunction with drifting forces due to wind. Results demonstrate that the MT scheme is able to able to reduce the likelihood of grounding compared to the ST scheme. Finally, this enhanced safety imposes no penalty on travel time, energy consumption, or path-tracking accuracy during scenarios that do not require a fail-safe action. The only significant trade-off is a two- to three-fold increase in average solver time compared to the ST scheme.

1 Introduction

The development of Autonomous Surface Vessels (ASVs) is driven by the urgent need to improve maritime safety [1]. The recent grounding of the MT Cheng Xian Feng 168 in January 2026 due to a loss of propulsion and strong drifting forces, highlights the limitations in standard navigation: the failure to account for uncertainties and reliance on human assessment [2, 3]. Risk-aware control of ASVs can account for uncertainties in maritime operations and instantly react based on real-time data [1]. When detecting hazardous situations, such as grounding, these risk-aware control systems can execute a fallback strategy. The appropriate response is dictated by the supervisory system. If the supervisory system determines that the vessel can safely continue the operation, possibly with reduced functionality, a fail-operational action is executed. In contrast, if the supervisory system determines that it cannot guarantee safe operation due to faults or environmental conditions, it must execute a fail-safe action. In this scenario, the focus shifts completely from following the mission to guiding the vessel towards a safe state and cease the operation [4].

The state-of-the-art in risk-aware control for ASVs model the risk of grounding using a safety domain violation term in the Particle Swarm Optimization objective function [5] and exponential risk terms in the MPC objective functions and account for vulnerability to steering and propulsion faults [1, 6]. While these methods consider mission progression and fail-operational actions like rerouting, they lack fail-safe capabilities. [7] proposes a MT MPC, which utilizes a dual-plan strategy to ensure smooth transition to a fault-tolerant control strategy. However, no fail-safe actions are included. Therefore, the identified gap in the state-of-the-art is that existing literature does not consider risk-aware control capabilities that can enable fail-safe actions. The use of risk-aware control schemes that optimize a single trajectory are are not well suited, as it is uncertain if a feasible fail-safe action exists until there is an attempt to find one [1, 5, 6]. This creates a blind spot where a vessel may unknowingly enter a state of no return. A MT MPC scheme provides a solution by simultaneously optimizing a mission-driven nominal trajectory and a fail-safe trajectory under the constraint that the initial control input is identical. This approach reduces the possibility that the nominal trajectory leads to a state from which an escape is impossible, since the previous fail-safe action is executed if no new, safe solution can be found [7]. Simultaneously, it provides the ability to check if a valid fail-safe action even exists. Therefore, to address the identified gap, this work proposes a MT risk-aware MPC scheme, which optimizes the mission-driven nominal trajectory and fail-safe trajectory simultaneously. Integrated into this scheme is a supervisory system designed to identify hazardous situations that require a fail-safe action and manages the activation of the fail-safe action. This work will consider hazardous situations that arise from unsafe actions taken during vessel operations originating from steering and propulsion faults and drifting forces due to wind. Fault detection and isolation is considered outside the scope of this work. The main research question that will be addressed in this work is: *How can a risk-aware control scheme be designed for fail-safe actions of ASVs during steering and propulsion faults?* This question is supported by the following sub-questions:

1. Which steering and propulsion fault scenarios create hazardous situations that require a fail-safe action?

2. How to design a risk-aware MPC in which both the nominal and fail-safe trajectories are considered?

3. How can the performance of the proposed Multi-Trajectory Risk-Aware MPC be verified through simulations and key performance indicators?

The key contribution of this work is the design and verification of the Multi-Trajectory Risk-Aware MPC scheme that enables fail-safe capabilities for ASVs.

2 Problem Formulation

2.1 Ship Model

For the multi-trajectory optimization, a dynamic model of the vessel is required. The model used in this work is a 3 Degree of Freedom (DOF) model [8]. The motion of the vessel in the horizontal plane is described by a state vector $\boldsymbol{x} = [\boldsymbol{\eta}^\top, \boldsymbol{\nu}^\top]^\top$. This vector combines the earth-fixed position and orientation $\boldsymbol{\eta} = [x, y, \psi]^\top$ with the body-fixed velocities $\boldsymbol{\nu} = [u, v, r]^\top$ (surge, sway, and yaw rate). The dynamics of the vessel are governed by the following equations:

$$\dot{\boldsymbol{\eta}} = \boldsymbol{J}(\psi)\,\boldsymbol{\nu} \qquad \text{(Kinematic equation)}$$
$$\dot{\boldsymbol{\nu}} = \boldsymbol{M}^{-1}(\boldsymbol{\tau} + \boldsymbol{d} - \boldsymbol{D}\,\boldsymbol{\nu}) \qquad \text{(Kinetic equation)}$$

This work focuses on an ASV equipped with a single rudder and propeller located at the stern. This propeller-rudder configuration is modeled as a single point force and the control forces are approximated using a trigonometric decomposition. In this model, the rudder is assumed to deflect the total propeller force f by an angle a, which are collected in the control input vector $\boldsymbol{u} = [f, a]^\top$.

