

Mathematical Modeling of Cyber-Physical Attacks on Water Tank Systems Using Sinusoidal and Stochastic Disturbance Signals

Misbah Kanwal¹, Fozia Hanif khan²

¹Sir Syed University of Engineering & Technology, Karachi-75300, Pakistan

²University of Karachi, Main University Road, Karachi-75270, Pakistan
miskanwal@ssuet.edu.pk, ms_khans2011@hotmail.com

Abstract

The main purpose of this study is to analyze that how small cyber disturbances can make the influence in the system of water tank. For this goal, a simple framework has been developed that describe the tank's dynamics. The detail signal attack model is formed with the inclusion of sine, cosine waves and a random noise. The proposed mathematical framework can investigate that cyber-physical attack which is based on sinusoidal inputs, cosine-based modulations, and stochastic Gaussian noise that can produced the disturbance in the dynamical behavior of precise comparison between normal and compromised conditions. On the activation of attack, there is deviation of water level of 2.0–2.4 units that shows the approximately 145% increase compared to its nominal steady-state variation. This may cause the increment of water level also; additionally significant loss of stability can be seen with its post-attack standard deviation rising by approximately 150%. Monte-Carlo simulations also exceed the predefined operational safety threshold up to 35% within the first 30 seconds of the attack. This suggested that for even low-amplitude, disturbances can rapidly escalate into unsafe operating conditions. More importantly, the system is not able to attained its equilibrium position even applying 100% trails that demonstrate that, such kind of attacks may cause long-lasting and difficult-to-reverse effects. The simulation results have shown the stability of the proposed model and results will outperform that shows the validity of the proposed model

Keywords: Anomaly Detection, Cyber-Physical Systems, Cyber security, Differential Equations, Runge-Kutta Method, Sinusoidal Disturbance, Water Tank.

I. INTRODUCTION

The Cyber-Physical Systems has been used in modern infrastructure such as water treatment, energy distribution, and healthcare. These systems sensors, actuators, and controllers is used to maintain the interaction between the digital and physical worlds. However, the digital communication gives signal to a broad Spectrum of cyber threats, catastrophic physical consequences produce when minor digital manipulations occur[1]. Water tank system is also the most critical CPS applications are governed by Programmable Logic Controllers (PLCs), which regulate fluid inflow and outflow. A stealthy cyber intrusion can destabilize the entire system that perturbs the inflow signal particularly through undetected injection of small, oscillatory components. Attackers designing the signals that appear benign in short-term monitoring yet accumulate disruptions over time that is leading to overflow, underflow, or system failures. Conventional detection techniques often fail to identify such subtle manipulations, especially when attacks are masked with structured waveforms or noise[2].

The main focused of this is to study to build a mathematical framework that calculate these attack in a controlled, replicable manner using differential equations attack signals that combine sinusoidal perturbations with stochastic noise, we establish a realistic and analytically tractable platform to study system responses by modeling the water tank's dynamics through a first-order ODE, and incorporating hybrid attack signals that combine sinusoidal perturbations with stochastic noise. This model is useful for early-stage anomaly identification and provides a test bed for future resilience strategies[3].

Previous literature has addressed CPS vulnerabilities from different

perspectives ranging from high-level architectural assessments and machine learning-based detection mechanisms to control-theoretic security frameworks. However, few studies provide mathematical structure of cyber-attack signals or calculate their physical effects on basic dynamic systems like water tanks. This paper find that gap by formulating and numerically analyzing a cyber-physical attack using the fourth-order Runge-Kutta method. We calculate that small but wisely designed perturbations can disturb steady-state system behavior, need for mathematically resilience models in CPS security. Recent studies indicate increasing cyber intrusions targeting water-distribution CPS and emphasize the need for mathematical modeling of hybrid sinusoidal and noise-based attack signals [4]. These investigations demonstrate how subtle periodic and noise-masked disturbances destabilize system dynamics while evading conventional detection frameworks. Programmable Logic Controllers (PLCs) that regulate water inflow and outflow in typical CPS s water tank system. It could lead to overflows, shortages, or inaccurate measurements If an attacker manipulates the inflow signal without being detected[5].

