



City Research Online

City, University of London Institutional Repository

Citation: Aarsal, M., Asad, H., Kamel, T. & Khan, A. Cyber-Safety Assessment of Wind Turbines: A Reachability Analysis Approach Against Cyber-Attacks. Paper presented at the Sensei 2025, 9 Sep 2025, Stockholm, Sweden.

This is the accepted version of the paper.

This version of the publication may differ from the final published version.





Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/35350/>

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Cyber-Safety Assessment of Wind Turbines: A Reachability Analysis Approach Against Cyber-Attacks

Muhammad Aarsal¹✉ , Hafizul Asad² , Tamer Kamel¹ , and Asiya Khan¹ 
muhammad.arsal@plymouth.ac.uk, hafizul.asad@citystgeorges.ac.uk
tamer.kamel@plymouth.ac.uk, asiya.khan@plymouth.ac.uk

¹ University of Plymouth, Plymouth, UK

² City St George's, University of London, London, UK

Abstract. Cyber threats to Wind Power Plants (WPPs) are progressively rising as they often rely heavily on numerous digital assets and interconnected control systems. This makes WPPs more attractive to cybercriminals, as sabotaging these facilities can disrupt grid stability and energy supply. Most risk analyses of WPPs use informal frameworks or simulations, which can miss rare but critical scenarios, especially during cyberattacks, due to their non-exhaustive nature. This can compromise both security and safety. However, formal methods like model checking and theorem proving provide us with guarantees to ensure safety and stability. This paper presents the application of formal methods, particularly reachability analysis, to highlight the risks associated with wind plants. The focus is on model-based safety analysis of a wind turbine, including its pitch control system, with an emphasis on scenarios involving cyberattacks. We model the wind turbine system as a hybrid automaton based on its different control regions. We then perform reachability analysis of the hybrid automaton to examine all system states over a finite horizon, thus addressing the verification challenges inherent in such nonlinear dynamical systems. We identify vulnerabilities present in the system that attackers may exploit to cause harm to the plant. We conclude by discussing the impact of two different cyber attacks on the safety of the system.

Keywords: Security and Safety · Reachability Analysis · Hybrid Automaton · Cyber-Risk Analysis · Wind Power Plant Security

1 Introduction

The widespread integration of Cyber-Physical Systems (CPS) has significantly expanded the attack surface across modern infrastructure, creating unprecedented vulnerabilities. As CPS and IoT technologies proliferate, they introduce novel threats that traditional IT security frameworks struggle to address [1]. While standards like IEC 62443 have emerged to enhance industrial automation security [2], they fail to fully accommodate the dynamic, autonomous nature of

contemporary CPS [3] - a critical gap in sectors like energy where renewable technologies are rapidly evolving.

WPPs exemplify these challenges. Their operation depends on standards like IEC 61400-25 that enable cross-vendor communication and SCADA integration [4], creating a complex security-safety interdependency. While security protects sensitive data and controls, safety prevents physical hazards - yet security breaches can directly compromise safety, potentially causing equipment damage, operational failures, or environmental harm. This nexus demands rigorous risk analysis methodologies that can systematically identify, assess, and prioritise cyber-physical threats.

Current approaches often rely on standardised frameworks and simulations [5–7], which, while flexible, cannot provide exhaustive safety guarantees due to their inability to account for all edge cases. Formal verification methods like reachability analysis address this limitation by systematically examining all possible system behaviours under uncertainty, including cyberattack scenarios. By determining whether the system can reach unsafe states, these approaches provide mathematical assurances about system safety while revealing critical cyber-physical vulnerabilities.

1.1 State of the Art

There has been a lot of work done in the domain of risk analysis for WPPs. Several studies highlight vulnerabilities in wind energy systems and their SCADA. A survey on cyber-physical challenges for wind energy can be found in [8]. *Moness et al.* discuss safety and security risks due to cyber-physical integration, while *Staggs et al.* in [9] identify cyberattack vectors that could control wind turbines maliciously, risking physical damage. Cyber-attacks on SCADA systems of wind farms have been widely explored. For instance, the effects of altered SCADA reference parameters are examined in [10]. *Sabev et al.* use the Cyber Kill Chain in [11] to expose SCADA vulnerabilities, particularly through phishing. *Yang et al.* employ STRIDE in [12] to systematically assess cybersecurity threats in wind farms. Their framework, leveraging STRIDE, offers a structured approach to identifying, evaluating, and mitigating vulnerabilities. The impact of cyber-attacks on Wind Farm Active Power Control (WFAPC) is explored via simulation in [13].

Formal risk analyses often require a precise model of the system to verify critical properties. For example, [14] presents a timed automaton model to verify the safety properties of a wind turbine. Although it excludes dynamical behaviour of the system and cyber attacks, the focus is just to utilise model checking to verify the safety properties based on the timing of the state control mechanism. From a risk analysis perspective, a Bayesian graph model approach is used to represent the cyber-attacks on wind farms in [15], and frequencies of successful cyber-attacks are estimated. This can change with the scores for the stochastic model developed, depending on different security mechanisms. Recently, a modelling approach for the threat analysis of offshore wind farms by Bayesian Belief Networks is also been presented [16]. *Gabriel et al.* investigated the probabilities

of compromising offshore wind power plants, considering seasonal parameters too. A formal constraint satisfaction and optimisation problem framework to represent UFDI (Undetected False Data Injection) Attacks on meteorological sensors is presented in [17]. It is demonstrated that both power loss and attack vectors increase proportionally with the adversary’s capabilities.