Table 1: Summary of the 3-DOF Vessel Model

Component Symbol	**Mathematical Definition**
Rotation Matrix $\boldsymbol{J}(\psi)$	$\begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Inertia Matrix $\boldsymbol{M}$	$\begin{bmatrix} m + X_{\dot{u}} & 0 & 0 \\ 0 & m + Y_{\dot{v}} & 0 \\ 0 & 0 & I_z + N_{\dot{r}} \end{bmatrix}$
Damping Matrix $\boldsymbol{D}$	$\begin{bmatrix} X_u & 0 & 0 \\ 0 & Y_v & 0 \\ 0 & 0 & N_r \end{bmatrix}$
Control Force Vector $\boldsymbol{\tau}$	$\begin{bmatrix} f\cos(a) \\ f\sin(a) \\ -l_x f\sin(a) \end{bmatrix}$
Wind Force Vector $\boldsymbol{d}$ ($\boldsymbol{\tau}_{\text{wind}}$)	$\frac{1}{2}\rho_a V_w^2 \begin{bmatrix} -c_x\cos(\gamma_w)A_{Fw} \\ c_y\sin(\gamma_w)A_{Lw} \\ c_n\sin(2\gamma_w)A_{Lw}L_{oa} \end{bmatrix}$

To provide a complete overview of the system, all matrices and vectors that make up the kinematic and kinetic equations are mathematically summarized in table 1. m and I_z are the mass and moment of inertia about the vertical axis of the vessel, respectively. $X_{\dot{u}}$, $Y_{\dot{v}}$, and $N_{\dot{r}}$ are the added mass coefficients and X_u, Y_v, and N_r are the linear damping coefficients. The yaw moment in the control force vector is given by $-l_x f\sin(a)$, where l_x denotes the lever arm distance between the center of gravity of the vessel and the rudder post. In the wind force vector, ρ_a represents the air density, V_w the relative wind velocity and ψ_w is the true wind direction. A_{Fw} is the frontal wind area of the vessel, while A_{Lw} is the lateral wind area. The overall length of the vessel, L_{oa}, acts as a lever in the calculation of the yaw moment. The forces and moment are scaled by a set of dimensionless wind force coefficients, c_x, c_y, and c_n and are the following [1]:

$$c_x = 0.7, \qquad c_y = 0.8, \qquad c_n = 0.1.$$

The ship dynamics are modeled using the three-dimensional kinematic and kinetic equations in the horizontal plane, which are the time derivatives of the state vector [1]. This results in the following first-order Ordinary Differential Equation (ODE):

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}) = \begin{bmatrix} \dot{\boldsymbol{\eta}} \\ \dot{\boldsymbol{\nu}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{J}(\boldsymbol{\eta})\boldsymbol{\nu} \\ \boldsymbol{M}^{-1}(\boldsymbol{\tau} + \boldsymbol{\tau}_{\text{wind}} - \boldsymbol{D}\boldsymbol{\nu}) \end{bmatrix}$$

To optimize the nominal and fail-safe trajectory, the continuous-time ODE is discretized into a finite-dimensional problem by dividing the prediction horizon into N intervals with a duration of Δt. The control inputs are the same for the entire time interval. The dynamic behavior of the vessel is approximated using the Fourth Order Runge Kutta method [1, 6].

2.2 Steering & Propulsion Fault Scenarios

This work focuses on an ASV operating in a confined channel. A hazardous situation that requires a fail-safe action can occur when there is a loss of effectiveness or stuck-rudder fault, while being subjected to drifting forces due to wind that overpower the remaining control and push the vessel towards grounding obstacles [1, 2]. During a loss of effectiveness fault, the steering and propulsion system can only deliver a fraction of the commanded input. This can be managed by scaling the commanded control input accordingly to compensate for the loss. However, a consequence is the reduction of the maximum propeller force to a new, lower limit [9]. With this reduction, the vessel potentially cannot generate the force required to counter the drifting forces. Upon the occurrence of the stuck-rudder fault, the rudder angle is stuck at the last applied control input, eliminating the steering capability of the vessel [1]. When subjected to drifting forces, the stuck fault might make it impossible for the vessel to generate the forces required to counter the drifting forces. This work assumes an ideal fault diagnosis (FD). Consequently, upon the occurrence of a fault it is assumed that the exact location, type, and magnitude are instantaneously and perfectly known. Furthermore, it is assumed that only a single fault can occur during a given operation and that the occurrence coincides precisely with the discrete time-steps.

