Chapter 6

ENGAGING EMPIRICAL DYNAMIC MODELING TO DETECT INTRUSIONS IN CYBER-PHYSICAL SYSTEMS

David Crow, Scott Graham, Brett Borghetti and Patrick Sweeney

Abstract Modern cyber-physical systems require effective intrusion detection systems to ensure adequate critical infrastructure protection. Developing an intrusion detection capability requires an understanding of the behavior of a cyber-physical system and causality of its components. Such an understanding enables the characterization of normal behavior and the identification and reporting of anomalous behavior.

This chapter explores a relatively new time series analysis technique, empirical dynamic modeling, that can contribute to system understanding. Specifically, it examines if the technique can adequately describe causality in cyber-physical systems and provides insights into it serving as a foundation for intrusion detection.

Keywords: Intrusion detection systems, empirical dynamic modeling

1. Introduction

Intrusion detection systems are commonly used to defend against cyber-physical system attacks and protect critical infrastructure assets. These systems monitor computer systems and networks, and report malicious activity to system administrators. In the cyber-physical system domain, an intrusion detection system can identify attempts by attackers to modify or misrepresent physical processes.

Consider an automobile as an example of a cyber-physical system. If an attacker intends to have the driver receive a speeding ticket, the attacker could inject packets that specify a lower speed to cause the speedometer to display incorrect information. In this case, an effective intrusion detection system would notice that the speed data does not conform to the expected behavior indicated by the related physical pro-

Published by Springer Nature Switzerland AG 2020
J. Staggs and S. Shenoi (Eds.): Critical Infrastructure Protection XIV, IFIP AICT 596, pp. 111–133, 2020.
https://doi.org/10.1007/978-3-030-62840-6_6

cess data such as engine and wheel rotational velocities, throttle position and fuel efficiency. In other words, the intrusion detection system would notice that the speedometer readings are anomalous.

In another example, if an intrusion detection system knows that a substantial increase in the automobile's brake pressure likely precedes a relative decrease in velocity, it could assert that no change, a small change or an increase in velocity (after the application of significant brake pressure) are anomalous. Of course, this requires the intrusion detection system to determine the expected behavior and identify anomalies.

Designing an effective intrusion detection system for a vulnerable cyber-physical system requires insights into the system dynamics or patterns, including how the current system state enables predictions about future states. An adequate quantity of data obtained under normal operating conditions is required to establish normal behavior. A process is needed to determine whether new traffic conforms to normal behavior. Also, an alerting system is necessary to report abnormal traffic behavior.

Intrusion detection architects need a strong understanding of cyber-physical systems or powerful computational resources to model system dynamics. However, the latter is often infeasible because many cyber-physical systems have limited hardware or are constrained by standards and regulations. Modern automobiles, for example, utilize small network packets and fairly simple hardware. For this reason, the former is often more attainable. A solid understanding of system dynamics – specifically, causality, such as how one signal affects another and how a current state predicts future states – is required to identify anomalous traffic, assuming that an ample quantity of normal data is available.

This research examines empirical dynamic modeling, an emerging technique that supports sophisticated time series analyses. Empirical dynamic modeling can contribute to the understanding of cyber-physical systems, and this research evaluates the feasibility of using the technique to detect intrusions in cyber-physical systems.

Two datasets are employed in the evaluation. The first is based on a simple linear model of the relationship between the steering wheel of an automobile and the relative velocities of its two front wheels. The second dataset was generated using a nonlinear flight simulator called the avionics vulnerability and assessment system (AVAS).

2. Background

This section discusses cyber-physical systems and time series, along with empirical dynamic modeling, an emerging technique for nonlinear forecasting and causality analysis.

2.1 Cyber-Physical Systems and Time Series

The journal *ACM Transactions on Cyber-Physical Systems* [1] defines cyber-physical systems as:

> "... systems where the cyber parts, i.e., the computing and communications parts, and the physical parts are tightly integrated, both at the design time and during operation. Such systems use computations and communications deeply embedded in and interacting with physical processes to add new capabilities to physical systems ... There is an emerging consensus that new methodologies and tools need to be developed to support cyber-physical systems."