Despite extensive studies, existing approaches fail to fully capture cyber risks in WPPs. There is no dedicated cyber-risk framework for WPPs or wind turbines. Frameworks like STRIDE, DREAD, MITRE ATT&CK, and the Cyber Kill Chain identify vulnerabilities, but have limitations. STRIDE and DREAD offer structured threat modelling but lack completeness [18]. Microsoft’s Security Development Lifecycle (SDL) advises documenting security notes before using STRIDE, but lacks guidance [19]. The Cyber Kill Chain focuses on attack stages but omits modern threats [5]. ATT&CK outlines adversarial tactics but lacks hierarchy, traceability, and structure [20]. These frameworks emphasise external threats while neglecting internal system properties, control specifications, and safety concerns. Simulations aid behavioural analysis but are constrained by predefined parameters, leaving gaps in vulnerability assessment [21].

Formal methods could enhance the safety of WPPs, but their hybrid, non-linear nature poses challenges. Safety can be framed as a reachability problem, yet reachability for such systems is undecidable due to complex state interactions [22]. One approach is to compute over-approximations of reachable states [23]. If this set avoids unsafe regions, safety is ensured; otherwise, the system’s safety remains undecidable. Barrier certificates provide another verification method, where a barrier function prevents system trajectories from reaching unsafe states [24]. However, finding a suitable barrier certificate is challenging. Such systems may also struggle with scalability and computational feasibility. This is why the safety analysis of WPPs, considering their physical dynamics, remains relatively unexplored.

1.2 Contributions

In this work, we propose a model-based approach for risk analysis and apply it to capture the cyber threats associated with the wind turbine system (WTS). We utilise hybrid system modelling to capture the important control modes and nonlinear dynamics of WTS. We define rotor overspeed as an unsafe state due to its potential to cause severe mechanical stress, component failure, safety hazards, and catastrophic turbine failure [25, 26]. We then explore the state space to determine whether any initial states can lead to rotor overspeed, with or without attacks, within a finite time horizon using reachability. The key contributions of this work are:

1. Model the WTS as a hybrid system to capture its various control modes alongside the nonlinear dynamical evolution of the system; representing the control modes as discrete states and the continuous evolution as the flow variables of a hybrid automaton.

2. Identification of vulnerable initial conditions using reachability; identifying the sets of initial states that are susceptible to exploitation if security is compromised.
3. Formal modelling of stealthy cyber-attacks (e.g., false data injection and parameter manipulation) with bounded adversarial capabilities, and verification of their impact on safety via reachability analysis.

The remainder of this paper is structured as follows: Section 2 reviews the preliminaries, presents the system and attacker model. Then, we model the hybrid automaton for WTS in Section 3. Accordingly, the experiment and analysis are demonstrated in Section 4. Section 5 briefly discusses the results of this experiment. Finally, Section 6 concludes this work and outlines the directions for future research.

2 Preliminaries

Definition 1 (Hybrid Automata). *Hybrid automata formally model systems that combine continuous dynamics (e.g., turbine rotation) with discrete transitions (e.g., mode switches). Formally, a hybrid automaton H is a tuple [27]:*

$$\langle Q, X, f, \text{Init}, \text{Dom}, E, G, R \rangle$$

where Q is the set of discrete states, $X = \mathbb{R}^n$ the continuous state space, and $f : Q \times X \rightarrow \mathbb{R}^n$ the vector field. $\text{Init} \subseteq Q \times X$ denotes initial states, $\text{Dom} : Q \rightarrow 2^X$ the domain, $E \subseteq Q \times Q$ the transitions, $G : E \rightarrow 2^X$ the guards, and $R : E \times X \rightarrow 2^X$ the reset map.

A hybrid state $(q, x) \in Q \times X$ evolves according to $\dot{x} = f(q, x)$ while $x \in \text{Dom}(q)$. When $x \in G(q, q')$, the discrete state may switch to q' , and x resets to a value in $R(q, q', x)$. The trajectory under input $u(\cdot) \in \mathbb{R}^m$ is denoted $\eta(t, x_0, q_0, u(\cdot))$.