2.3 Reference Path & Parameter Vector

The reference path is parameterized using a two-dimensional vector function $\boldsymbol{r}(\alpha)$, which maps an advancement variable α to a specific coordinate in the global reference frame, using piecewise linear functions [1, 6].

$$\boldsymbol{r}(\alpha) = \begin{bmatrix} r_x(\alpha) \\ r_y(\alpha) \end{bmatrix}, \quad \text{for } \alpha \in [0, L_{total}],$$

The solution, provided by the MPC, is governed by a set of parameters that define the mission objectives and operational constraints. These parameters are combined into a single parameter vector, $\boldsymbol{\theta}_k$ and consists of the reference transit speed s_{ref}, the advancement rate α_{step}, and the obstacle vector $\boldsymbol{\sigma}$ [1]. Together, the parameter vector is defined as:

$$\boldsymbol{\theta}_k = \begin{bmatrix} s_{ref} \\ \alpha_{step} \\ \boldsymbol{\sigma} \end{bmatrix} \in R^{2+3J}.$$

Here, the obstacles are modeled as a set of J circular exclusion zones, which are assumed to account for the dimensions of the vessel. The obstacle vector $\boldsymbol{\sigma}$ is a vector defining the characteristics for all J obstacles:

$$\boldsymbol{\sigma} = \begin{bmatrix} \boldsymbol{\sigma}_1 \\ \boldsymbol{\sigma}_2 \\ \cdots \\ \boldsymbol{\sigma}_J \end{bmatrix}, \quad \text{with } \boldsymbol{\sigma}_j = \begin{bmatrix} x_j \\ y_j \\ r_j \end{bmatrix}, \quad \text{for } j \in \{1, ..., J\}.$$

Here, x_j and y_j denote the global coordinates of the center of obstacle j, while r_j represents the radius.

3 Methodology

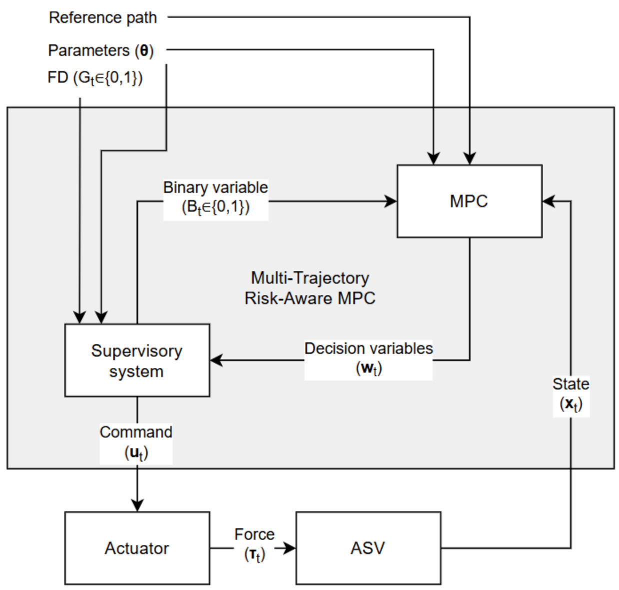


Figure 1: Overview of the control structure

Figure 1 provides the control structure of the Multi-Trajectory Risk-Aware MPC scheme. An important aspect of this control structure is the distinction between the two time horizons. The outer loop, indexed by $t \in \{1, \ldots, T\}$, represents the discrete optimization time-steps. Within each of these time-steps t, the MPC performs an optimization over an internal prediction horizon, indexed by $k \in \{0, \ldots, N\}$. Here, N represents the look-ahead capability of the MPC.

3.1 Decision Variable Vector

The decision variable vector $\boldsymbol{w}$ represents the complete set of variables which the MPC needs to optimize and consist of the state and control input vectors of the nominal (denoted by superscript I) and fail-safe trajectory (denoted by superscript II) and the advancement variable. This results in the following vector:

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{x}_0^{I\top} & \boldsymbol{x}_0^{II\top} & \alpha_0 & \boldsymbol{u}_0^{I\top} & \boldsymbol{u}_0^{II\top} & \ldots & \boldsymbol{x}_{N-1}^{I\top} & \boldsymbol{x}_{N-1}^{II\top} & \alpha_{N-1} & \boldsymbol{u}_{N-1}^{I\top} \\ \boldsymbol{u}_{N-1}^{II\top} & \boldsymbol{x}_N^{I\top} & \boldsymbol{x}_N^{II\top} & \alpha_N \end{bmatrix}^\top$$