Analyses of cyber-physical systems require high-fidelity models. However, the models are often difficult to articulate and replicate. For this reason, it is necessary to analyze the inputs and outputs of a cyber-physical system to develop a model of the system.

Often, the output of a cyber-physical system is time series data that expresses the values of its processes over time. An example time series in the case of an aircraft is the instantaneous revolutions per minute (rpm) of the propeller over time as measured by the aircraft sensors. The National Institute of Standards and Technology (NIST) [9] observes that "[t]ime series analysis accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for."

Kotu and Deshpande [6] differentiate between time series analysis and time series forecasting. Time series analysis involves the extraction of meaningful non-trivial information and patterns from time series. Time series forecasting involves the prediction of future time series data based on past observations and other inputs.

Most time series analysis and forecasting techniques require data stationarity for the time series in question. A stationary process has the property that the mean, variance and autocorrelation do not change over time; the time series data is flat without trends, and has constant variance over time, constant autocorrelation over time and no periodic fluctuations (seasonality) [9].

Figure 1 presents examples of time series plots [5]. Figure 1(a) shows Google stock prices over 200 consecutive days. Figure 1(b) shows the annual numbers of labor strikes in the United States. Figure 1(c) shows the annual prices of a dozen eggs in the United States (constant dollars). Figure 1(d) shows the monthly totals of pigs slaughtered in Victoria, Australia. Figure 1(e), which represents a stationary time series, shows the annual totals of lynx trapped in the McKenzie River District of Northwestern Canada. Figure 1(f) shows the monthly electricity

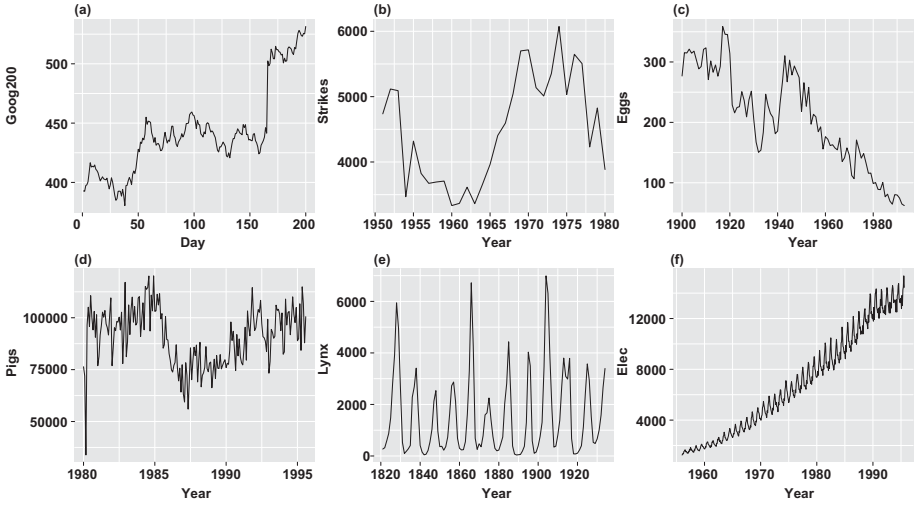

Figure 1. Example time series plots [5].

production in Australia. These plots demonstrate the diversity of time series models.

Time series data generated by the cyber-physical systems of an aircraft are non-stationary, so analysis and forecasting techniques that require stationarity are not viable. However, empirical dynamic modeling supports non-stationary time series analyses and forecasting.

2.2 Empirical Dynamic Modeling

Takens [16] introduced the delay embedding theorem in 1981. The theorem deals with mathematical attractors, where an attractor is the value or set of values that a system settles towards over time. Empirical dynamic modeling is an application of the delay embedding theorem. Sugihara et al. [15] state that empirical dynamic modeling "is based on the mathematical theory of reconstructing system attractors from time series data." In practice, empirical dynamic modeling is used to capture nonlinear dynamical systems with observational time series data.

Figure 2 provides visual representations of the main ideas underlying Taken's delay embedding theorem and empirical dynamic modeling [15, 19]. Figure 2(a) shows a Lorenz attractor, a set of solutions to a Lorenz dynamical system that is modeled as a set of ordinary differential equations [8]. The attractor manifold M is the set of states that the system progresses through time. The figure shows that a time series for a given dimension can be identified by recording observations in the dimension over time.

