Aalto-yliopisto
Aalto-universitetet
Aalto University

Bachelor’s Programme in Electrical Engineering

# Estimation Based Methods for Cyber-Attack Resilience

System resiliance under false data injection attacks on sensors

**Jonathan Ramstedt**

Bachelor’s Thesis
2025

Aalto-yliopisto
Aalto-universitetet
Aalto University

**Author** Jonathan Ramstedt

**Title** Estimation Based Methods for Cyber-Attack Resilience — System resiliance under false data injection attacks on sensors

**Degree programme** Electrical Engineering

**Major** Automation and Robotics

**Teacher in charge** D. Sc. Pekka Forsman

**Advisor** D.Sc. Shreya Das

**Date** 12 December 2025 **Number of pages** 22+1 **Language** English

**Abstract**

Control systems are widely used in industry as a means of controlling processes and actuators. A control system uses an internal mathematical model for steering the system output towards a desired value based on input. These systems appear in factories and similar sites. Within these, the control system is often part of the local network. In the network, the system receives signals as input. The source of these signals is often sensors that are also connected to the network. This gives the control system wider exposure to cyber-security-related risks. False data injection (FDI) attacks are a common form of attack performed against control systems. FDI attacks against sensors have been investigated as part of this thesis. Estimating the internal state reduces the dependence of accurate sensor data and can therefore theoretically be viewed as a way to build resiliency against FDI. Methods for state estimation are already in wide use in modern control systems. The goal of this thesis is to evaluate the effectiveness of state estimation as an additional layer of security against FDI attacks. The thesis covers the threat landscape of FDI attacks and a deeper understanding of state estimation through a literature review. In addition, simulation is used as a research method. A control system for a device moving in a plane is subject to FDI. By implementing state estimation methods for the control system, these are evaluated in their ability to build resilience against FDI attacks. The nature of the measurement data allowed for a linear control system, allowing the implementation of a Luenberger observer and a Kalman filter. The simulation proves that the chosen methods can provide security for the chosen system when the injected data is noise with zero mean. The same could not be proven for the situation when the data is constant bias. Presentation and analysis of the results indicate that the effectivness of these methods should be viewed as system specific. In addition, the studied methods provided do not provide comprehensive nor general proof of resilience against FDI attacks.

**Keywords** Cyber-security, Estimation, Operational Technology, False Data Injection

Aalto-yliopisto
Aalto-universitetet
Aalto University

**Författare** Jonathan Ramstedt

**Titel** Uppskattning som Metod för Resiliens mot Cyberattacker

**Utbildningsprogram** Kandidatprogrammet i elektroteknik

**Huvudämne** Automation och Robotik

**Ansvarslärare** D. Sc. Pekka Forsman

**Datum** 12 December 2025 **Sidantal** 22+1 **Språk** engelska

**Sammandrag**

Kontrollsystem som möjliggör automatisk styrning av system och processer är i användning inom många tillämpningsområden i industrin. Ett kontrollsystem använder en intern matematisk modell för att, på basis av systemets indata, styra systemet mot ett måltillstånd. Dessa system förekommer i fabriker och andra liknande anläggningar. Inom dessa tenderar kontrollsystemen att vara kopplade till ett lokalt nätverk. Via nätverket mottar kontrollsystem data som ofta härstammar från sensorer. Som följd av detta exponeras kontrollsystem till ett bredare spektrum av cyberattacker. En vanlig attack mot kontrollsystem är falsk datainjektion (FDI). I detta arbete har FDI attacker där data från sensorer manipuleras forskats. Uppskattning av det interna tillståndet minskar beroendet på sensorers mätdata och kan därför i teorin bygga resiliens mot FDI. Uppskattning av det interna tillståndet har redan bred tillämpning inom moderna kontrollsystem. Syftet med detta arbete är att evaluera uppskattningens effektivitet som ett extra lager av säkerhet mot FDI attacker. Genom litteraturöversikt bygger arbetet en definition och hotbild av FDI attacker samt djupare förståelse för uppskattning av tillstånd hos kontrollsystem. Utöver detta har simulation använts som metod för forskning. Ett system med mål att röra sig i ett plan utsätts i virtuell simulering för injektion av falsk data. Genom att implementera metoder för att uppskatta tillståndet kunde dessa evalueras i sin förmåga att ge systemet resiliens mot injektioner av data. Mätdatan var sådan att systemet i fråga är linjärt och tillståndet uppskattas separat med hjälp av en Luenberger-observatör och ett Kalman-filter. I simuleringen påvisas att de valda metoderna kan förse systemet med säkerhet i situationer då datan tar form av brus. Det kan också konstateras att då injicerade datan var en konstant förskjutning hade attacken klara negativa effekter på systemet. Framställning och analys av resultaten hänvisar till att dessa metoder ska beaktas som systemspecifika och att ingen heltäckande lösning framställs i detta arbete.

**Nyckelord** Informationssäkerhet, Uppskattning, Operativ Teknologi, Falsk Data Injektion

# Contents

# Symbols and abbreviations

## Symbols

| Symbol | Description |
|---|---|
| $x_k$ | System state at discrete time $k$ |
| $\hat{x}_{k\|k-1}$ | A priori state estimate |
| $\hat{x}_{k\|k}$ | A posteriori state estimate |
| $A$ | State transition matrix of the control system |
| $B$ | Input matrix describing how control actions affect the state |
| $C$ | Output/measurement matrix mapping states to measured data |
| $y_k$ | True system output |
| $\tilde{y}_k$ | Measured output including attack noise |
| $\hat{y}_k$ | Estimated output $C\hat{x}_{k\|k}$ |
| $u_k$ | Control input at time step $k$ |
| $r_k$ | Reference value or setpoint for a controller |
| $w_k$ | Process noise, typically assumed $w_k \sim \mathcal{N}(0, Q)$ |
| $v_k$ | Measurement noise, typically assumed $v_k \sim \mathcal{N}(0, R)$ |
| $Q$ | Process noise covariance |
| $R$ | Measurement noise covariance |
| $b_k$ | Measurement bias (systematic error added to the sensor output) |
| $d_k^a$ | Injected false data |
| $D_{\mathrm{FDI}}$ | Covariance of the injected false data distribution |
| $K_k$ | Kalman gain at time step $k$ |
| $P_{k\|k-1}$ | Predicted estimation error covariance |
| $P_{k\|k}$ | Updated estimation error covariance |
| $L$ | Luenberger observer gain matrix |
| $e_k$ | Estimation error, $e_k = x_k - \hat{x}_{k\|k}$ |

## Operators

| Operator | Description |
|---|---|
| $(\dots)^{-1}$ | Inverse |
| $\sum$ | Sum |
| $(\dots)^{\mathsf{T}}$ | Transpose |

## Abbreviations

| Abbreviation | Full name |
|---|---|
| CKF | Cubature Kalman Filter |
| CV | Constant Velocity |
| DoS | Denial-of-Service |
| EKF | Extended Kalman Filter |
| FDI | False Data Injection |
| ICS | Industrial Control System |
| IT | Information Technology |
| KF | Kalman Filter |
| OT | Operational Technology |
| RMSE | Root-Mean-Square Error |
| RSE | Root-Square Error |

# 1 Introduction

Attacks on industrial control systems (ICS) have become a focus of cyber-security measures. The reason being that they often rely on devices with old hardware, firmware, and software [1]. Anton *et al.* analyze and describe how most devices in a production environment are exposed to known vulnerabilities. These devices also use insecure protocols that open the possibility of attack chains [2]. For the aforementioned reasons, it is important to find ways to secure production environments.