3.2 Constraints and Objective Function

At every optimization time-step t, the MPC solves a receding horizon problem to determine the best course of action. This process involves finding the optimal decision variable vector $\boldsymbol{w}$ that minimizes the total cost over the prediction horizon while satisfying all constraints. The mathematical formulation of this problem is defined as follows:

$$\begin{aligned} \min \quad & \varphi(\boldsymbol{w}, \boldsymbol{\theta}) \\ \text{s.t.} \quad & \boldsymbol{x}_{k+1}^I = f(\boldsymbol{x}_k^I, \boldsymbol{u}_k^I) \\ & \boldsymbol{x}_{k+1}^{II} = f(\boldsymbol{x}_k^{II}, \boldsymbol{u}_k^{II}) \\ & \alpha_k \geq \alpha_{k-1} \\ & \boldsymbol{u}_0^I = \boldsymbol{u}_0^{II} \\ & \alpha_{\text{step}} = s_{\text{ref}} \cdot \Delta t \\ & -\boldsymbol{u}_{\text{limit}} \leq \boldsymbol{u}_k^I \leq \boldsymbol{u}_{\text{limit}} \\ & -\boldsymbol{u}_{\text{limit}} \leq \boldsymbol{u}_k^{II} \leq \boldsymbol{u}_{\text{limit}} \end{aligned}$$

where $\boldsymbol{u}_{limit} = [f_{limit}, a_{limit}]^\top$. The objective function $\varphi(\boldsymbol{w}, \boldsymbol{\theta})$ is derived from the objective function proposed in [1] and is defined as:

$$\begin{aligned} \varphi(\boldsymbol{w}, \boldsymbol{\theta}) = & \sum_{k=0}^{N} \Big[\underbrace{B\xi(\boldsymbol{x}_k^I, \boldsymbol{\theta}_k, \alpha_k, \alpha_{k-1})}_{\text{Progression}} + \underbrace{B\rho^I(\boldsymbol{x}_k^I, \boldsymbol{\theta}_k)}_{\text{Nominal Risk}} \Big] \\ & + \sum_{k=0}^{N-1} \Big[\underbrace{B\epsilon^I(\boldsymbol{u}_k^I, \boldsymbol{u}_{k-1}^I)}_{\substack{\text{Nominal} \\ \text{Control Input}}} + \underbrace{\epsilon^{II}(\boldsymbol{u}_k^{II}, \boldsymbol{u}_{k-1}^{II})}_{\substack{\text{Fail-Safe} \\ \text{Control Input}}} + \underbrace{\rho^{II}(\boldsymbol{x}_k^{II}, \boldsymbol{\theta}_k)}_{\substack{\text{Fail-Safe} \\ \text{Risk}}} \Big] \\ & + \underbrace{\rho^T(\boldsymbol{x}_N^{II}, \boldsymbol{\theta}_N)}_{\substack{\text{Terminal} \\ \text{Risk}}} + \underbrace{\zeta(\boldsymbol{x}_N^{II})}_{\substack{\text{Terminal} \\ \text{Velocity}}}. \end{aligned}$$

3.2.1 Progression Cost Term

$$\xi(\boldsymbol{x}_k^I, \boldsymbol{\theta}_k, \alpha_k, \alpha_{k-1}) = \boldsymbol{\kappa} \begin{bmatrix} \|\boldsymbol{r}(\alpha_k) - \boldsymbol{p}_k^I\|^2 \\ \|\alpha_k - \alpha_{k-1} - \alpha_{\text{step}}\|^2 \end{bmatrix},$$

where $\boldsymbol{\kappa}$ is a weight vector. $\|\boldsymbol{r}(\alpha_k)-\boldsymbol{p}_k^I\|^2$ quantifies the squared error between the predicted position of the vessel $\boldsymbol{p}_k^I=[x_k^I,y_k^I]^\top$ and the reference position on the path $\boldsymbol{r}(\alpha_k)$. The second term, $\|\alpha_k-\alpha_{k-1}-\alpha_{\text{step}}\|^2$, encourages the ship to advance along the reference path at the desired rate α_{step}.

3.2.2 Control Input Cost Terms

$$\begin{aligned}\epsilon^I(\boldsymbol{u}_k^I,\boldsymbol{u}_{k-1}^I)&=\boldsymbol{u}_k^{I\top}\boldsymbol{\Lambda}^I\boldsymbol{u}_k^I\\&\quad+(\boldsymbol{u}_k^I-\boldsymbol{u}_{k-1}^I)^\top\boldsymbol{\Delta}^I(\boldsymbol{u}_k^I-\boldsymbol{u}_{k-1}^I)\\\epsilon^{II}(\boldsymbol{u}_k^{II},\boldsymbol{u}_{k-1}^{II})&=(\boldsymbol{u}_k^{II}-\boldsymbol{u}_{k-1}^{II})^\top\boldsymbol{\Delta}^{II}(\boldsymbol{u}_k^{II}-\boldsymbol{u}_{k-1}^{II})\end{aligned}$$

The component of the control input cost term, weighted by the diagonal matrix $\boldsymbol{\Lambda}\in R^{2\times 2}$, penalizes the magnitude of the control inputs to conserve power. The component, weighted by the diagonal matrix $\boldsymbol{\Delta}\in R^{2\times 2}$, penalizes the difference between consecutive control inputs to improve stability [1].