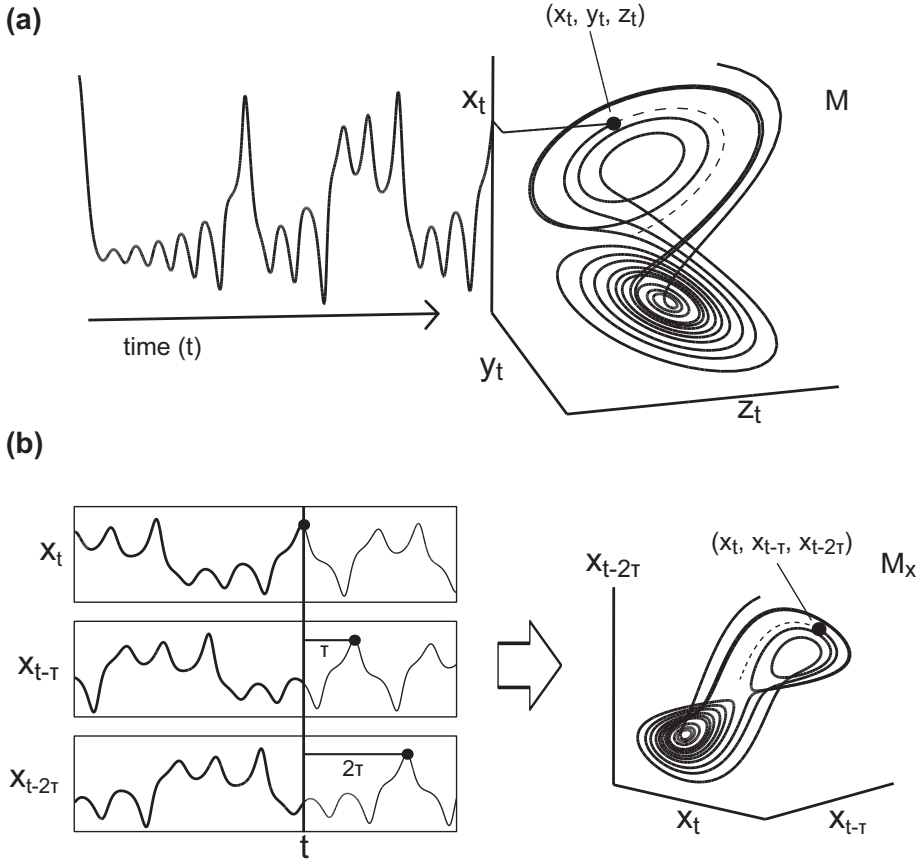

Figure 2. Lorenz attractor and shadow manifold [15, 19].

Figure 2(b) shows how a univariate time series can be converted to a higher dimensional representation using time-lagged versions of itself as additional dimensions [15, 19]. Projecting the system states from M to the coordinate axis X generates a time series. The time series is not the manifold, but it is used to create it using lags. The figure shows how lags of the time series X are used as coordinate axes to construct the shadow manifold M_X. The visual similarity between M_X and M is apparent. Takens [16] showed that the shadow manifold M_X is diffeomorphic (maps one-to-one) to its original attractor manifold M.

Sugihara et al. [14] have also shown that the diffeomorphic property between M and its shadow manifolds – one for each dimension – implies that the shadow manifolds are diffeomorphic with respect to each other; the opposite is also true. Thus, if two shadow manifolds are diffeomorphic with respect to each other, it can be assumed that they belong