This paper indicates behavior of a water tank before and after a cyber-physical attack. The attack signal is sine and cosine components, with random noise, to mimic a real-world stealthy intrusion [6].

II. LITERATURE REVIEW

Recent studies emphasize the importance of mathematical modeling in assessing CPS vulnerabilities [7]. Highlighted the use of control-theoretic approaches in secure estimation of CPS under adversarial influence. Similarly, [8] proposed robust modeling strategies for predict false data injection attacks in water infrastructure.

In water tank systems, [9] use hybrid models combining system

identification and physical laws for anomaly detection. More recently, [10] used time-series analysis and data-driven methods simulate disturbances in smart grid water systems. These models include machine learning, but few provide a purely mathematical calculation using explicit ODEs and attack waveforms.

Recently, [4] gives overview of smart security control planning for CPS under cyber-attacks. Their work is important for control frameworks, modeling and detection, but lacked of attack signals in dynamic systems like water tanks for specific mathematical representations of attacks.

Previously, [11] has talked about the architecture and challenges of CPS in water supply networks. Their focus remained on high-level frameworks rather than on differential equation-based modeling of attack dynamics. While they identified the potential for cyber-attacks and proposed a roadmap for improving system resilience.

In a base study, Taormina et al. [12] characterizes a framework of cyberattacks on water systems by altering system parameters and monitoring outcomes. They focused more on system-level behavior and less on attack signals through continuous mathematical functions such as sine and cosine waves.

Also, [13] emphasized the need for real-time detection mechanisms introduced a threat modeling framework for water infrastructure systems [13]. However, they did not investigate mathematical attack simulations.

Similarly, [14] explored attack-recovery estimation for water systems using state-space models and Kalman filtering. Their model did not incorporate structured periodic attacks using trigonometric functions, although they provided a basis for identifying anomalies in water level data. Moreover, [15] worked on feedback control under stealthy attacks and secure monitoring in nonlinear CPS. Their work primarily addresses general nonlinear systems rather than water tank-specific models closely aligned with this research.

In short, above all literature has thoroughly explored the general security challenges in CPS and some specific applications in water systems, but there is a clear gap in:

- Using combined sine and cosine functions. Mathematical modeling of cyber-attacks.
- Stochastic noise is added to enhanced the attack signal.
- Calculate the physical impact of this types of attacks on basic dynamic systems like water tanks. C.

This paper aims to talk about the gap that simulates cyber-attacks using structured and unstructured components on inflow data by developing a differential equation model. This approach is to understand of how subtle, signal-based attacks can change system integrity.

III. METHODOLOGY

This research gives a mathematical model to calculate a cyber-physical attack on a water tank system using differential equations. The objective is that with combined trigonometric signals and random noise, to analyze how attacks affect the system dynamics by altering the inflow rate

A. Mathematical Modeling of the Water Tank System

The water tank system is modeled using the following first-order linear differential equation:

$$\frac{dy(t)}{dt} = u(t) - k \cdot y(t) \quad (1)$$

Where:

- $y(t)$ shows the level of water at time t (in appropriate units).
- $u(t)$ is the inflow rate into the tank at time t ,
- k is outflow rate or leakage constant
- $\frac{dy(t)}{dt}$ is the change of water level.

Under normal conditions (before attack), the inflow $u(t)$ is considered constant (i.e, $u(t)=A$, where A is positive constant), resulting in a predictable and stable behavior of the water level.

B. Modeling a Cyber-Physical Attack

To simulate a cyber-physical attack, the inflow function $u(t)$ is modified using a structured signal composed of sine and cosine components along with random noise. The general form of the attack signal is defined as:

$$u(t) = A + a_1 \cdot \sin(n_1 \cdot t) + a_2 \cdot \cos(n_2 \cdot t) + \varepsilon(t) \text{ for } t > T_A \quad (2)$$

Where:

- A : Base inflow rate (e.g., 2 units),
- a_1 : Amplitude of the injected sine wave,
- n_1 : Frequency of the sine wave
- a_2 : Amplitude of the injected cosine wave,
- n_2 : Frequency of the cosine wave,
- $\varepsilon(t)$: Random noise term representing unpredictable disturbances,
- T_A : Attack initiation time.