3.2.3 Risk Cost Terms

The generalized risk cost term, with $s\in\{I,II,T\}$, is defined as:

$$\rho^s(\boldsymbol{x}_k^s,\boldsymbol{\theta}_k)=\sum_{j=1}^{J}\left(\mu_1^s+\mu_2^s\,V_w\,\max(0,\hat{\boldsymbol{i}}_{j,k}^s\cdot\hat{\boldsymbol{\omega}}_{\text{wind}})\right)e^{-\Xi^s(\|\boldsymbol{p}_k^s-\boldsymbol{c}_j\|-r_j)}$$

$e^{-\Xi^s(\|\boldsymbol{p}_k^s-\boldsymbol{c}_j\|-r_j)}$ calculates the Euclidean distance between the current position of the vessel $\boldsymbol{p}_k$ and the j-th obstacle. The magnitude of the risk penalty is scaled dynamically by the term, $(\mu_1^s+\mu_2^s\,V_w\,\max(0,\hat{\boldsymbol{i}}_{j,k}^s\cdot\hat{\boldsymbol{\omega}}_{\text{wind}}))$, which accounts for both static risk and additional risk introduced by the wind. $\hat{\boldsymbol{i}}_{j,k}$ represents the direction from the current position of the vessel to the center of the obstacle, while $\hat{\boldsymbol{\omega}}_{\text{wind}}=[\cos(\psi_w),\sin(\psi_w)]^\top$ is the unit vector representing the global wind direction ψ_w. The function $\max(0,\cdot)$, ensures that the additional risk due to wind is only included when the vessel is positioned upwind of the grounding obstacle. Ξ, μ_1 and μ_2 are tuning parameters.

3.2.4 Terminal Velocity Cost Term

$$\zeta(\boldsymbol{x}_N^{II})=\Omega((u_N^{II})^2+(v_N^{II})^2)$$

The term is defined by the squared velocity of the ship at the end of the horizon, $\sqrt{(u_N^{II})^2+(v_N^{II})^2}$, calculated from the body-fixed frame velocity components u_N^{II} and v_N^{II}. The magnitude of this penalty is controlled by the weighting scalar Ω.

3.3 The Supervisory System

The supervisory system is used to identify hazardous situations that require a fail-safe action and manages the activation of the fail-safe action.

3.3.1 Solver Convergence

First, the supervisory system verifies if the MPC is able to find a solution. If the solver fails to find a solution ($E_t=0$), the supervisory system checks if no steering and propulsion fault has occurred in the current time-step ($G_t\geq G_{t-1}$) and if there are no consecutive solver failures ($E_{t-1}=1$). If the conditions are met, the system bypasses the failed optimization by retrieving the solution from the previous time-step ($\boldsymbol{w}_{t-1}$). In the nominal mode ($B=1$), $\boldsymbol{u}_1^I$ is executed and in the fail-safe mode ($B=0$), $\boldsymbol{u}_1^{II}$ is executed. Conversely, if one of these conditions are not met, the supervisory system forces an immediate transition to the fail-safe mode ($B\to 0$) when operating in the nominal mode. When already in the fail-safe mode ($B=0$), the system terminates the optimization.

3.3.2 Safety Check

If the MPC is able to find a solution ($E_t=1$), the MPC returns the decision variable vector, denoted as $\boldsymbol{w}_t$. The supervisory system decomposes $\boldsymbol{w}_t$ into two matrices: the nominal position matrix $\boldsymbol{P}^I=[\boldsymbol{p}_0^I,...,\boldsymbol{p}_N^I]\in R^{2\times(N+1)}$ and the fail-safe position matrix $\boldsymbol{P}^{II}=[\boldsymbol{p}_0^{II},...,\boldsymbol{p}_N^{II}]\in R^{2\times(N+1)}$. First, the supervisory system evaluates the trajectories for grounding using the function:

$$\mathcal{S}(\boldsymbol{P})=\min_{k\in\{0,...,N\}}\left(\min_{j\in\{1,...,J\}}(\|\boldsymbol{p}_k-\boldsymbol{c}_j\|-r_j)\right).$$

The second evaluation assesses the stopping capability by extracting the terminal velocity of the fail-safe trajectory, $v_{terminal}=\sqrt{(u_N^{II})^2+(v_N^{II})^2}$. And finally, the risk gradient along the nominal trajectory is assessed using the function, $\Delta\rho=\rho^I(\boldsymbol{x}_N^I,\boldsymbol{\theta}_N)-\rho^I(\boldsymbol{x}_0^I,\boldsymbol{\theta}_0)$.