It is important to apply common information technology (IT) defense techniques such as patching and monitoring, where applicable. However, in practice, isolation has become central in cyber-security in operational technology (OT). This mitigation technique starts with separating OT from IT. The idea behind this is to close off the production network from the outside world. OT networks also benefit from internal segmentation. In his article [3], Belev explains how this can be implemented for a ship when treated as an ICS. This architecture reduces cyber-attack risks and makes response easier. An example is that incidents are more narrowly confined, making both detection and response faster. One of the most common conduits (allowed connections) is for data, which often comes from sensors. Sensors are a key part of production and are often highly optimized in terms of their design.

The importance of data integrity in production varies, but for the most part, it is crucial. Controllers control actuators based on sensor readings. Reliance on data opens up an attack vector that information security measures or segmentation are not necessarily suited to handle. A false data injection (FDI) attack [4, 5] often occurs at the network level of sensors and actuators. These systems have less security built in [1], and the network protocols are sometimes not secure at all.

One way to ensure resilience when dealing with noisy or manipulated data is state estimation. State estimation means predicting the internal values of a system, *i.e.* estimating its state based on the previous state, current state, and the relationship between internal variables. W. Zheng *et al.* [6] provide a demonstration of state-estimation in a power system. The effect of bad data was partially eliminated by using a state-estimation model to provide the control values instead of relying solely on the data readings.

This thesis aims to investigate secure implementations of control systems. The goal is to investigate the threat landscape of FDI attacks and how state estimation could provide resilience against them. The second chapter investigates control systems, FDI attacks, and state estimation through an in-depth literature review. The third chapter motivates the choice of research and explains the objective in more detail. The fourth chapter explains the research process and methods. The fifth chapter discusses the results obtained from the research involved in this work. The final chapter summarizes the thesis and includes conclusions that have been and can be made from the research.

# 2 Literature review

A typical control system uses sensor data as input to some defined model. To obtain a correct model output, the system is modeled with enough variables to account for different factors affecting the system. A linear system can be modeled in discrete time as [7]

$$x_{k+1} = Ax_k + Bu_k, \quad (1)$$
$$y_k = Cx_k + D,$$

where $u_k$ is a variable based on the measurement. In a proportional closed feedback controller implementation, the response to the data would come from

$$u_k = K_p e_k, \quad (2)$$

$$e_k = r_k - y_k,$$

where $y_k$ is the measurement (output), $r_k$ is the desired value and $e_k$ the error between these. $K_p$ is a chosen proportional control parameter stabilizing the system appropriately. The control of the internal state $x$ is herein directly based on the data points of $y_k$. The goal of the system is to minimize the error $e_k$.

## 2.1 False data injection attacks

When implementing automatic data-based control, the trust by default is total. The problem that arises with the above-described system is that if the data is from a bad source or tampered with before it reaches a control system. In their article [8], Das *et al.* present a false data injection attack, where the system output is replaced with tampered data points. An attack could also be performed on sensor output, which in many cases is controller input [9, 10]. Converting from the continuous-time model to its discrete feedback equivalent and introducing the false data to the feedback controller system above gives

$$\bar{y}_k = y_k + F^a d_k, \quad (3)$$
$$x_{k+1} = Ax_k + BK_p(r_k - \bar{y}_k),$$
$$y_k = Cx_k + D,$$

where $y_k$ is the output of the system under normal conditions. The system now responds to $\bar{y}_k$, leading to the controller applying changes according to a completely different state than the actual $y_k$.

An FDIA could have systems compromised if targeted on sensor data or control signals sent to actuators [8, 9]. In their article [10], Fuxi Wen and Wei Liu argue that taking control of field devices, including sensors, is to be accounted for as the most likely entry of intrusion. In their example with power grids, the geographical spread of the area makes this the case, but this also applies to other industrial settings. The role of physical security is very important [11], and this is the result of access control not being well implemented in OT, at least compared to the IT standards of today. This physical access, combined with unsafe protocols, allows for man-in-the-middle attacks to occur in the lower levels [1, 11]. It is most likely for the access breach to occur at the industrial communication level on insecure channels. Industrial networks consisting of wireless sensors without proper security face the same risks on wireless channels [12].

There are a few different ways to inject data, and given different systems and tools for disrupting them, the possibilities are often limited. Based on previous research [4, 8, 9, 13, 14], a summary of different types of FDI attacks with a brief description.

**Table 1:** A table of common FDI attack variations.

| **Name** | **Description** | **Idea** |
|---|---|---|
| Noise injection | Introduces noisy data into a control system. | Disrupts operation by making measurements less accurate. |
| DoS (denial-of-service) attacks | Overflow signals, and potentially making data noisy or causing measurements to stop momentarily or totally. | Bringing the system to a halt with inaccurate or missing data points. |
| Bias injection | Introduces a constant bias and thereby makes signals wrong by the amount of bias introduced. | Causing the system to operate on biased data to introduce chaotic measurements or to put it in a desirable state for the attacker. |
| Stealthy attack | Often carefully crafted technique where the system internals can be taken into account. A sophisticated approach where the goal is to make subtle and undetectable injections. | Moving the system into some desired state for the attacker. |

Most attacks have some limitations when considering that industrial settings enforce safety. Different types of thresholds force a system to halt, slow down, or otherwise stop actuating when dangerous circumstances arise. Unfortunately, research shows that a malicious actor can use this to their advantage. In their research [14], Liu *et al.* explain how attacks can be stealthy, which means they can go undetected, also explained in Table 1. This can be achieved by finding ways of operating within the allowed thresholds of a system or trick the system that measurements are within the operating margin.

With the growth of Industry 4.0 [15] and the integration of IT and OT systems, computational resources available in production have become larger. This has given room for advanced data science concepts, previously too computationally expensive to become available in industry as well. A useful technology for control systems is machine learning [16]. Kumar *et al.* [13] prove machine learning to be a viable option for FDIA detection. In their article [12], a model is also trained for better accuracy when building an FDIA resilient control system. The downside of machine learning is non-determinism and its unclarity compared to predetermined mathematical models.

## 2.2 State estimation

The base state estimation method in modern control systems is the Luenberger observer [7]. It works by introducing a state-observer for x. Consequently, an estimation can be calculated for the output y as well. The Luenberger observer estimation in parallel to equation (1) for x, based on the previously introduced control system, gives

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + L(y_k - \hat{y}_k), \tag{4}$$

$$\hat{y}_k = C\hat{x}_k,$$

where $\hat{x}_k$ denotes the estimated state and $L$ is the observer gain matrix. The *innovation* $y_k - C\hat{x}_k$ is the difference between the estimated and measured value for the output. When adjusting the observer gain L, the innovations' influence on state correction changes. The error in this case becomes

$$e_{k+1} = (A - LC)e_k. \tag{5}$$

### 2.2.1 State estimation methods and resiliance against attacks

Control systems, especially critical ones with sensors, can benefit from an extra layer of security. Layered vulnerabilities allow malicious data to enter systems and be read as acceptable values if an attack is well-constructed. State estimation introduces state awareness in a system. Chan et. Al [17] reviews multiple ways in which a system can keep track of multiple system variables. Utilizing Kullback-Leibler divergence/distance (KLD) when analyzing current and historical data can reveal faults in the current measured state. State estimation builds on calculating what the next state of the system should be. Comparing this state with the actual one can reveal FDIAs if multiple consecutive deviations of large magnitude occur.

The observer system in equation (4) introduces new calculations for each iteration, but it is based on parameters already tracked for the system in equation (1). As introduced in [18], a deterministic system with noiseless data can use a Luenberger observer to improve system resilience in abnormal conditions. This is because the error $e_k$ should converge to 0 as $k \to \infty$ as this is the aim of the observer. In the case of false data being introduced, as in equation (3) $e_k$ starts to deviate. A threshold value $\tau$ can be used for $||e_k||$, where $||e_k|| > \tau$ would indicate anomalies in the measurement data.