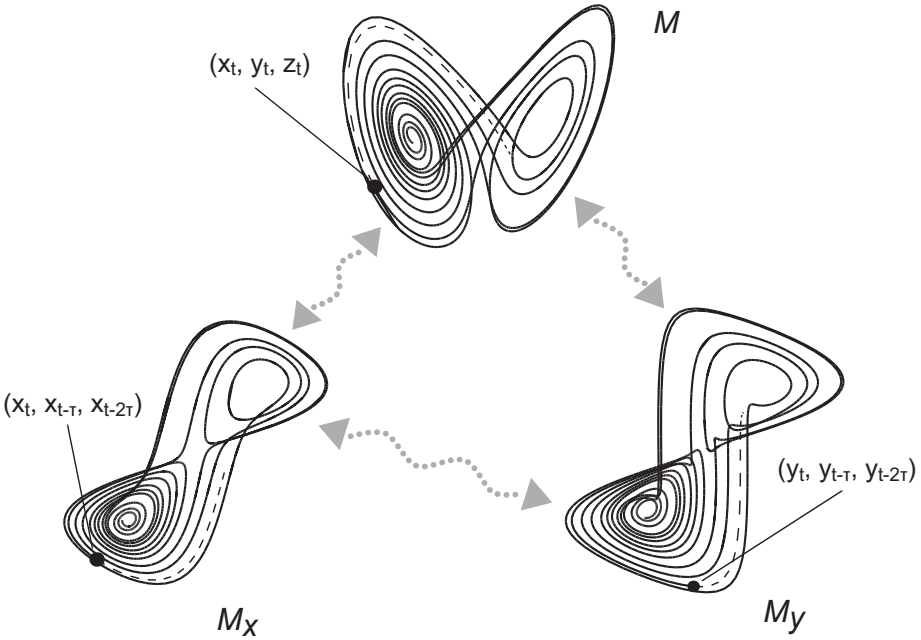

Figure 3. Convergent cross-mapping.

to the same dynamical system. Convergent cross-mapping (CCM), a mathematical technique developed by Sugihara et al. [14], can then be used to identify and quantify the causality between the two original time series. In short, convergent cross-mapping seeks to determine whether an arbitrary point and its nearest neighbors in one shadow manifold can accurately predict a point and its neighbors in another shadow manifold.

Figure 3 summarizes these concepts. Convergent cross-mapping tests the correspondence between shadow manifolds. The figure shows the attractor manifold of the original Lorenz system in three-dimensional space and two shadow manifolds, M_X and M_Y, constructed via lagged coordinate embeddings of X and Y, respectively. The arrows between the manifolds represent the diffeomorphic properties of the attractors. Because X and Y are dynamically coupled, nearby points in M_X correspond temporally to nearby points in M_Y. This enables the estimation of states across manifolds using Y to estimate the state of X, and vice versa, using the nearest neighbors. In the case of longer time series, the shadow manifolds become denser and the neighborhoods (ellipses of the nearest neighbors) shrink, allowing more precise cross-mapping estimates [14, 19]. Sugihara et al. [14] have shown that increasing the sample sizes of shadow manifolds improves the predictive power of con-

vergent cross-mapping and that the predictive power converges to some maximum as the sample size increases to infinity.

2.3 Related Work

A survey of the literature reveals that no published research has focused on applying empirical dynamic modeling to automobile or aircraft time series data, or even cyber security problems in general. The vast majority of applications are in the areas of economics and natural sciences. For example, the Sugihara Lab [15], where empirical dynamic modeling originated, primarily applies the technique to problems in ecology. This research explores the application of empirical dynamic modeling to intrusion detection in cyber-physical systems.

3. Proposed Methodology

The proposed methodology employs empirical dynamic modeling, a relatively new statistical analysis tool, to obtain insights into the characteristics of cyber-physical systems, including their nonlinearity, deterministic chaos and causality. This section discusses the nature and origins of the experimental data. Additionally, it describes the techniques used to analyze the data.

3.1 Datasets

This research has employed datasets generated from two simulated cyber-physical systems, an automobile and an aircraft. The two datasets are employed to evaluate empirical dynamic modeling techniques on linear and nonlinear cyber-physical systems.

The first dataset is based on simple relationships between the steering wheel of an automobile and the relative velocities of its two front wheels. The steering dataset is considered to be linear because the relationships between the pairs of time series are linear or nearly linear. Specifically, the relationship between the velocities of the two front wheels is linear and the relationships between the steering input and the velocity of each of the two wheels are almost linear. The latter two relationships are linear for steering wheel angles of small magnitude, but grow in nonlinearity as the steering wheel angle increases.

The second dataset was generated using the AVAS nonlinear flight simulator that employs real-world physics and flight dynamics for research purposes. Since the focus was on airplane airspeed, altitude and pitch, simulated data pertaining to airspeed, angle of attack, position, heading and wind angle was collected. The dataset is nonlinear because the relationships between the time series are nonlinear.