In the nominal mode, the supervisory system evaluates both trajectories for grounding, the terminal velocity and the risk gradient. The logic in the nominal mode ($B=1$) is defined as follows:

$$B_{next}=\begin{cases}1 & \text{if } \mathcal{S}(\boldsymbol{P}^I)>0\wedge\mathcal{S}(\boldsymbol{P}^{II})>0\wedge v_{\text{terminal}}\leq v_{\text{th}}\\0 & \text{if } \mathcal{S}(\boldsymbol{P}^I)\leq 0\vee\mathcal{S}(\boldsymbol{P}^{II})\leq 0\vee(v_{\text{terminal}}>v_{\text{th}}\wedge\Delta\rho>\Delta\rho_{\text{th}})\end{cases}$$

If $B_{next}=1$, the system executes the nominal control input $\boldsymbol{u}_t=\boldsymbol{u}_0^I$. If $B_{next}=0$, the system discards the current solution, updates the binary variable to $B=0$, and triggers an immediate re-optimization in the current optimization time-step. In the fail-safe mode, the supervisory system checks if the fail-safe trajectory is grounding-free. The logic in the fail-safe mode is defined as:

$$\boldsymbol{u}_t=\begin{cases}\boldsymbol{u}_0^{II} & \text{if } \mathcal{S}(\boldsymbol{P}^{II})>0\\0 & \text{if } \mathcal{S}(\boldsymbol{P}^{II})\leq 0\end{cases}$$

In the scenario $\boldsymbol{u}_t=0$, the system determines that no feasible fail-safe trajectory exists to avoid grounding from the current state and stops the optimization.

4 Verification

The performance of the MT scheme is verified against a ST scheme (based on [1]), which is identical to the nominal trajectory of the MT scheme, to verify the fail-safe capabilities. The no-fault, stuck-rudder fault, and loss of effectiveness fault scenarios are simulated under varying wind velocities, with the wind going in the positive y direction (see figure 7). Furthermore, the simulations include different moments of fault initiation and duration, as well as a scenario where the wind velocity changes mid-operation. The performances are quantitatively compared based on five Key Performance Indicators (KPIs): safety via Minimum Distance to Grounding (D_{min}), tracking accuracy through Average Cross-Track Error (CTE_{avg}), operational efficiency using Travel Time (T_{travel}) and Accumulative Energy Consumption (E_{total}), and computational cost via Average Solver Time ($t_{comp,avg}$). Finally, implementation is conducted in a MATLAB simulation environment using the CasADi framework and IPOPT solver. To reduce solver time, a warm start strategy is deployed, which shifts the previous optimal solution forward as the initial guess for the current step. Upon the occurrence of the stuck fault, the rudder angle is stuck at the last applied control input a_{stuck}. The stuck-rudder fault is modeled as an additional equality constraint, $a_k = a_{\text{stuck}}$. Upon occurrence of the loss of effectiveness fault, the maximum available propeller force is reduced to a new lower level f_{LoE}. This fault is modeled as a constraint on the propeller force, $-f_{LoE} \leq f_k \leq f_{LoE}$.

4.1 Results

Table 2: Simulation Results: No-Fault, Stuck Rudder, and Loss of Effectiveness Scenarios

Scenario	Scheme	T_{travel} (Steps)	E_{total} $\times 10^8$ (J)	D_{min} (m)	CTE_{avg} (m)	$t_{comp,avg}$ (s)
No-Fault Scenario with Varying Wind Velocity						
10 m/s	ST	53	2.2	122.16	17.89	0.2363
	MT	53	2.2	121.97	17.41	0.4030
15 m/s	ST	53	2.2	122.23	19.84	0.1589
	MT	53	2.2	121.99	19.21	0.4437
20 m/s	ST	54	2.2	121.36	26.01	0.1548
	MT	53	2.2	122.22	23.52	0.4531
Stuck Rudder Fault with Varying Wind Velocity						
0 m/s	ST	68	3.3	53.29	67.15	0.0810
t = 1-50	MT	69	3.1	34.04	67.69	0.1994
10 m/s	ST	-	-	-101.62	-	0.1387
t = 1-50	MT	-	-	42.23	-	0.2876
Loss of Effectiveness Fault with Moments of Fault Initiation						
20 m/s	ST	-	-	-10.52	-	0.2113
t = 25-70	MT	-	-	57.16	-	0.7008
20 m/s	ST	-	-	-13.50	-	0.2065
t = 30-	MT	-	-	-3.43	-	0.3134
20 m/s	ST	70	1.9	121.36	56.69	0.1830
t = 35-70	MT	70	2.0	122.22	52.98	0.5164
Loss of Effectiveness Fault with Constant and Changing Wind Velocity						
16 m/s	ST	-	-	-5.32	-	0.2296
t = 20-	MT	-	-	160.40	-	0.3829
16-5 m/s	ST	-	-	-5.32	-	0.2408
t = 20-	MT	-	-	160.40	-	0.3380