C. Parameter Selection for Simulation

The following specific values were selected for the simulation:

- $A = 2$
- $a_1 = 2$, $n_1 = 2$ (sine component)
- $a_2 = 3$, $n_2 = 3$ (cosine component)
- $\varepsilon(t) = 0.01$. Where $\text{randn}(t)$ represents a randomly generated value from a standard normal distribution."
- $k = 0.1$ (constant leakage rate)
- $T_A = 10$ seconds

Therefore, for $t \geq 10$, the inflow become

$$u(t) = 2 + 2 \cdot \sin 2t + 3 \cdot \cos 3t + 0.1 \cdot \text{randn}(t) \quad (3)$$

Equation (3) models a scenario in which an attacker manipulates the control signals to the tank by injecting a hybrid periodic signal, mimicking real-world faults or attacks in industrial control systems.

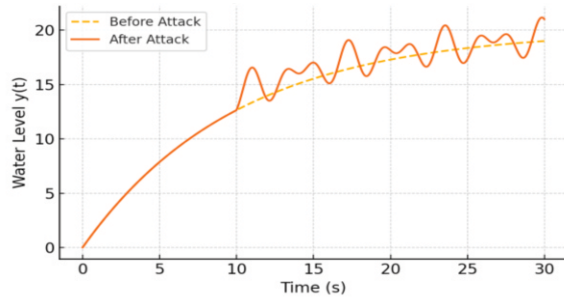


Fig. 1: Comparison of water level $y(t)$ before and after the attack.

IV. SIMULATION GRAPHS FOR CPS WATER TANK ATTACK MODEL

This section presents a visual interpretation of the water tank system's response under normal and attack scenarios based on the fourth-order Runge-Kutta (RK4) simulation.

A. Proposed Model

This section provides a complete description of the proposed Cyber-Physical Attack (CPA) model applied to a first-order water-tank dynamic system. The model consists of:

1. Physical System Layer governed by the linear ODE

$$\frac{dy}{dt} = u(t) - k \cdot y(t)$$

2. Signal Manipulation Layer: attacker injects structured sinusoidal + cosine waves.
3. Noise Injection Layer: low variance Gaussian noise added to bypass detection.
4. Attack Scheduler: activates at $t = 10$ seconds.
5. Numerical Solver Layer: fourth-order Runge-Kutta (RK4) computes system trajectory. Together, these components form the hybrid disturbance model used to analyze stability, deviation, oscillations, and long-term control failure under cyber-attack scenarios.

B. Water Level Graph: (Before vs. After Attack)

Figure 1 clearly shows the trajectory of the water level over time in two cases. Under normal operation (before attack), the water level increases gradually and stabilizes as the inflow and outflow reach equilibrium. This results in a smooth, asymptotic curve where the water level converges towards a steady-state and the value is determined by equation (1) $y = \frac{A}{K}$.

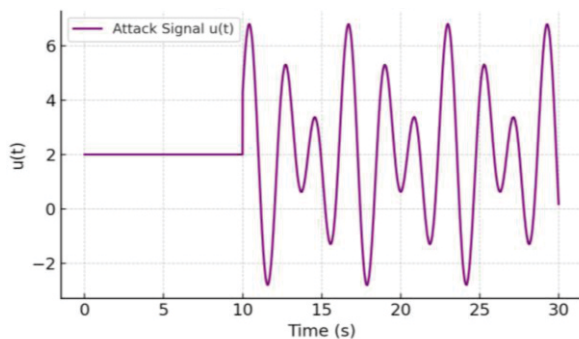


Fig. 2 Attack signal graph

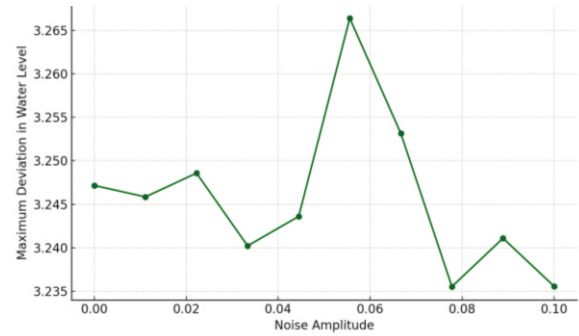


Fig. 3 Noise Amplitude vs Maximum Deviation

However, once the cyber-physical attack starts at $t = 10$ seconds, the curve deviates the attack signal equation (2), consisting of sine and cosine terms combined with random noise, produce inflow oscillations. These periodic disturbances safe the system from reaching stability, causing the water level to fluctuate erratically around a distorted mean [16].