### 2.2.2 Kalman filter family

The Luenberger, while a good tool with a persistent offset, does not handle noise well unless the noise is heavily skewed. The Kalman filter, originally introduced in [19] is smoothing and cancels out stochastic disturbance by using estimations. A Basic Kalman filter for equation (1) includes a prediction step and covariance calculation step

**Prediction step:**

$$\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1} + Bu_{k-1}, \tag{6}$$

**Covariance calculation:**

$$P_{k|k-1} = AP_{k-1|k-1}A^{\mathsf{T}} + Q_k. \tag{7}$$

$$S_k = CP_{k|k-1}C^{\mathsf{T}} + R_k.$$

where $Q_k$ is the covariance of $u_k$, $P_k$ the covariance of $x_k$ and $R_k$ that of $y_k$.

**Kalman gain:**

$$K_k = P_{k|k-1}C^{\mathsf{T}}S_k^{-1}. \tag{8}$$

**Update step:**

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - C\hat{x}_{k|k-1}), \tag{9}$$

$$P_{k|k} = (I - K_kC)P_{k|k-1}.$$

As seen in the update step of equation (9), the next estimate is a weighted average of the predicted state $\hat{x}_{k|k-1}$ and the latest output $y_k$. This effect cancels out noise, assuming a random distribution. This means that Kalman filtering theoretically works against FDIA where the signal is being disturbed or noise is otherwise introduced. This would not mean detection of the attack, but continued smooth operation of the system. In their article [9], K. Kumar *et al.* use the KF to provide system stability and inaction to false data during an FDI attack.

For non-linear systems, *i.e.,* most real-world systems, the extended Kalman filter (EKF) can be used instead of the above-described Kalman filter that works for linear data. The non-linearities in a system are estimated using a first-order Taylor series as one additional step. A nonlinear system is, unlike linear systems, modeled $x_{k+1} = f(x_k, u_k) + w_k$, where $w_k$ represents system noise. An for the output $y_k = h(x_k) + v_k$ where $v_k$ is measurement noise. Thus, the calculation of covariance and Kalman gain is changed to

**Jacobian calculation:**

$$F_k^x = \frac{\delta f(\hat{x}_{k|k-1})}{\delta x}, H_k^x = \frac{\delta h(\hat{x}_{k|k-1})}{\delta x}. \tag{10}$$

**Covariance calculation:**

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^{\mathsf{T}} + Q_k. \tag{11}$$

$$S_k = H_k^x P_{k|k-1} H_k^{x\mathsf{T}} + R_k.$$

**Kalman gain:**

$$K_k = P_{k|k-1} C^{\mathsf{T}} S_k^{-1}. \tag{12}$$

Another filter in the Kalman filter family is the cubature Kalman filter (CKF), first introduced by Arsasartnam *et al.* in [20]. Given $n = dim(x)$, the CKF uses **2n** cubature points to estimate the covariance matrix and Kalman gain. The cubature points are symmetric pairs formed by the mean and covariance of each state variable. Calculating a numeric integral and a weighted average of the points gives the next state estimate. The method, as discussed in the original article, provides an accurate and computationally efficient method for filtering non-linear systems.

Other variations also exist. In later work [12], Kumar *et al.* combine Kalman filtering with GLSR to further prove how cyber-attacks can be mitigated with an example scenario. The example scenario is one where a wireless channel is the link between a sensor and another device and is exploited to inject false data. By using state estimation, including a Kalman filter, they manage to prove theoretical resilience against this attack.

### 2.2.3 System resiliance against FDIA

Below is a table with different approaches for state-estimation. Each technique comes with an explanation of the methodology. Each row also includes which types of attacks this technique could defend against.

**Table 2:** State-estimation method versus FDIA attack types and sources.

| **Method(s)** | **Principle** | **Effective against (attack types)** | **Source(s)** |
|---|---|---|---|
| Luenberger Observer | Deterministic observer. Any persistent false data biases can be detected via a threshold. Low complexity and low cost. | Persistent attacks that assume no guardrails. | [17] |
| Luenberger Observer and Statistical test | Bias detected by statistically analyzing features of estimation, such as convergence and mean. | Sneaky attacks if they cause conflict in statistical tests. | [17] |
| Kalman Filter | Filters out stochastic data, providing resilience against noisy attacks. | Random data injections or data-loss (such as DoS attack related scenarios.) | [9, 12] |
| Extended Kalman Filter | Implements the KF after linearizing the non-linear systems using the first-order Taylor series approximation.* | Noisy data injection attacks. | [21] |
| Cubature Kalman Filter | Uses numeric integration to allow the benefits of the KF filter for non-linear systems. | Noisy data injection attacks. | [22] |

*Loses accuracy with large non-linearities.

# 3 Motivation and objective

This part of the thesis is a brief motivation of the chosen research, its relevance, how it is done and the process of choosing methods. It is clear that many aspects affect the safe and correct operation of industrial systems. Even with many methods available, it is important to have the ability to validate data going through the system. This is essential when operating based on sensor readings or other actuator input. Assuring this has significant challenges if the checks are not made on the control devices themselves. Validating data this way should be a priority since it adds an extra layer of security. Ignoring security risks introduces business risk, such as production downtime. Additionally, security risks can become physical risk when the systems operate in the real world.

The solution researched in this paper is state estimation as part of the system itself. State estimation is already implemented in a multitude of applications in industry and is therefore a very suitable choice, as it is an area with vast previous research and implementation examples. As discussed in the literature review, there are many different methods available. Each comes with its strengths as highlighted in table 2. Especially, the Kalman filter family has shown promise in establishing resiliency in FDI attack scenarios.

A scenario where both DoS attacks and sensor manipulation attacks could occur is an automatically guided system. It also introduces safety concerns since it operates in the physical world. It is also a system which can operate with a variety of different sensors, and a fair amount of new technology emerges in this field. If state estimation can be implemented for the steering system, it is a system where safe operation would be allowed even with emerging technologies in communication and sensing.

The objective of this research is to simulate a toy example situation where a control system is used, subject it to FDI attacks, and apply state estimation techniques to test and prove the basic principles in different scenarios. While previous research exists, much of it is focused on a particular type of system or device with an optimized solution for that particular system. The investigated system and state estimation methods implemented for it are aimed at highlighting common attacks and defensive techniques for each.

# 4 Research material and methods

This section explains the scenario chosen for the research. It also highlights the system implementation and key aspects of the work. Initially, the chosen example is explained, including simulation setup, state estimation, and attack methods. In this simulation, the measurements are given directly as global coordinates. The control system has linear internal dynamics.

## 4.1 System selection

For simulating attacks, an example system with a moving device utilizing two sensors for calculating its position in a two-dimensional fashion. It has transparent dynamics, is easy to reproduce, and is common in many OT environments.

## 4.2 System dynamics

In their paper [23], Baisa derives the constant velocity (CV) model for two-dimensional navigation. He describes how the position and velocity in two directions can be iteratively updated. The simulation follows the same physical model for describing the state of the system and its update steps. To include a reference signal, the one-dimensional case available about kinematic models in [24] can be applied to this two-dimensional case.

The system has a state vector

$$X_k = \{x_k, y_k, v_k^x, v_k^y\}^\mathsf{T} \tag{13}$$

where $(x_k, y_k)$ are the coordinates and $(v_k^x, v_k^y)$ the respective velocities.

The system is updated at each step k using

**System internal update:**

$$\begin{aligned} X_k &= AX_{k-1} + Bu_k \\ Y_k &= CX_k + v_k \end{aligned} \tag{14}$$

**State-space parameters:**

$$A = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} \frac{\Delta t^2}{2} & 0 \\ 0 & \frac{\Delta t^2}{2} \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{15}$$

where $v_k$ is $\mathcal{N}(0, Q_{k-1})$ and represents noise in the process and $u_k$ is the reference signal. In this research, Q is chosen to be random Gaussian noise with zero mean.