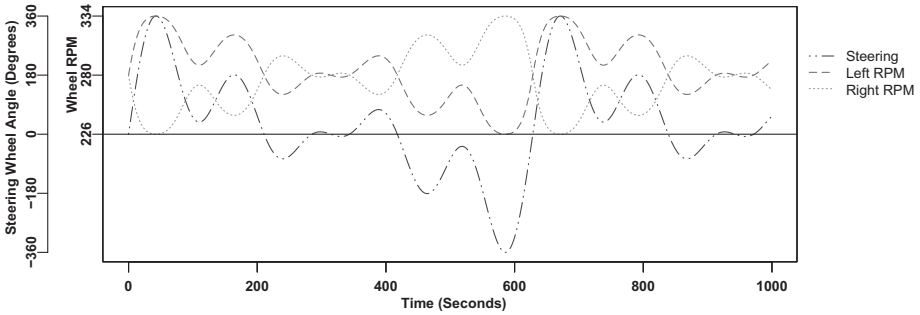

Figure 4. Plots of the steering system time series.

Linear Data. Linear data corresponding to the steering wheel angle and front wheel velocities of a passenger vehicle was employed to assess empirical dynamic modeling techniques. The time series were generated for a vehicle with 30-inch wheel radius including the tires, 72-inch wheelbase, 60-inch track, maximum steering wheel turning angle of 360°, steering ratio of 8:1 (maximum wheel angle of 45°) and constant forward speed of 25 mph.

Under the assumptions, a sum of sines function loosely represents a hypothetical driving scenario. That is, the steering line in Figure 4 serves as a potential steering wheel angle time series, and the inside and outside wheel velocities (in rpm) are computed using the following equations:

$$rpm_{inside} = \frac{60sb}{\pi r\left(b + \cos\left(90 - \frac{\theta}{t}\right)\sqrt{b^2 + \left(k + b\tan\left(90 - \frac{\theta}{t}\right)\right)^2}\right)} \quad (1)$$

$$rpm_{outside} = \frac{60s\sqrt{b^2 + \left(k + b\tan\left(90 - \frac{\theta}{t}\right)\right)^2}}{\pi r\left(b\sec\left(90 - \frac{\theta}{t}\right) + \sqrt{b^2 + \left(k + b\tan\left(90 - \frac{\theta}{t}\right)\right)^2}\right)} \quad (2)$$

Note that, when the steering wheel angle θ is less than zero, the left wheel corresponds to the inside wheel; otherwise, the right wheel corresponds to the inside wheel.

Table 1 describes the variables used in Equations (1) and (2) and specifies their values.

The automobile steering system model is rather rudimentary. It does not account for the physical properties of the real system and the effects of other relevant variables. Despite its simplicity, it is possible to draw

Table 1. Automobile steering system variables.

Variable	Description	Defined Value
r	Wheel radius	15 in
b	Wheelbase	12 in
k	Track	60 in
t	Steering ratio	8:1
s	Forward speed	25 mph
θ	Current steering wheel angle	NA

conclusions about the applicability of empirical dynamic modeling to linear cyber-physical systems.

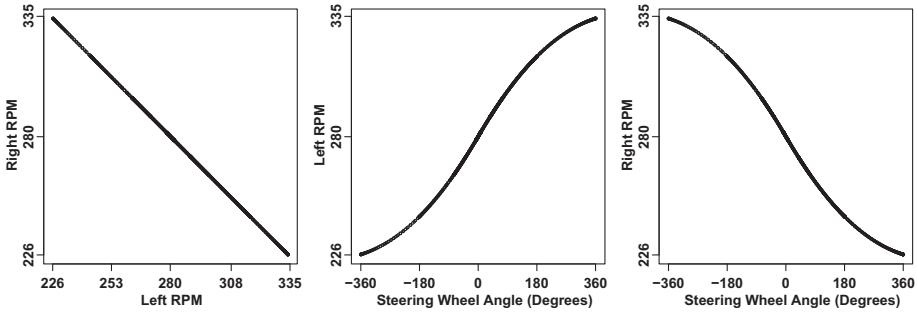

Figure 5. Scatter plots of the relationships between steering system variables.