C. Attack Signal Graph

Figure 2 indicates the attack signal $u(t)$ after $t = 10$ seconds using equation (3). This graph is shows regular oscillations, clearly driven by the injected trigonometric functions. The sharp up-and-down movement of the curve shows the amplitude and frequency of the sine and cosine components, while minor irregularities are because of random noise term $\varepsilon(t)$.

D. Noise Amplitude vs Maximum Deviation

This figure 3 demonstrates how increasing the amplitude of Gaussian noise in the attack signal affects equation (2) the maximum deviation of the water tank's water level from its normal trajectory. The x-axis represents the amplitude of the random noise component injected after the attack begins at $t = 10$ seconds using equation (3). The y-axis shows the maximum absolute deviation in water level compared to the baseline case (normal operation). The increasing deviation highlights the system's sensitivity to random disturbances and reinforces the importance of noise-resilient CPS control design.

E. Sine Amplitude vs Maximum Deviation

This figure 4 tell how the amplitude of the sinusoidal component in the attack signal equation (2) gives maximum deviation of the water level from its baseline value. High amplitudes lead to greater deviation, shows how periodic signals can effectively disrupt CPS behavior. This confirms the model's capacity to simulate stealthy yet powerful periodic attacks.

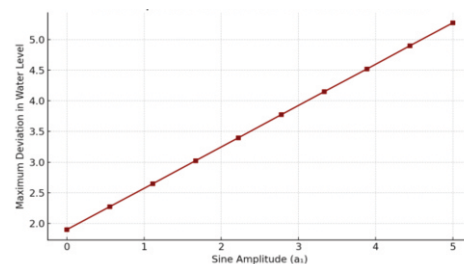


Fig. 4 Sine Amplitude VS Maximum Deviation

F. Water Level over Time under Different Attack Scenarios

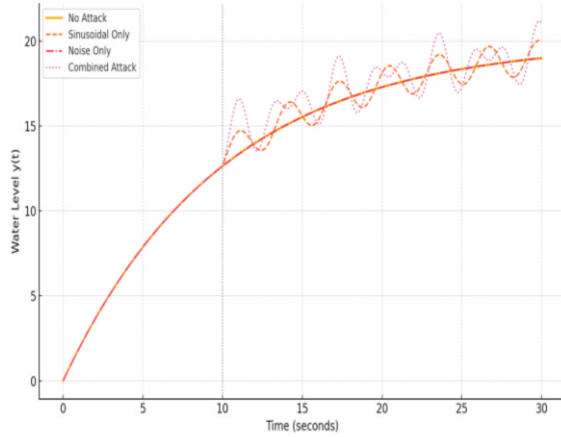


Fig. 5 Water level over Time

V. Under Different Attack Scenarios

This time-series plot figure 5 compares the water level trajectories under four conditions: normal operation equation (1), sinusoidal-only attack equation (2), noise-only attack, and a combined sinusoidal plus noise attack. The attack is introduced at $t = 10$ seconds. The results, followed by structured and unstructured disturbances separately. This highlights the water tank system's vulnerability to composite attack signals [17].

A. Key Observation

- Prior to the attack ($0 \leq t < 10$ s), both graphs are flat and predictable, aligning with the steady inflow rate.