For making the system move in a desired way, a PD feedback controller is implemented.

$$u_k = K_P(r_k - Y_k) + K_D(\dot{r}_k - \dot{Y}_k) \tag{16}$$

where $r_k$ is a given position $\{x_k^{ref}, y_k^{ref}\}^\mathsf{T}$ that the system should move towards.

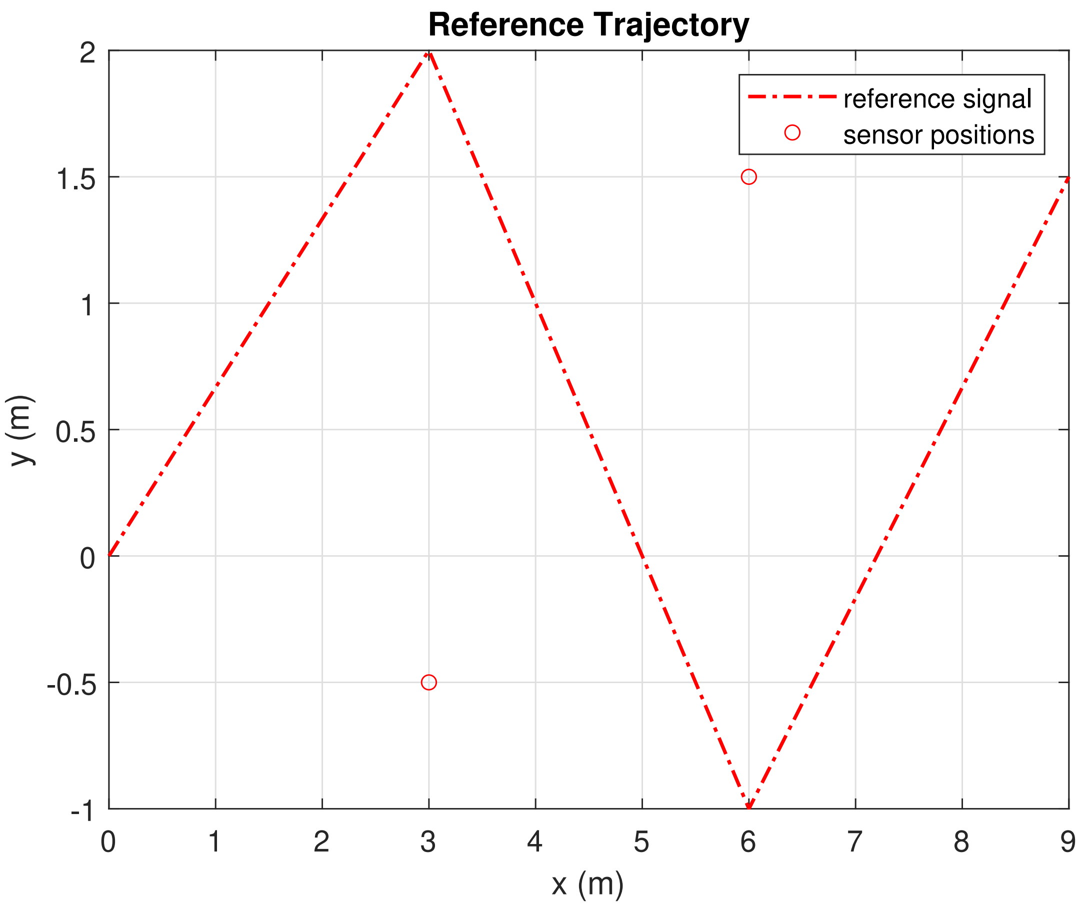


**Figure 1:** Reference signal for the system. Desired position in the 2D plane.

In the first scenario with global position measurements, the two sensors will be fused by simple averaging, meaning

$$x_k = \frac{(x_k^{(1)} + v_{x,k}^{(1)}) + (x_k^{(2)} + v_{x,k}^{(2)})}{2} \tag{17}$$

$$y_k = \frac{(y_k^{(1)} + v_{y,k}^{(1)}) + (y_k^{(2)} + v_{y,k}^{(2)})}{2}$$

where $(x_k^{(1)}, y_k^{(1)})$ are the measurements of the first sensor and $(x_k^{(2)}, y_k^{(2)})$ of the second. For each measurement, some measurement noise, $v_{x,k}$ and $v_{y,k}$, is introduced by the sensor.

To simulate the controller, the following parameters are used and disclosed for reproducibility. Stepsize $dt = 0.1s$, number of steps $N = 150$ and the controller proportional gain, $K_P = \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}$, and derivative gain, $K_D = \begin{bmatrix} 3.25 & 0 \\ 0 & 3.25 \end{bmatrix}$.

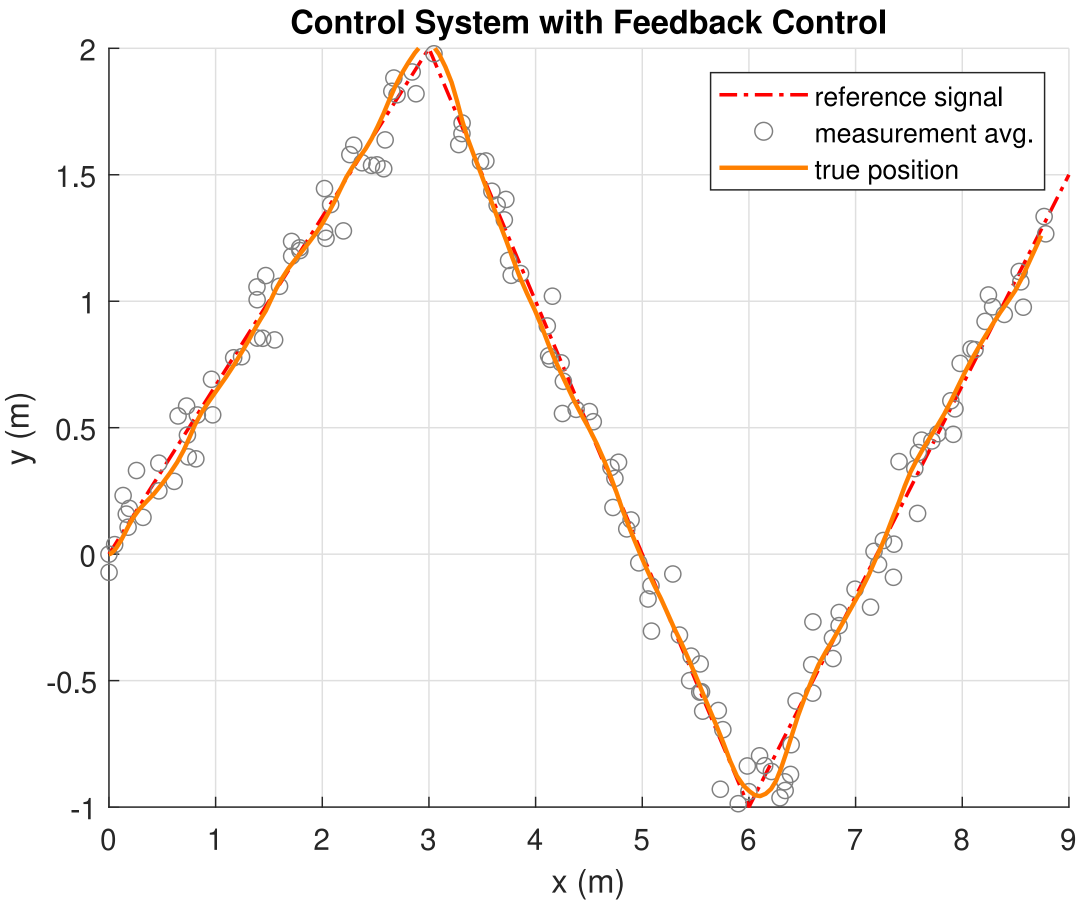


**Figure 2:** Simulation with sensor data and PD controller in the 2D plane.

The system is stable, and the state evolves with regard to the reference signal as represented in figure A1.