Definition 2 (Reachability). *A state $(\hat{q}, \hat{x}) \in Q \times X$ is reachable if there exists a trajectory from $(q_0, x_0) \in \text{Init}$ that reaches it in finite time under some input $u(\cdot)$. Given initial states X_0 and input set \mathcal{U} , the exact reachable set at time t is [28]:*

$$\mathcal{R}^e(t) := \{\eta(t, x_0, q_0, u(\cdot)) \mid x_0 \in X_0, u(\kappa) \in \mathcal{U}, \kappa \in [0, t]\}$$

Since exact computation is intractable for general hybrid systems [29], we compute tight over-approximations: $\mathcal{R}(t) \supseteq \mathcal{R}^e(t)$. The cumulative reachable set over time horizon $[0, \kappa]$ is:

$$\mathcal{R}(0, \kappa) = \bigcup_{t \in [0, \kappa]} \mathcal{R}(t)$$

2.1 System Description

Wind turbines harness wind energy to power a generator, producing electricity. The aerodynamic power extracted by a wind turbine is given by [30]:

$$P_a = \frac{1}{2} C_p \rho \pi R^2 V^3$$

where ρ is air density, V is wind speed, and R is the turbine radius. The Betz coefficient C_p , which depends on turbine speed, wind speed, and blade pitch angle, is highly nonlinear. We use a one-mass drive train model expressed as:

$$\dot{\omega} = \frac{1}{J}(T_m - T_g) \quad (1)$$

where $T_m = \frac{P_m}{\omega}$ is the aerodynamic torque, T_g is the generator's reaction torque, J the moment of inertia, and ω the rotor speed. The WTS control policy varies with wind speed. i.e., below rated speeds; it maximises C_p using a near-zero pitch and a power controller [31]. The generator reaction torque is:

$$T_g = G_{opt} \cdot \omega^2$$

where G_{opt} is a constant, making torque proportional to rotor speed. Above rated wind speeds, pitch angle control is used to maintain constant speed, adjusting C_p by varying the pitch angle β . The pitch actuator, modelled as a first-order system:

$$\dot{\beta} = \frac{\beta_d - \beta}{\tau} \quad (2)$$

controls the pitch angle demand β_d from the controller, where τ is the actuator's time constant. A gain-scheduled PI controller for pitch angle, based on rotor speed, is given by:

$$\beta_d = K_p(\omega - \omega_{rated}) + K_i \int (\omega - \omega_{rated}) \quad (3)$$

where ω_{rated} is the rated rotor speed, and K_p, K_i are the proportional and integral gains of the PI controller. Equations (1), (2), and (3) define a system amenable to verification and validation.

$$\dot{x} = f(x) \quad (4)$$

where $x = [\omega, \beta, \beta_d]^T$. We consider rotor overspeed $\omega > \omega_{rated+}$ (speed more than maximum allowed speed) a violation of our safety property and want to check for its occurrence from a bounded set of initial conditions, particularly in the presence of cyber-attacks.

2.2 Attack Model

We define data integrity attacks (DIA) on the WTS and assess their impact on safety. We consider these attacks to be stealthy, meaning the attacker must avoid detection by anomaly detection or bad data detection mechanisms. Figure 1 illustrates such an attack. To explore scenarios leading to rotor overspeed (similar attacks on other system-level parameters can be constructed), the attacker launches a UFDI attack at time h , spoofing sensor data $y = \{\omega\}$ or altering control outputs $u = \{\beta_d, P\}$ via the communication channel, such that:

$$u_a(h) = \begin{cases} u(h), & \text{if } h \notin T_{\text{attack}} \\ u(h) - \Delta u, & \text{if } h \in T_{\text{attack}} \end{cases} \quad (5)$$

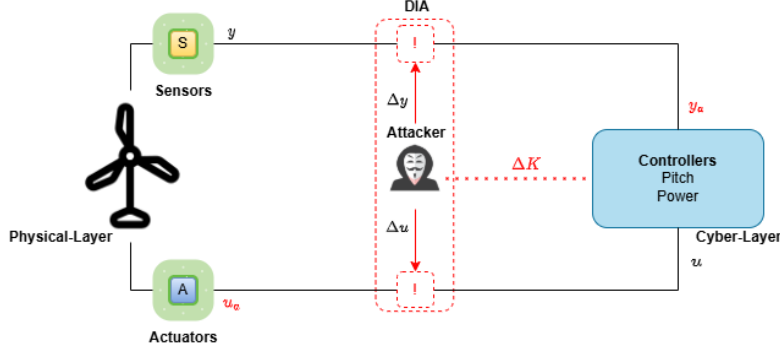


Fig. 1: Cyber attack on a Wind Turbine System

where Δu represents the injected attack signal and T_{attack} is the attack time interval. Such attacks introduce uncertainty into the system, potentially compromising safety. Additionally, we consider a parameter modification attack, wherein the attacker gains access to control parameters K through firmware manipulation:

$$K_a(h) = \begin{cases} K(h), & \text{if } h \notin T_{attack} \\ K(h) - \Delta K, & \text{if } h \in T_{attack} \end{cases} \quad (6)$$

where ΔK denotes the malicious parameter perturbations. The characteristics of the attack model are discussed as follows:

Attacker's Capability: An attacker can maliciously manipulate the controller's gain and control command along with the rotor speed measurements as long as it is below the unsafe threshold. We assume that the adversary is capable of making a successful attack every time as we are searching for the worst-case scenario.