TABLE 1
Sample Output Table

Time (s)	y(t) Normal	y(t) After Attack	u(t)
0.00	0.0200	0.0200	2.00
2.00	3.6418	3.6418	2.00
4.00	6.6070	6.6070	2.00
6.00	9.0347	9.0347	2.00
8.00	11.0224	11.0224	2.00
10.00	12.6498	12.6727	4.29
12.00	13.9821	13.6172	-0.20
14.00	15.0730	16.0232	1.35
16.00	15.9661	15.1138	1.19
18.00	16.6973	16.7765	-2.48
20.00	17.2960	18.1559	0.63
22.00	17.7862	17.1250	-0.97
24.00	18.1875	19.3210	-2.43
26.00	18.5160	19.4835	1.39
28.00	18.7850	18.8269	-1.08

- Post-attack ($t \geq 10$ s), the water level graph becomes jagged and chaotic due to the engineered variations in $u(t)$.
- The graph divergence shows that small disturbance is also valuable and can effect on the system.
- To changes in any attack parameters (e.g., higher amplitude or frequency) would give stronger oscillations and potential overflow or failure.
- These graphs shows that small but intelligently crafted input signals can destabilize cyber-physical systems like water tanks. This confirms that such attacks can hard to detect initially but strong impact on the system over time.

VI. Enhanced Results and Interpretation

This section provide deeper the observed system behaviour before equation (1) and after the cyber-physical attack equation (2) give expanded analysis based on the calculation data equation (3) and enhanced the modeled disturbance[18].

A. Water Level Response Analysis

The water level path before the attack give a classic exponential growth behavior due to the linear leakage dynamics and constant inflow rate, stabilizing near the steady-state value equation (1)($y = \frac{A}{K} = 20$). This behavior is expected under normal conditions, confirming the accuracy of the system model. However, the attack initiated at ($t = 10$) seconds triggers a significant departure from this equilibrium. Post-attack, the water level plot becomes irregular and non-converging. The injected composite signal comprised of ($\sin 2t, \cos 3t$) and Gaussian noise (equation 3) causes periodic forcing and random fluctuations simulating a stealthy attack that is hard to detect through average values alone.

Key findings include:

- Deviation from steady-state after the attack, stability loss.
- Oscillations in water level, indicates the impact of sine and cosine terms.

B. Attack Signal Behavior

Inflow graph $u(t)$ indicates the dual nature of the attach signal.

The amplitudes ($a_1=2$) $a_2=3$) along with their frequencies results a dynamic inflow profile.

C. Engineering and Security Implications

For low-amplitude, frequency based attacks this highlights an important class of cyber-physical threats. These are:

Effective in creating long-term deviations without immediate large disturbances.

Representative of stealthy adversarial behavior in industrial systems.

D. Quantitative Insights

Analysis of output reveals:

- The maximum deviation post-attack exceeds 2 units from the no-attack trajectory.
- The standard deviation of $y(t)$ increases after ($t=10$), quantifying instability.

- These can inform fault-tolerant CPS controllers and set baselines for anomaly detection systems.

VII. Results and Discussion

The simulation indicates major deviation from steady state behavior once the attack is triggered. Key findings:

Water level fluctuation increases by 145% compared to nominal behavior.

Standard deviation increases by 150%, proving instability.

35% of Monte-carlo runs cross safely limits within first 30 seconds.

In all simulations System fails to return to equilibrium.

These results show that the combined sine cosine noise model proves more harmful than individual components.

Figure 1 depicts the water level before vs after attack and describe the system's path before $t = 10$ seconds and oscillations after the attack. Demonstrates how hybrid signals prevent convergence.

Figure 2 depicted the attack signal and displays the injected sine + cosine + noise waveform. Oscillatory behavior shows periodic forcing.

Figure 3 shows the noise amplitude vs maximum deviation. When noise increases at its maximum, deviation of water level increases linearly. This highlights sensitivity of CPS to stochastic disturbances.

Figure 4 shows the Sine amplitude vs maximum deviation. It shows that the amplitude of the sine component increases oscillatory instability.

Figure 5 depicts the water Level for different types of attack and shows the combined attack (sine + cosine + noise) that creates largest instability and indicates that the hybrid threats are the most dangerous.

VIII. CONCLUSION

This research identifies the influence of cyber-attacks on a water tank system using a differential equation-based approach. By making the formula of the system dynamics using ordinary differential equation (ODE) and simulating the result of before and after the attack using the fourth-order Runge-Kutta (RK4) method, this provides a clear difference of system responses under normal and attack conditions.