## 4.3 State estimation implementation

Although the system works well in the base case, there is noticeable disturbance from the noisy data. This is indicative of the fact that a noisy FDI attack could destabilize the system. To test the hypothesis of system resilience through state estimation, the promising and applicable methods from the literature review are implemented for the system.

The scenario selected in this thesis builds upon the previously outlined measurements. A Luenberger observer and the Kalman filter are implemented for the linear system with global position measurements. In this system, it is assumed that the sensor provides global position measurements to the controller, indicating that these positions are computed within the sensor itself. The calculation of the global coordinate is not regarded as relevant.

Implementing a Luenberger observer gives estimates

$$\hat{X}_{k+1} = A\hat{X}_k + Bu_k + L(Y_k - \hat{Y}_k), \tag{18}$$

$$\hat{Y}_k = C\hat{X}_k,$$

where $L = \begin{bmatrix} 0.11 & 0 \\ 0 & 0.11 \\ 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}$.

To extend research, a Kalman filter is implemented following equations (6)-(9). For the KF, the set noise parameters were

$$Q = 1 \times 10^{-5} \times I, \tag{19}$$

$$R = 0.1 \times 10^{-1} \times I.$$

In both cases, the state update is

$$u_k = K_P(r_k - \hat{Y}_k) + K_D(\dot{r}_k - \widehat{\dot{Y}}_k). \tag{20}$$

$$X_{k+1} = AX_k + Bu_k.$$

The results of both state-estimation implementations are seen in figure 3.

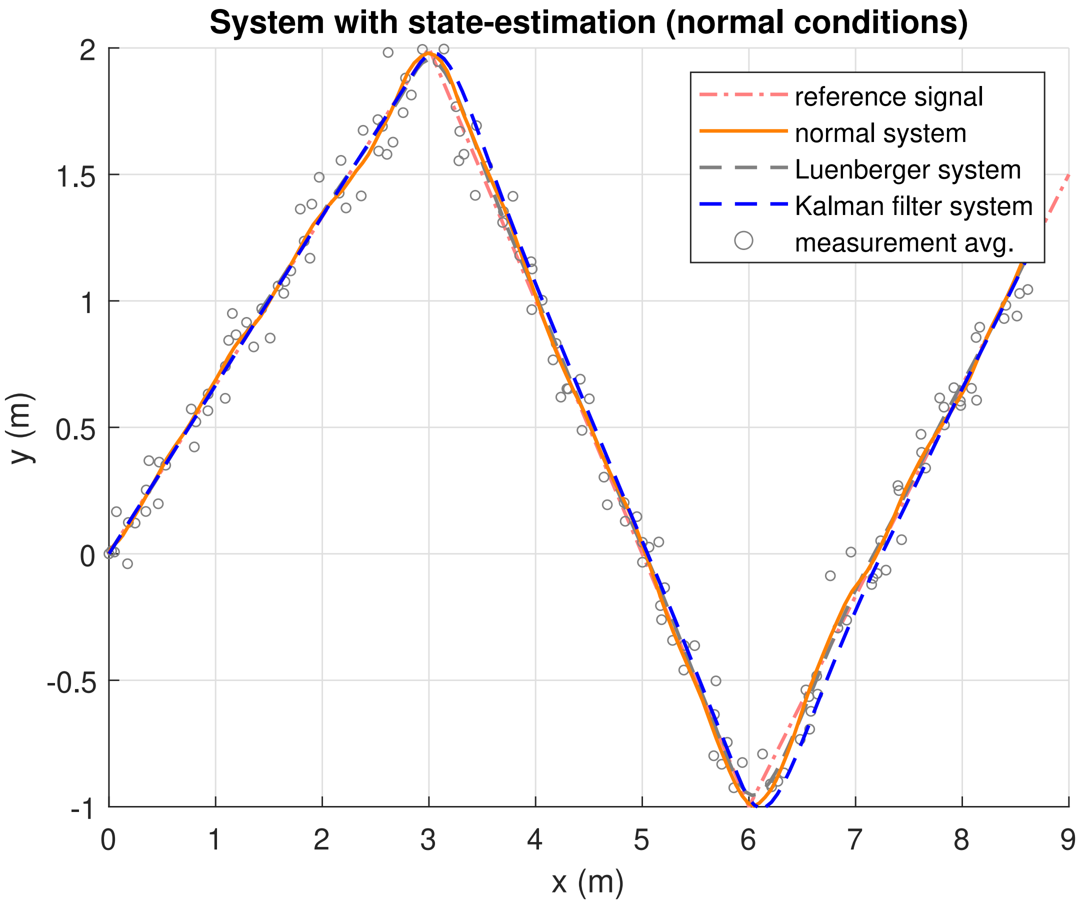


**Figure 3:** Trajectory using Kalman filtering (blue) and a Luenberger observer (grey.)

### 4.4 False data injection

To test the system, an FDI attack is simulated on it. The simulated injection introduces a factor $d_k^a$ in each update of the system by manipulating the data of one of the sensors directly. This aims to represent the situation of an actor manipulating the communication channel from one sensor to the controller. The measurement vector $\{x_k^{(1)}, y_k^{(1)}\}^\top$ is thereby replaced with

$$\{\overline{x}_k^{(1)}, \overline{y}_k^{(1)}\}^\top = \{x_k^{(1)}, y_k^{(1)}\}^\top + d_k^a \tag{21}$$

$$d_k^a = \{d_{x,k}^a, d_{y,k}^a\}^\top$$

To extend the research, two different variations of $d_k^a$ are introduced to the system. This allows simulating both a noisy injection and a bias injection. The chosen variations are

**Injected noisy data:**

$$d_{x,k}^a, d_{y,k}^a \in \mathcal{N}(0, D_{FDI}). \tag{22}$$

**Injected bias data:**

$$d_k^a = \{a, b\}^\top. \tag{23}$$

In both cases, the manipulation on the system level becomes

$$\overline{Y}_k = CX_k + \omega_k + \frac{d_k^a}{2} \tag{24}$$

Causing the trajectory to react to the noisy attack as in figure 4 and to the biased attack as in figure 5. The effect of introducing the same data when using state-estimation is discussed in the fifth section covering results.