Attacker's Target: The attacker aims to damage the wind turbine while remaining undetected within the system. The objective is to induce rotor over-speed by either manipulating control parameters through firmware modification or tampering with actuator inputs via the communication channel, such that $\omega > \omega_{rated+}$.

Attack Constraints: For a successful attack, the attacker must not manipulate the controller's gain and commands to such an extent that they can be detected. Furthermore, the attack term Δu or ΔK must be positive, as an increase in these parameters is usually capped and would be caught by detection mechanisms. i.e.,

$$\begin{aligned} 0 &\leq \Delta u \leq \alpha u \\ 0 &\leq \Delta K \leq \beta K \end{aligned}$$

where $\alpha \geq 0$ and $\beta \geq 0$ are constants that enable it to evade detection through bad data detection mechanisms. K can be either K_p or K_i . Even though these

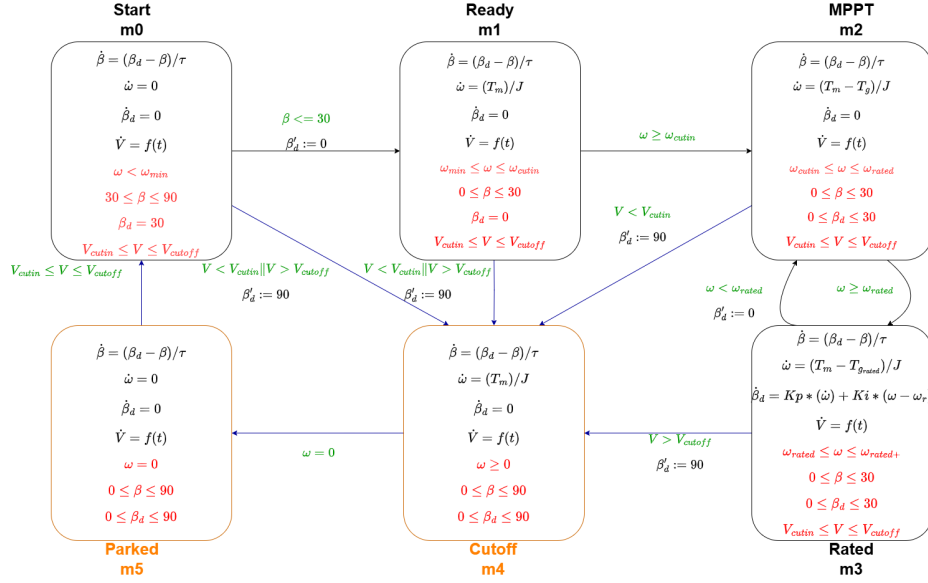


Fig. 2: Hybrid Automaton of a WTS

uncertain inputs are bounded by attack constraints, they raise verification challenges (infinitely many possible cases) that cannot be effectively addressed by conventional simulation-based testing alone.

3 Methodology

3.1 Hybrid Automaton of WTS

We propose a hybrid automaton to formally model the various control modes of the WTS. As the WTS inherently exhibits hybrid behaviour—integrating continuous dynamics and discrete events such as parking, startup, and power generation—it is naturally suited to such modelling. The parameters and structure of each subsystem adapt to transitions between working states identified through system analysis. Accordingly, we define six discrete operational states: *park*, *start*, *ready*, *mppt*, *rated*, and *cutoff*.

Figure 2 illustrates the hybrid automaton model of the WTS. Here, V_{cutin} and V_{cutoff} denote the startup and cutoff wind speeds, respectively. The relevant rotor speeds include ω_{\min} (startup threshold), ω_{cutin} , ω_{rated} , and $\omega_{\text{rated}+}$ (maximum safe limit). Discrete states m_0, m_1, \dots, m_5 represent operational modes, with transitions labelled by guard and reset conditions. The continuous dynamics in each mode are governed by ODEs derived from the WTS model.

The system transitions to *Start* when wind conditions become favourable. As the pitch angle decreases to its cut-in value and brakes are released, it enters *Ready*, where the turbine accelerates toward cut-in speed. Upon reaching it,

MPPT mode begins, maintaining minimal pitch to maximise power extraction via the power controller. At rated speed, the system switches to *Rated* mode, adjusting blade pitch to stabilise power output. If wind speeds exceed safe limits, *Cutoff* mode activates, increasing pitch to reduce aerodynamic torque. Finally, in *Park* mode, brakes engage and the pitch angle is maximized to stop the turbine.