Across all scenarios that do not require a fail-safe action, the data confirms that the MT scheme matches the ST scheme in path tracking accuracy (CTE_{avg}), total travel time (T_{travel}) and total energy consumption (E_{total}). However, there was a significant difference in the computational cost. The average solver time for the MT scheme is two to three times higher than for the ST scheme.

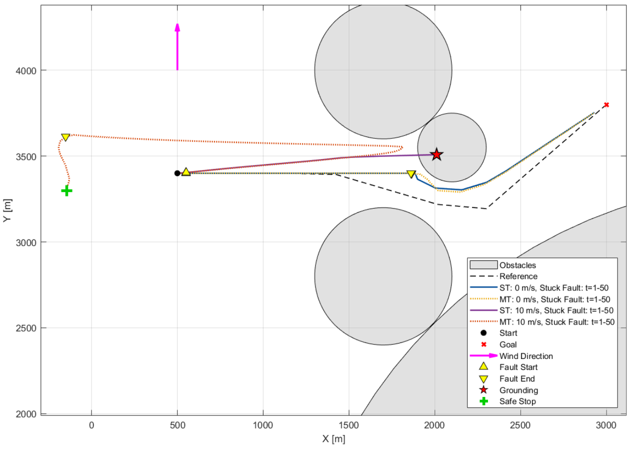


Figure 2: Trajectories during stuck-rudder fault scenario with varying wind velocity

Figure 2 shows the results from the stuck-rudder fault scenario under varying wind velocities. During the stuck rudder fault in calm conditions (0 m/s wind, $t = 1$ to 50), the vessel retains the ability to bring itself to a stop. Consequently, the fail-safe action is not triggered, allowing the vessel to complete the mission. During the stuck rudder fault under moderate wind (10 m/s, $t = 1$ to 50), the use of the ST scheme results in grounding. The MT scheme correctly identifies the hazardous situation and is able to prevent grounding.

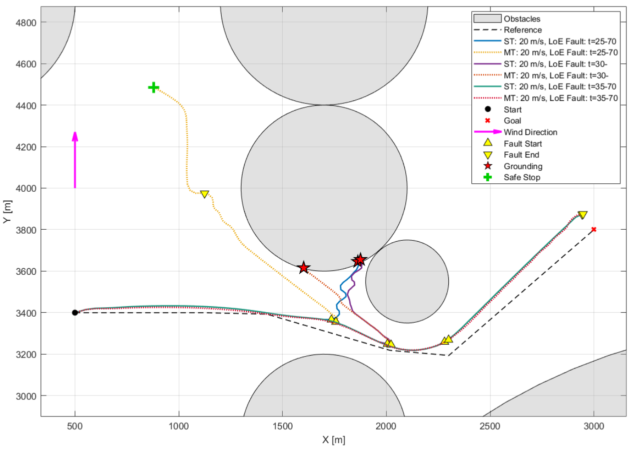


Figure 3: Trajectories during loss of effectiveness fault scenario with varying moments of fault initiation

Figure 3 shows the results from the loss of effectiveness fault scenario with varying moments of fault initiation. During the loss of effectiveness fault under strong wind (20 m/s, $t = 25$ to 70), the use of the ST scheme results in grounding. The MT scheme correctly identifies the hazardous situation and is able to prevent grounding. However, when the fault is initiated at $t = 30$, it occurs when the vessel is approximately at the deepest, narrowest point inside the narrow passage. The fail-safe action is immediately activated after the fault occurs. Although the fail-safe action is immediately activated with the MT scheme after the fault occurs, the vessel is physically unable to exit the narrow passage before grounding occurs. With the fault occurring at $t = 35$, the fault occurs as the vessel is exiting the narrow passage. Because the risk decreases over the nominal trajectory, the vessel is allowed to proceed in the nominal mode.