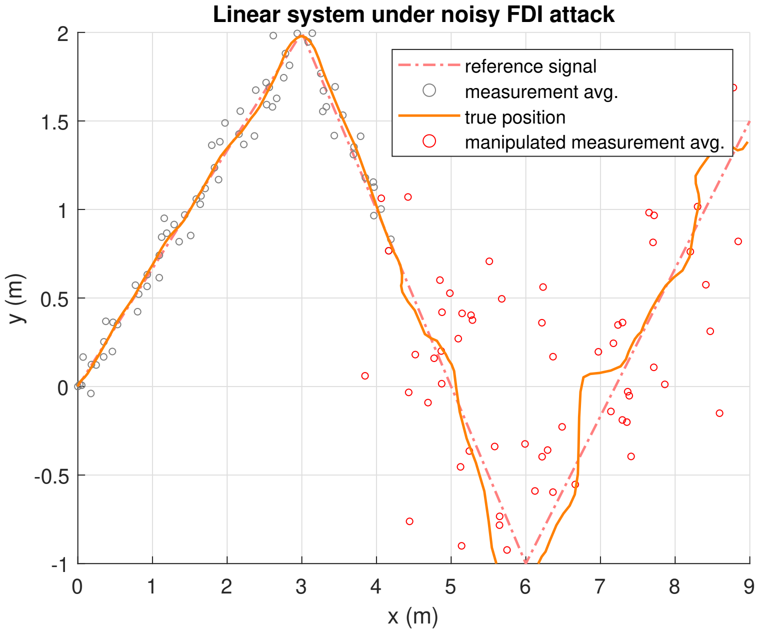


**Figure 4:** CV model reaction to noisy FDIA without state estimation.

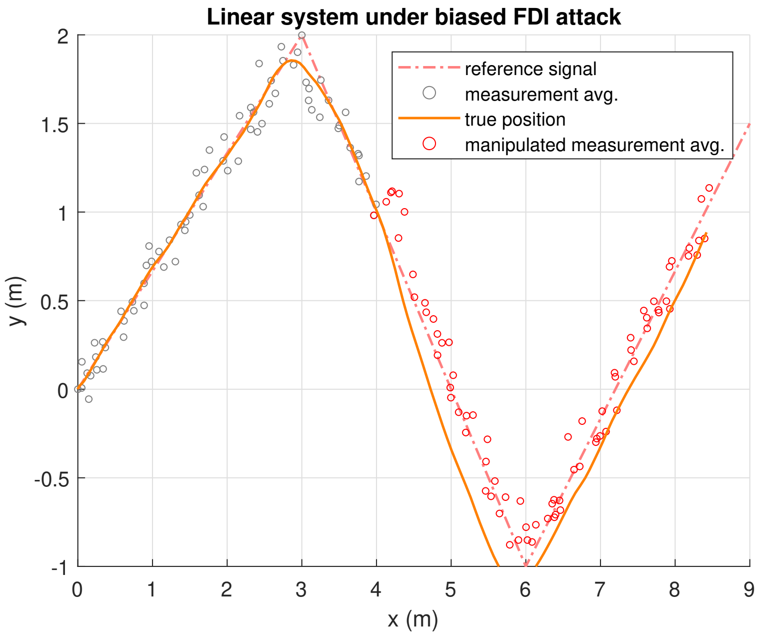


**Figure 5:** CV model reaction to noisy FDIA without state estimation.

# 5 Results

This fifth section analyzes the results of the simulations described in Section 4. Specifically, it analyzes how well the outlined goal is reached, *i.e.*, how well the two different state-estimation algorithms provide resilience to FDI attacks when implemented for a control system. The section has been divided into three subsections. The first two cover the results with both attack variations, and the third analyzes the used systems and methods, simulation results, motivated conclusions, and their validity.

As metrics, the root mean-squared error (RMSE) between reference and system state $\sqrt{(x_k - x_k^{ref})^2 + (y_k - y_k^{ref})^2}$ and the standard deviation (STD) of the error $\sigma_e = \sqrt{\frac{1}{N}\sum(e_k - \overline{e})}$ where $\overline{e}$ is the error mean.

## 5.1 System behaviour during noisy data injection

For the simulation of noisy data, the equation (22) is used with $D_{FDI} = I_2$ at halfway through the movement $k = 75$. Figure 6 visualizes the ability to stay on course with a large amount of suddenly injected noise.

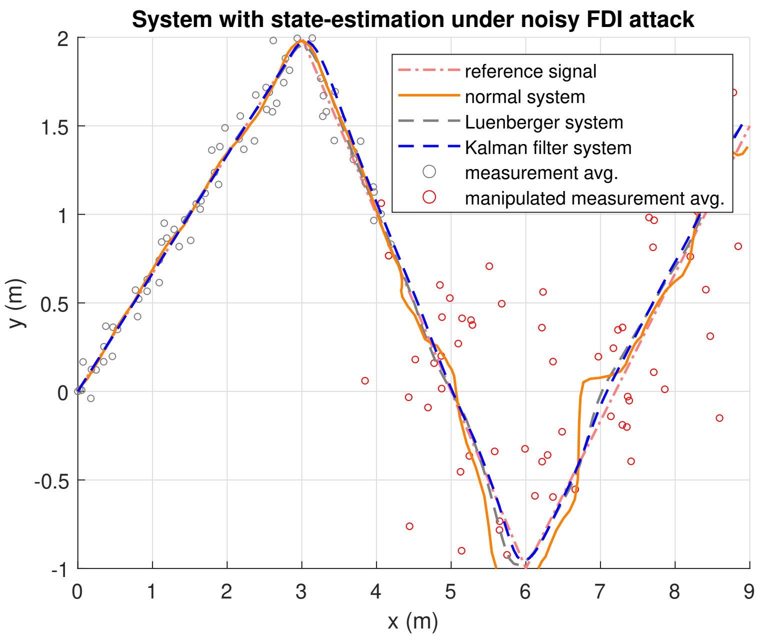


**Figure 6:** CV model reaction to noisy FDIA with state-estimation.

The tracked metrics have been documented for each system in table 3. The root-squared error RSE over time evolved as seen in figure A2.

**Table 3:** Noisy attack simulation metrics.

| System | RMSE in normal conditions | RMSE under attack | Error STD |
|---|---|---|---|
| Without state-estimation | 0.352 m | 0.404 m | +/- 0.217 m |
| With Luenberger observer | 0.348 m | 0.416 m | +/- 0.178 m |
| With Kalman filter | 0.275 m | 0.355 m | +/- 0.180 m |

## 5.2 System behaviour during biased data injection

As elaborated in equation (23), the other type of injection is a constant bias. This means that the mean of the measurements will be shifted by a along the x-axis and b along the y-axis. For the simulation, values $a = 0.2, b = 0.5$ are chosen, and the attack is also introduced here at $k = 75$. The trajectory for the different systems is available in figure 7.

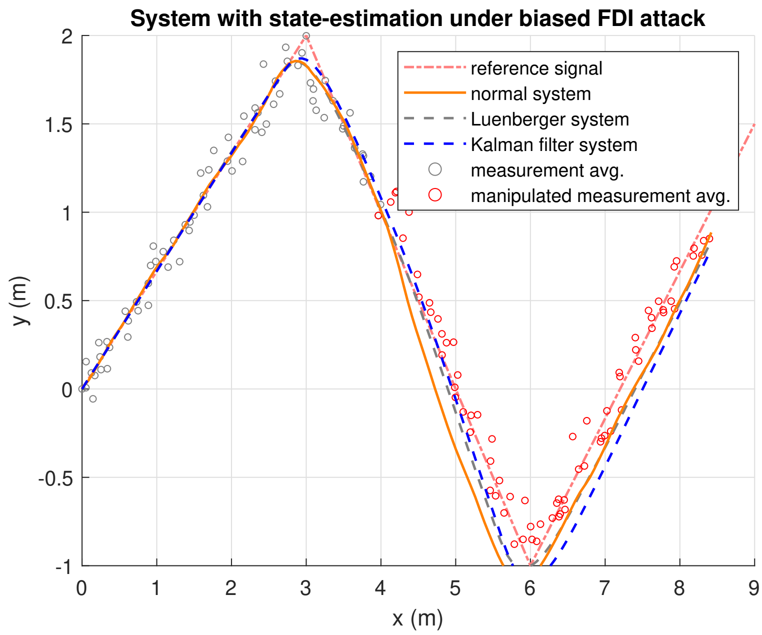


**Figure 7:** CV model reaction to biased FDIA with state-estimation.

RMSE prior to and during the attack and the STD of the error for the simulated bias injection attack are available in table 4. The RSE over time evolved as shown in A3.

**Table 4:** Biased attack simulation metrics.

| System | RMSE in normal conditions | RMSE under attack | Error STD under attack |
|---|---|---|---|
| Without state-estimation | 0.352 m | 0.557 m | +/- 0.121 m |
| With Luenberger observer | 0.348 m | 0.572 m | +/- 0.134 m |
| With Kalman filter | 0.275 m | 0.541 m | +/- 0.176 m |

## 5.3 Result analysis

In the measured trajectory, it is evident that state-estimation helps de-escalate when measurements start behaving chaotically or in a new way. This reaction can be described as stiffness in the system. This can be seen during the noisy FDI attack simulated above, where even high levels of noise do not confuse the system, since the noise is not skewed.