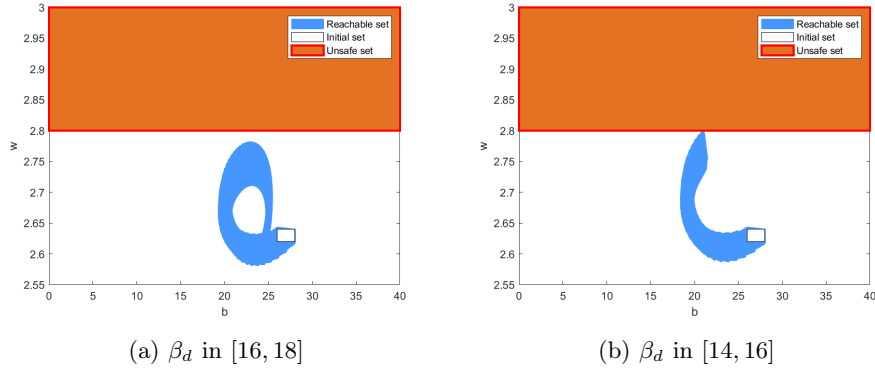
Some transitions depend on wind speed—a stochastic and uncontrolled variable. While high-fidelity models (e.g., polynomials or stochastic processes) could be used, they are computationally prohibitive for reachability analysis. To mitigate this, we fix the wind speed at a representative constant value, yielding a simplified but analyzable hybrid automaton. Importantly, our primary objective is to verify violations of the safety property, specifically rotor overspeed ($\omega > \omega_{\text{rated}+}$). Thus, certain modes, such as *Cutoff*, become less relevant in this context, as the analysis focuses on malicious interventions under constant wind conditions.

4 Experiment

This experiment conducts a reachability analysis for a WTS, focusing on identifying potentially unsafe states, especially under cyber attacks. An unsafe state occurs when rotor speed exceeds its upper limit, which happens in the *Rated* mode—our system’s most critical state. If one can guarantee the safety for all the possible initial conditions of this state, we may extend this to all other states. For our three-dimensional system, the invariant set is defined within the variable ranges $[2.58, 0, 0]$ to $[2.8, 30, 30]$, where 2.8 rad/s is unsafe as overspeed for SCIG wind turbines typically ranges from 5–8% of rated speed [32]. We explore the entire state space over 100 seconds.

We use **CORA** [33], a MATLAB toolbox for CPS verification via reachability analysis, computing over-approximated system states to ensure safety. The **Conservative Linear Algorithm** provides a safe approximation, with a **Taylor order of 5** for numerical computation of the next states of flow variables, a **Zonotope order of 20** for set-based enclosures, and a **0.1s time step** for high-resolution analysis. Experiments run on a **Core i5-1245U (12 CPUs), 16GB RAM, Windows 10**.

It must be clear that safety is ensured only if the unsafe set does not intersect with the reachable set. However, an intersection does not confirm unsafety—only that a definitive conclusion cannot be drawn. Large state variable ranges increase over-approximation errors and computational complexity, leading to false positives. To mitigate this, we slice ranges into a $10 \times 15 \times 15$ grid and analyse seven wind speeds above the rated speed to detect unsafe initial conditions. We also utilise reachability analysis to identify vulnerabilities that may emerge from malfunctions or security breaches by exploring whole ranges of these variables. For instance, Figure 3(a) shows the reach set when $\beta_d \in [16, 18]$ at 18 m/s wind speed. A small change, such as $\beta_d \in [14, 16]$, can shift the system toward unsafe behaviour, as in Figure 3(b), highlighting potential exploits for cyber-attacks like UFDI attacks.

Fig. 3: Analysis at ω in $[2.62, 2.64]$, β in $[26, 28]$ for $V = 18m/s$

As we have computed such reach sets for 2250 cases per wind speed, figure 4 presents a heatmap of unsafe initial conditions at different wind speeds, revealing that higher wind speeds increase the likelihood of unsafe states. This demonstrates the effectiveness of our approach in uncovering hidden system vulnerabilities, as these initial conditions could lead to unsafe behaviour in case of a cyber-attack or any control malfunction.

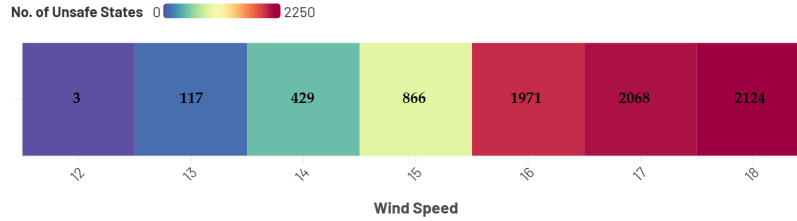


Fig. 4: Heatmap of Unsafe Initial Sets at Various Wind Speeds

4.1 Reachability Analysis under Cyber-Attacks

Upon identifying a set of vulnerable initial states, we exclude them from the set of initial conditions to get a set of conditions under which the system remains safe for a finite duration. These safe sets would represent the normal conditions for the operation of our system. Motivated by the literature, we investigate the impact of cyber-attacks on the system while considering these safe initial conditions.