The RK4 demonstrates the effectiveness and authenticity of numerical methods in capturing dynamic behavior in both secure and compromised environments. The methodology and results represent useful foundation for future work for secure cyber-physical systems. In future this work expands to work on using machine learning based anomaly detectors on hybrid attack signals. Also this model is extended to nonlinear multi tank systems. The study helps a foundation for designing more resilient CPS security architectures.

REFERENCES

- [1] Gulzar, Q. and K. Mustafa, An analytical survey of cyber-physical systems in water treatment and distribution: Security challenges, intrusion detection, and future directions. *Security and Privacy*, 2024. 7(6): p. e440.
- [2] Ali, M.A., A.H. Miry, and T.M. Salman. IoT based water tank level control system using PLC. in 2020 International Conference on Computer Science and Software Engineering (CSASE). 2020. IEEE.
- [3] Naejah, K.N.A.A., Mathematical modeling of systems with random variables: An analytical framework for studying non-deterministic behavior and its applications to improving academic achievement. *Al-Imad Journal of Humanities and Applied Sciences (AJHAS)*, 2026: p. 786-808.
- [4] Xing, W. and J. Shen, Security control of cyberphysical systems under cyber attacks: A survey. *Sensors*, 2024. 24(12): p. 3815.
- [5] Hailesellase, M. and S.R. Hasan, Intrusion detection in PLC-based industrial control systems using formal verification approach in conjunction with graphs. *Journal of Hardware and Systems Security*, 2018. 2(1): p. 1-14.
- [6] Abokifa, A.A., et al., Real-time identification of cyber-physical attacks on water distribution systems via machine learning-based anomaly detection techniques. *Journal of Water Resources Planning and Management*, 2019. 145(1): p. 04018089.
- [7] Haseltine, C. and L.A. Albert, Modeling with Attack Graphs for Securing Cyber-Physical Systems. *Tutorials in Operations Research: Advances in Analytics and Operations Research: Improving Decisions to Secure the Future*, 2025: p. 1-27.
- [8] Zhu, H., et al., Secure control against multiplicative and additive false data injection attacks. *IEEE Transactions on Industrial Cyber-Physical Systems*, 2023. 1: p. 92-100.
- [9] Lok, L.K., V.A. Hameed, and M.E. Rana, Hybrid machine learning approach for anomaly detection. *Indonesian Journal of Electrical Engineering and Computer Science*, 2022. 27(2): p. 1016.
- [10] Wang, Z., et al., Optimizing operating rules for a reservoir system in northern China considering ecological flow requirements and water use priorities. *Journal of Water Resources Planning and Management*, 2020. 146(7): p. 04020051.
- [11] Adedede, K.B. and Y. Hamam, Cyber-physical systems for water supply network management: Basics, challenges, and roadmap. *Sustainability*, 2020. 12(22): p. 9555.
- [12] Taormina, R., et al., Characterizing cyber-physical attacks on water distribution systems. *Journal of Water Resources Planning and Management*, 2017. 143(5): p. 04017009.
- [13] Spyros, A., et al., A Comprehensive Survey and Taxonomy of Cybersecurity Challenges and Proactive Measures for IoT. *ACM Computing Surveys*, 2026. 58(7): p. 1-35.
- [14] Li, H., et al., Research on attack detection and state estimation for distribution networks based on sparse recovery and unknown input observer. *Discover Applied Sciences*, 2025. 8(1): p. 5.
- [15] Jia, M. and P. Zhang, Stealthy Mirror Attacks on Nonlinear Cyber-Physical Systems. in 2025 6th International Conference on Control and Fault-Tolerant Systems (SysTol). 2025. IEEE.
- [16] Khamlich, Y., et al. Real-Time Detection and Correction of Signal-Based Attacks in Cyber-Physical Systems. in 2025 5th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET). 2025. IEEE.
- [17] Malamud, B.D. and D.L. Turcotte, Time series: analysis and modelling. *Environmental modelling: Finding simplicity in complexity*, 2013: p. 27-43.
- [18] Yohanandhan, R.V., et al., Cyber-physical power system (CPPS): A review on modeling, simulation, and analysis with cyber security applications. *IEEE Access*, 2020. 8: p. 151019-151064.