The trajectories seem to indicate that state-estimation produces a more accurate following of the reference path. This in itself is not conclusive evidence that state-estimation would provide more resilient systems. Reaction to noise can be assumed to be important in some systems, and the slow reaction to changes might be unacceptable. In the outlined scenario, the implemented state-estimation method's action is clearly safer, and it acts more predictably. The Luenberger observer proved to be as inaccurate as without it, but the lower STD indicates a more stable system. The KF implementation proved the hypothesis that a system with KF state-estimation would be more resilient against a noisy data injection attack with zero mean. The KF system showed better following of the trajectory, lower RMSE, and lower STD of error, indicating both better performance and stability.

As discussed in the literature review, the state-estimation methods had little ability to provide resilience against bias attacks directly. Analyzing the metrics for the systems under attack, it seems to align with those conclusions. Although state-estimation initially does not react as fast to the injected bias, the steady error grows with time. The slower reaction to bias would allow for increased reaction time, which could be argued to improve safety. That is, however, a case-specific benefit.

The choice of methods results in tackling the topic without unnecessary abstraction. Based on conclusions drawn in prior sections of the thesis, the results were in line with the other research and as expected. The simulations could benefit from varying sensor types, such as ones introducing non-linear but more accurate control mechanisms. This research focused on two separate attack scenarios for the same system. More conclusions might be drawn when using a similar attack scenario for a variety of systems with the same state-estimation techniques.

# 6 Summary/Conclusions

This thesis aimed at reviewing state-estimation methods as an extra layer of security for control systems. The threat vector for the thesis was narrowed to FDI attacks. The research, motivated largely by safety, aimed at studying whether or not control systems can use state-estimation to increase resilience against such attacks. The research initially covered existing research as a literature review in section two. Section three covered the main focus, scope, and methods of this thesis. Chapter four opened up how the researched methods could be applied in an example scenario and discussed ways in which system resilience was tested. Section five analyzed the obtained results from these tests and stated what conclusions could and could not be drawn. Additionally, the results section included suggestions for improvement.

This study shows that there is a place for state-estimation when implementing resilient systems. There is, however, further research required to find better coverage of system and attack scenario combinations. The Kalman filter provided good results against the outlined noisy FDI attack, highlighting that there is a place for such methods when designing attack-resistant systems. The lack of leverage against the biased injection highlighted their limitations and their case-specific restrictions. Better resilience and generalization could be proved by researching more systems and other state-estimation techniques, such as those covered in the literature review.

# References


[1] S. D. D. Anton et al. "The Global State of Security in Industrial Control Systems: An Empirical Analysis of Vulnerabilities Around the World". In: *IEEE Internet of Things Journal* 8.24 (Dec. 2021), pp. 17525–17540. DOI: 10.1109/JIOT.2021.3081741.

[2] M. Bhole, W. Kastner, and T. Sauter. "A Model Based Framework for Testing Safety and Security in Operational Technology Environments". In: *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*. Stuttgart, Germany, 2022, pp. 1–4. DOI: 10.1109/ETFA52439.2022.9921549.

[3] B. C. Belev. "Purdue Model Implementation in the Shipping Control Systems". In: *2022 10th International Scientific Conference on Computer Science (COMSCI)*. Sofia, Bulgaria, 2022, pp. 1–4. DOI: 10.1109/COMSCI55378.2022.9912594.

[4] D. Serpanos. "False Data Injection Attacks on Sensor Systems". In: *2022 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*. San Diego, CA, USA, 2022, pp. 1–5.

[5] D. Mukherjee. "Data-Driven False Data Injection Attack: A Low-Rank Approach". In: *IEEE Transactions on Smart Grid* 13.3 (May 2022), pp. 2479–2482. DOI: 10.1109/TSG.2022.3145633.

[6] W. Zheng et al. "A robust bilinear three-phase state estimation method for power systems". In: *2016 IEEE Power and Energy Society General Meeting (PESGM)*. Boston, MA, USA, 2016, pp. 1–4. DOI: 10.1109/PESGM.2016.7741284.

[7] R. C. Dorf and R. H. Bishop. *Modern control systems*. eng. Fourteenth edition / Richard C. Dorf, Robert H. Bishop. Harlow, United Kingdom: Pearson, 2022 - 2022. ISBN: 9781292422374.

[8] S. Das, A. Ghosh, and D. Chatterjee. "Detection of False Data Injection Attacks in Cyber-Physical Systems". In: *2024 IEEE International Symposium on Information Theory (ISIT)*. 2024, pp. 3249–3254. DOI: 10.1109/ISIT57864.2024.10619152.

[9] K. Kumar, M. Iqbal, and S. Särkkä. "Risk-Sensitive Filtering under False Data Injection Attacks". In: *2024 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. 2024, pp. 1–6.

[10] F. Wen and W. Liu. "An Efficient Data-Driven False Data Injection Attack in Smart Grids". In: *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*. 2018, pp. 1–5. DOI: 10.1109/ICDSP.2018.8631857.

[11] S. Adepu, G. Mishra, and A. Mathur. "Access Control in Water Distribution Networks: A Case Study". In: *2017 IEEE International Conference on Software Quality, Reliability and Security (QRS)*. 2017, pp. 184–191. DOI: 10.1109/QRS.2017.58.

[12] K. Kumar, M. Iqbal, and S. Särkkä. "Statistical Linear Regression Approach to Kalman Filtering and Smoothing under Cyber-Attacks". In: *arXiv preprint arXiv:2504.08404* (2025).

[13] A. Kumar et al. "Improving Detection of False Data Injection Attacks Using Machine Learning with Feature Selection and Oversampling". In: *Energies* 15.1 (2022). ISSN: 1996-1073. DOI: 10.3390/en15010212. URL: https://www.mdpi.com/1996-1073/15/1/212.

[14] Y. Liu, P. Ning, and M. K. Reiter. "False data injection attacks against state estimation in electric power grids". In: *Proceedings of the 16th ACM Conference on Computer and Communications Security*. CCS '09. Chicago, Illinois, USA: Association for Computing Machinery, 2009, pp. 21–32. ISBN: 9781605588940. DOI: 10.1145/1653662.1653666. URL: https://doi.org/10.1145/1653662.1653666.

[15] F. J. Folgado et al. "Review of Industry 4.0 from the Perspective of Automation and Supervision Systems: Definitions, Architectures and Recent Trends". In: *Electronics* 13.4 (2024). ISSN: 2079-9292. DOI: 10.3390/electronics13040782. URL: https://www.mdpi.com/2079-9292/13/4/782.

[16] B. Huang et al. "Modern Machine Learning Tools for Monitoring and Control of Industrial Processes: A Survey". In: May 2020.

[17] S. Chan and P. Nopphawan. "Bespoke Mitigation Framework for False Data Injection Attack-Induced Contingency Events". In: *2023 International Conference On Cyber Management And Engineering (CyMaEn)*. 2023, pp. 492–499. DOI: 10.1109/CyMaEn57228.2023.10050946.

[18] Z. Zhang et al. "On Feasibility of Coordinated Time-Delay and False Data Injection Attacks on Cyber–Physical Systems". In: *IEEE Internet of Things Journal* 9.11 (2022), pp. 8720–8736. DOI: 10.1109/JIOT.2021.3118065.

[19] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45. ISSN: 0021-9223. DOI: 10.1115/1.3662552. eprint: https://asmedigitalcollection.asme.org/fluidsengineering/article-pdf/82/1/35/5518977/35_1.pdf. URL: https://doi.org/10.1115/1.3662552.