Case Study 1: We assume that the attacker can decrease the output power by tampering with the electrical power command from the controller through an

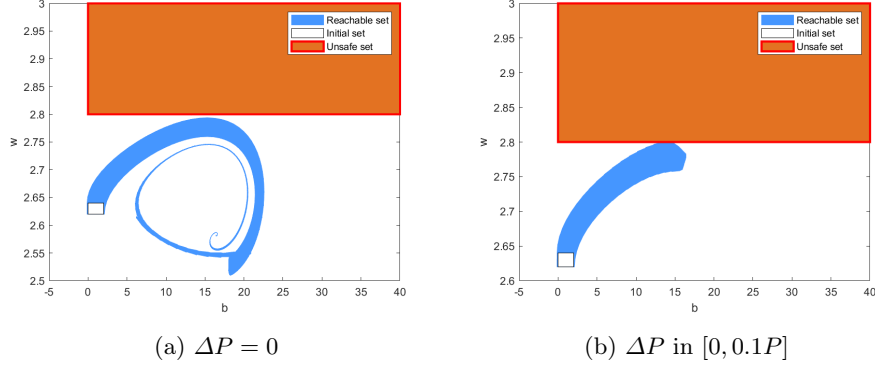


Fig. 5: Analysis at ω in $[2.64, 2.66]$, β in $[0, 2]$, and β_d in $[0, 2]$ for $V = 14$ m/s

unprotected communication channel. This scenario models a **Power Manipulation Attack**, where the attacker's capability is limited to altering the power signal by at most 10%, i.e.,

$$P_a = P - \Delta P, \quad \text{where} \quad \Delta P \in [0, \alpha P]$$

where $\alpha = 0.1$ to ensure the change remains stealthy and within system limits. This resembles a UFDI attack and effectively reduces the electrical torque, potentially leading to rotor overspeed due to insufficient electromagnetic braking. We thoroughly explore the state space by considering all safe states under this bounded power attack to evaluate the resiliency of the wind turbine control system. Figure 5(a) shows the system's behaviour when there is no cyber-attack, providing a baseline. In contrast, Figure 5(b) shows how the system behaves under the above-described uncertain power manipulation model.

Case Study 2: We now consider an adversary who has acquired knowledge of the control system and is capable of reducing the proportional gain of the pitch controller through firmware manipulation. This scenario models a **Gain Reduction Attack**, where the attacker maliciously alters the gain schedule within a feasible bound:

$$K_{p_a} = K_p - \Delta K_p, \quad \text{where} \quad \Delta K_p \in [0, \beta K_p]$$

where $\beta = 0.1$ represents the maximum tolerated reduction. This form of attack is subtle and likely harder to detect, especially since gain-scheduled controllers often have upper limits, making full compensation infeasible. We assess the impact of this attack through reachability analysis. As illustrated in Figure 6(a), the system remains within safe bounds when operating normally. However, under a 10% gain reduction attack, shown in Figure 6(b), the reach set intersects with the unsafe region. By performing such analysis over the full range of initial safe configurations, we identify regions where the control system may fail to ensure safety if such an attack occurs.

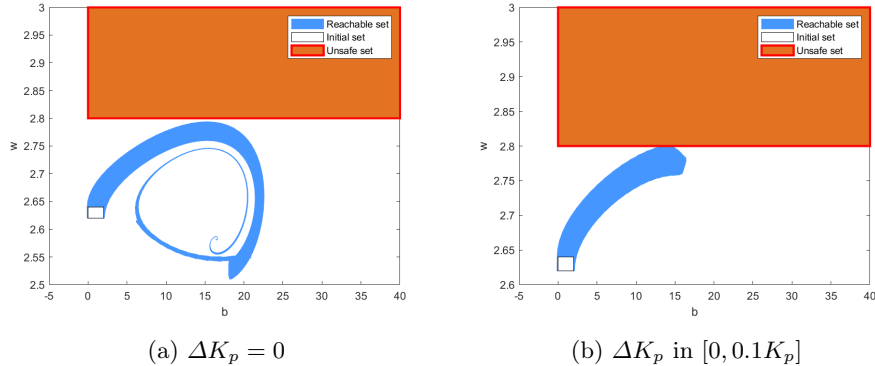


Fig. 6: Analysis at ω in $[2.62, 2.64]$, β in $[0, 2]$, and β_d in $[0, 2]$ for $V = 14$ m/s

5 Results and Discussion

We present a comprehensive formal modelling and verification approach to deal with cyber threats to WTS. Using hybrid automata, we construct WTS models that are well-suited for exhaustive verification. The verification process is based on reachability analysis, which identifies cyber attack scenarios that can lead towards rotor overspeed. Although in our experiments, we have investigated the rated mode because our unsafe state exists in that domain, it is certain that starting from this mode, WTS often visits the MPPT mode, describing the necessity of a Hybrid automaton. Our approach uncovers risks associated with system dynamics. Reachability results from other states are not shown due to the limitation of space as well. In contrast to the aforementioned informal frameworks and simulation, our work covers the whole state-space, exploring potential unsafe initial conditions, showing that small disturbances in controller or sensor values can lead to potentially catastrophic states where system safety cannot be guaranteed. The experimental results highlight the need for a highly resilient controller for the system. These findings indicate that the current controller, designed for normal operating conditions, lacks the resilience needed to address such scenarios effectively. We also consider two cyber attack case studies. The first scenario involves an uncertain power attack that reduces system power by 0–10%. We tested the system’s response to this UFDI attack at seven wind speeds, assuming initially safe conditions. As shown in Figure 7(a), the system is sensitive to wind speed variations, with a notable vulnerability at 15m/s , where the controller struggles to maintain safety. While this observation may vary between turbines, our approach lays the foundation for uncovering similar vulnerabilities in other systems for WPPs. The second scenario examines a controller gain manipulation attack. We consider the attacker to decrease the proportional gain of the system by 0 to 10% to check for such an attack whether the system remains safe or not. Figure 7(b) illustrates the reduction in safe states after the attack, similar to the power manipulation attack results. Most unsafe

cases occur at $15m/s$, indicating a lack of resilience in the controller’s design at this wind speed.