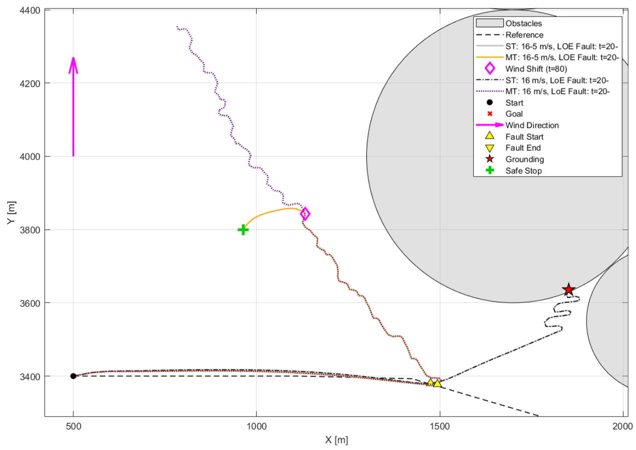


Figure 4: Trajectories during loss of effectiveness fault scenario with constant and changing wind velocity

Furthermore, as shown in figure 4, under a constant 16 m/s wind velocity with a persistent loss of effectiveness fault initiated at $t = 20$, the ST scheme results in grounding. Conversely, the MT scheme activates the fail-safe action and utilizes the remaining control to guide the vessel out of the narrow passage, allowing the vessel to drift safely for the remainder of the simulation ($D_{min} = 160.40$ m). The adaptability of the MT scheme is further highlighted in the scenario where the wind velocity drops from 16 to 5 m/s at $t = 80$. Once the drifting forces decrease, the degraded propeller force becomes sufficient to counter the drifting forces, allowing the MT scheme to bring the vessel to a stop.

5 Conclusion

This work proposes a Multi-Trajectory Risk-Aware MPC scheme that simultaneously optimizes a nominal and fail-safe trajectory. A shared initial control input constraint ($u_0^I = u_0^{II}$) is enforced, ensuring that every mission-driven action is also the first step of a fail-safe action. A supervisory system, which evaluates the predicted trajectories for grounding, the stopping capability and risk, utilizes a binary variable (B) to transition to the fail-safe mode. Simulations demonstrate that while the MT scheme increases the average solver time by two to three times compared to the ST scheme, the MT scheme reduces the likelihood of grounding compared to the ST scheme. Crucially, in situations where the remaining control is insufficient to counter drifting forces while executing a fail-safe action, the MT scheme focuses on guiding the vessel out of the narrow passage to drift where there is no immediate risk of grounding. Once the drifting forces do not overpower the remaining control, the vessel is brought to a stop. When the remaining control is insufficient to exit the narrow passage, grounding is inevitable. During scenarios that do not require a fail-safe action, the MT scheme matches the ST scheme in operational efficiency and tracking performance. Finally, to make real-world deployment possible, future research should build upon this work by addressing the following: soft grounding strategies, a fault diagnosis module, dynamic obstacles, complex drifting forces and sensor data handling.

References

[1] S. Blindheim, S. Gros, and T. A. Johansen, "Risk-based model predictive control for autonomous ship emergency management," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 14 524–14 531, 2020.

[2] SAFETY4SEA, "Tanker runs aground in the philippines; pcg monitors for oil spills," 2026. [Online]. Available: https://safety4sea.com/tanker-runs-aground-in-the-philippines-pcg-monitors-for-oil-spills/

[3] Davis Instruments, "What is navigation at sea?" 2026. [Online]. Available: https://www.davisinstruments.com/pages/what-is-navigation-at-sea

[4] N. Lodder, C. van der Ploeg, L. Ferranti, and E. Silvas, "Optimization-based fault mitigation for safe automated driving," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 1094–1100, 2023.

[5] M. Akdağ, T. A. Pedersen, T. I. Fossen, and T. A. Johansen, "A decision support system for autonomous ship trajectory planning," *Ocean Engineering*, vol. 292, p. 116562, 2024.

[6] S. Blindheim, T. A. Johansen, and I. B. Utne, "Risk-based supervisory control for autonomous ship navigation," *Journal of Marine Science and Technology*, vol. 28, no. 3, pp. 624–648, 2023.

[7] A. Tsolakis, "Rule-compliant and fault-tolerant motion planning: With application to autonomous surface vehicles," Dissertation (TU Delft), Delft University of Technology, 2025.

[8] T. I. Fossen, *Handbook of marine craft hydrodynamics and motion control*. John Wiley & Sons, 2011.

[9] Y. Song, Y. Chen, J. Gao, Y. Wang, and G. Pan, "Collision avoidance strategy for unmanned surface vessel considering actuator faults using kinodynamic rapidly exploring random tree-smart and radial basis function neural network-based model predictive control," *Journal of Marine Science and Engineering*, vol. 11, no. 6, p. 1107, 2023.