[20] I. Arasaratnam and S. Haykin. "Cubature Kalman Filters". In: *IEEE Transactions on Automatic Control* 54.6 (2009), pp. 1254–1269. DOI: 10.1109/TAC.2009.2019800.

[21] S. Sathish and M. A. Aneesa Farhan. "Detection of False Data Injection in Cyber Physical Power Systems using Extended Kalman Filter". In: *2023 IEEE International Conference on Energy Technologies for Future Grids (ETFG)*. 2023, pp. 1–6. DOI: 10.1109/ETFG55873.2023.10408571.

[22] S. Chen et al. “A Class of Nonlinear Kalman Filters Under a Generalized Measurement Model With False Data Injection Attacks”. In: *IEEE Signal Processing Letters* 29 (2022), pp. 1187–1191. DOI: 10.1109/LSP.2022.3172623.

[23] N. L. Baisa. *Derivation of a Constant Velocity Motion Model for Visual Tracking*. 2020. arXiv: 2005.00844 [cs.CV]. URL: https://arxiv.org/abs/2005.00844.

[24] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan. *Estimation with applications to tracking and navigation*. eng. New York: Wiley, 2001. ISBN: 9786610541898.

[25] L. Hu et al. “State estimation under false data injection attacks: Security analysis and system protection”. In: *Automatica* 87 (2018), pp. 176–183.

[26] Z.-H. Pang et al. “False Data Injection Attacks Against Partial Sensor Measurements of Networked Control Systems”. In: *IEEE Transactions on Circuits and Systems II: Express Briefs* 69.1 (2022), pp. 149–153. DOI: 10.1109/TCSII.2021.3073724.

[27] M. Iranpour and M. R. Narimani. “Analyzing the Impact of AC False Data Injection Attacks on Power System Operation”. In: *2025 IEEE Texas Power and Energy Conference (TPEC)*. 2025, pp. 1–6. DOI: 10.1109/TPEC63981.2025.10906949.

[28] M. El-Hajj. “Preventing Data Integrity Breaches in IoT Applications Using Digital Twins”. In: *2025 23rd Mediterranean Communication and Computer Networking Conference (MedComNet)*. 2025, pp. 1–6. DOI: 10.1109/MedComNet65822.2025.11103570.

[29] K. Guibene, N. Messai, and M. Ayaida. “Securing Cyber-Physical Industrial Systems Against False Data Injection Attacks: A Hybrid Detection Approach”. In: *2023 Fifth International Conference on Blockchain Computing and Applications (BCCA)*. 2023, pp. 142–149. DOI: 10.1109/BCCA58897.2023.10338925.

[30] F. E. Tosun et al. “Kullback-Leibler Divergence-Based Observer Design Against Sensor Bias Injection Attacks in Single-Output Systems”. In: *IEEE Transactions on Information Forensics and Security* 20 (2025), pp. 2763–2777. DOI: 10.1109/TIFS.2025.3546167.

# A Appendix

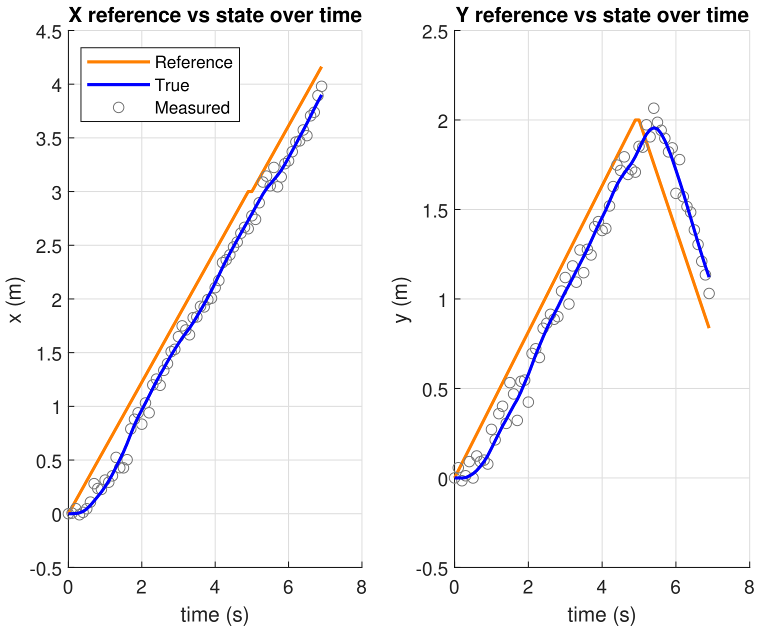


**Figure A1:** State variables vs reference signal.

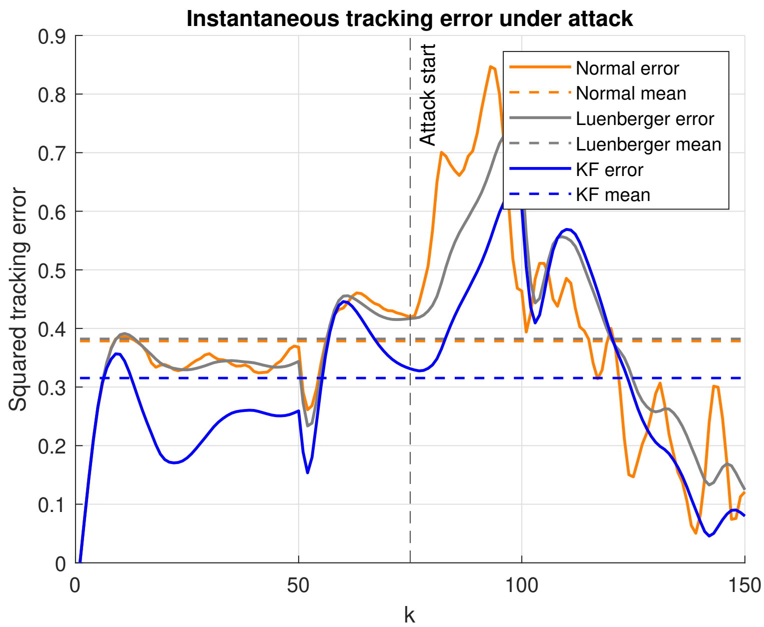


**Figure A2:** Root-squared error during noisy attack over time (k.)

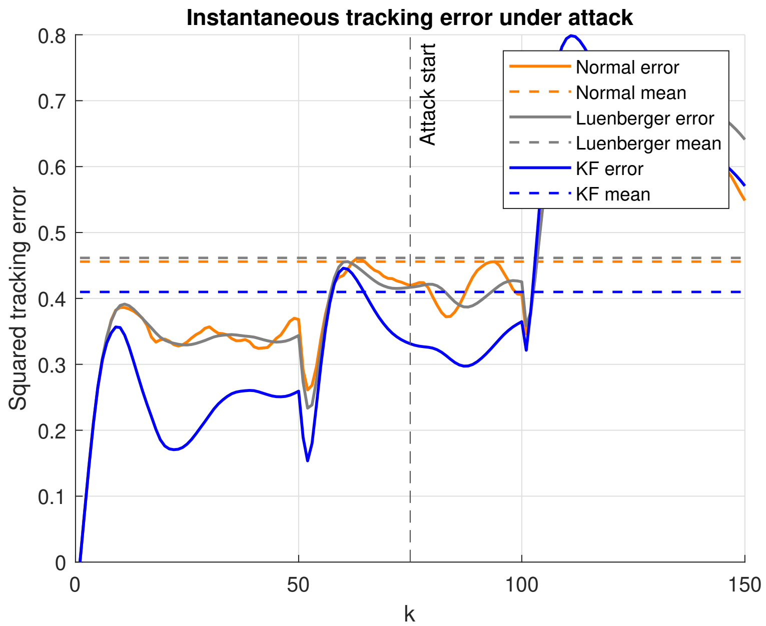


**Figure A3:** Root-squared error during biased attack over time (k.)