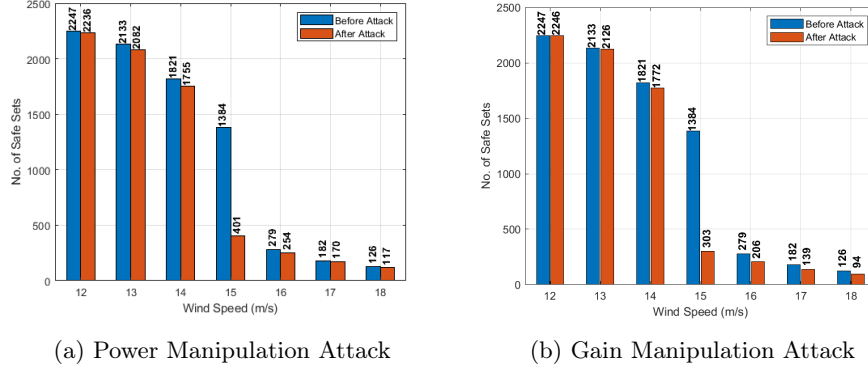


Fig. 7: Safe Sets before and After attacks

6 Conclusion

In this paper, we pursued an integrated approach to check the effects of cyber attacks on the safety of a WTS using reachability analysis. This study is the first of its kind to undertake the risk analysis of wind power plants at a highly granular level of abstraction while also accounting for their non-linear dynamics. The key finding is the identification of vulnerable initial states that an attacker may leverage to accomplish his malicious objectives. Conducting reachability analysis at different wind speeds also helps us identify the threats related to the gain-scheduled pitch control of the WTS. Moreover, a model of a stealthy attacker for DIA-type attacks is presented. It leads us to formalise cyber-attacks to find their impact on systems’ safety and reliability is prime as it uncovers the weakness related to control of the system. The adoption of hybrid automata and reachability analysis provides the basis to capture the risk associated with WPP dynamics.

This work opens several research directions. First, being time-bounded, it cannot guarantee long-term safety beyond computed reach sets. Second, treating wind speed as fixed limits the model’s scope; incorporating wind variability could improve accuracy. A promising approach is to model the stochastic nature of cyber attackers for enhanced security analysis. Another research direction can be the formal verification of the ML-based controllers of wind power plants because most of these systems are now data-driven. This verification can include ensuring that the model behaves correctly under adversarial inputs, falls within safe operating ranges, and satisfies specific functional properties.

References

1. Ayan Banerjee, Krishna K Venkatasubramanian, Tridib Mukherjee, and Sandeep Kumar S Gupta. Ensuring safety, security, and sustainability of mission-critical cyber-physical systems. *Proceedings of the IEEE*, 100(1):283–299, 2011.
2. Eider Iturbe, Erkuden Rios, Jason Mansell, and Nerea Toledo. Information security risk assessment methodology for industrial systems supporting isa/iec 62443 compliance. In *2023 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, pages 1–6. IEEE, 2023.
3. Withdrawal Date and Withdrawal Note. Archived nist technical series publication. *NIST Special Publication*, 800:60, 1992.
4. Brent Summerville. Distributed wind monitoring best practices. Technical report, National Renewable Energy Laboratory (NREL), Golden, CO (United States), 09 2024.
5. CYPHERE. What is cyber kill chain framework: Stages examples. <https://thecyphere.com/blog/cyber-kill-chain/>.
6. MITRE. Mitre att&ck. <https://attack.mitre.org/>.
7. Rafiullah Khan, Kieran McLaughlin, David Lavery, and Sakir Sezer. Stride-based threat modeling for cyber-physical systems. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6. IEEE, 2017.
8. Mohammed Moness and Ahmed Mahmoud Moustafa. A survey of cyber-physical advances and challenges of wind energy conversion systems: Prospects for internet of energy. *IEEE Internet of Things Journal*, 3(2):134–145, 2016.
9. Jason Staggs, David Ferlemann, and Sujeet Sheno. Wind farm security: attack surface, targets, scenarios and mitigation. *International Journal of Critical Infrastructure Protection*, 17:3–14, 2017.
10. Jie Yan, Chen-Ching Liu, and Manimaran Govindarasu. Cyber intrusion of wind farm scada system and its impact analysis. pages 1–6, 2011.
11. Evgeni Sabev, Galya Pavlova, Roumen Trifonov, Kamelia Raynova, and Georgy Tsochev. Analysis of practical cyberattack scenarios for wind farm scada systems. pages 420–424, 2021.
12. Baihua Yang and Yue Zhang. Cybersecurity analysis of wind farm industrial control system based on hierarchical threat analysis model framework. In *2022 International Conference on Computing, Communication, Perception and Quantum Technology (CCPQT)*, pages 6–13, 2022.
13. Mohammad Ali Ansari, Mohsen Ghafouri, and Amir Ameli. Cyber-security vulnerabilities of the active power control scheme in large-scale wind-integrated power systems. *Electrical Power and Energy Conference*, 2022.
14. Jagadish Suryadevara, Gaetana Sapienza, Cristina Seceleanu, Tiberiu Seceleanu, Stein-Erik Ellevseth, and Paul Pettersson. Wind turbine system: An industrial case study in formal modeling and verification. In Cyrille Artho and Peter Csaba Ölveczky, editors, *Formal Techniques for Safety-Critical Systems*, pages 229–245, Cham, 2014. Springer International Publishing.
15. Yichi Zhang, Yingmeng Xiang, and Lingfeng Wang. Power system reliability assessment incorporating cyber attacks against wind farm energy management systems. *IEEE Transactions on Smart Grid*, 8(5):2343–2357, 2017.
16. Alexander Gabriel, Babette Tecklenburg, Yann Guillouet, and Frank Sill Torres. Threat analysis of offshore wind farms by bayesian networks-a new modeling approach. In *Proceedings of the ISCRAM 2021 Conference Proceedings-18th International Conference on Information Systems for Crisis Response and Management, Omaha, NE, USA*, pages 28–31, 2021.

17. Amarjit Datta and Mohammad Ashiqur Rahman. Cyber threat analysis framework for the wind energy based power system. In *Proceedings of the 2017 Workshop on Cyber-Physical Systems Security and Privacy*, CPS '17, page 81–92, New York, NY, USA, 2017. Association for Computing Machinery.
18. Jarmo Alanen, Joonas Linnosmaa, Juha Pärssinen, Adrian Kotelba, and Eetu Heikkilä. Review of cybersecurity risk analysis methods and tools for safety critical industrial control systems. 2022.
19. Thomas Heyman. A formal analysis technique for secure software architectures. *KU Leuven*, 2013.
20. M Ruef and M Schneider. Mitre att&ck flaws of the standardization. URL: <https://www.scip.ch/en>, 2021.
21. Charles Burke Dawson. *Breaking things so you don't have to: risk assessment and failure prediction for cyber-physical AI*. PhD thesis, Massachusetts Institute of Technology, 2024.
22. Thomas A Henzinger, Peter W Kopke, Anuj Puri, and Pravin Varaiya. What's decidable about hybrid automata? In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, pages 373–382, 1995.
23. Meilun Li, Peter N Mosaad, Martin Fränzle, Zhikun She, and Bai Xue. Safe over- and under-approximation of reachable sets for autonomous dynamical systems. In *International Conference on Formal Modeling and Analysis of Timed Systems*, pages 252–270. Springer, 2018.
24. Stephen Prajna and Ali Jadbabaie. Safety verification of hybrid systems using barrier certificates. In *International Workshop on Hybrid Systems: Computation and Control*, pages 477–492. Springer, 2004.
25. Enyu Cai, Yunqiang Yan, Lei Dong, and Xiaozhong Liao. A control scheme with the variable-speed pitch system for wind turbines during a zero-voltage ride through. *Energies*, 13(13), 2020.
26. Fan Xinkai, Emanuele Crisostomi, Baohui Zhang, and Dimitri Thomopoulos. Rotor speed fluctuation analysis for rapid de-loading of variable speed wind turbines. pages 482–487, 06 2020.
27. Jean-François Raskin. An introduction to hybrid automata. In *Handbook of networked and embedded control systems*, pages 491–517. Springer, 2005.
28. John Lygeros. Lecture notes on hybrid systems. In *Notes for an ENSIETA workshop*, 2004.
29. Gerardo Lafferriere, George J Pappas, and Sergio Yovine. A new class of decidable hybrid systems. In *Hybrid Systems: Computation and Control: Second International Workshop, HSCC'99 Berg en Dal, The Netherlands, March 29–31, 1999 Proceedings 2*, pages 137–151. Springer, 1999.
30. Magdi Ragheb and Adam M Ragheb. Wind turbines theory-the betz equation and optimal rotor tip speed ratio. *Fundamental and advanced topics in wind power*, 1(1):19–38, 2011.
31. Saravanakumar Rajendran and Debashisha Jena. Control of variable speed variable pitch wind turbine at above and below rated wind speed. *Journal of Wind Energy*, 2014(1):709128, 2014.
32. Nikolai Kulev and Frank Sill Torres. Simulation of the impact of parameter manipulations due to cyber-attacks and severe electrical faults on offshore wind farms. *Ocean Engineering*, 260:111936, 2022.
33. Matthias Althoff. An introduction to CORA 2015. In *Proc. of the 1st and 2nd Workshop on Applied Verification for Continuous and Hybrid Systems*, December 2015.