Figure 5 shows the scatter plots of the relationships between the steering system variables, confirming that the time series are fairly linearly related. The somewhat nonlinear behavior between the steering wheel and each front wheel is due to the mechanics of the standard Ackermann automobile steering mechanism. The values of the variables cover significantly different ranges. For this reason, all the variables were standardized using the `scale` function in the R programming language to ensure that each variable would be equally important in the analysis. Given a time series as input, the `scale` function z-scales it by subtracting its mean and dividing by its standard deviation.

Nonlinear Data. The second dataset was created by guiding an AVAS-simulated aircraft through takeoff, low-altitude cruising and multiple shallow banked turns. The data collection yielded 7,582 observations from a 14-minute flight. Each observation included eight flight metrics with a timestamp relative to the start of the simulation. The metrics

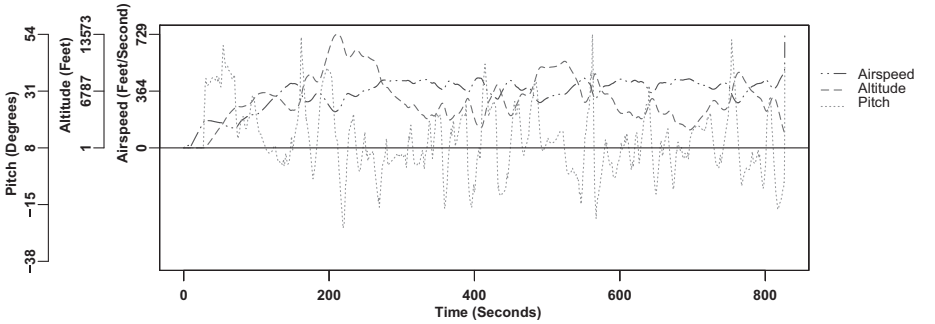

Figure 6. Plots of the selected AVAS time series.

included roll (deg) and pitch (deg), altitude (ft) and airspeed, vertical velocity and velocity along each of the three coordinate axes (ft/s). The roll and pitch values ranged from $-180°$ to $180°$. Altitude, airspeed and the directional velocities were floating point values (airspeed and altitude were nonnegative values). Note that yaw was excluded because, in the simulator, it is simply a measurement of aircraft heading relative to north, not a characteristic of aircraft dynamics.

The variables in the nonlinear dataset were also z-scaled using the `scale` function in R. A subset of variables – airspeed, altitude and pitch – were selected before conducting the analysis. Other subsets of the eight variables likely exhibit the desired dynamics, but the three selected variables were expected to best demonstrate a tightly-coupled system.

Figure 6 presents the three time series prior to scaling. Figure 7 shows the scatter plots of the relationships between each pair of AVAS variables, clearly demonstrating that the system is highly nonlinear.

3.2 Empirical Dynamic Modeling Techniques

Ye et al. [19] suggest that the following empirical dynamic modeling techniques be applied in sequence to best interpret the characteristics of a dataset:

1. Conduct nearest neighbor forecasting via simplex projection to identify the embedding dimension E that maximizes the forecast skill ρ [13].

2. Use simplex projection and E to determine whether the system exhibits deterministic chaos.

3. Employ sequential locally-weighted global linear maps (S-maps) to characterize the nonlinearity of the data [12].

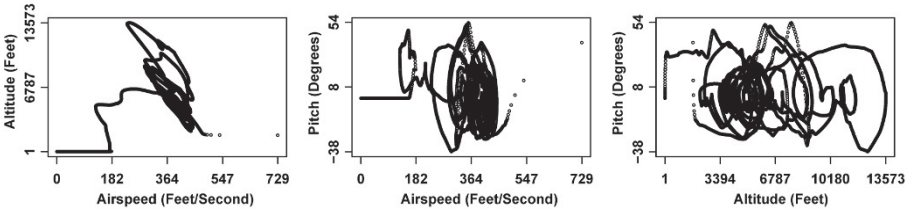

Figure 7. Scatter plots of the relationships between each pair of AVAS variables.

4. Use convergent cross-mapping to generate shadow manifolds, evaluate predictive accuracy and quantify causality [14].

A simplex is a generalization of a triangle or tetrahedron to an arbitrary number of dimensions. Simplex projection iteratively selects a point Y_t in a shadow manifold and b other points whose histories over time t are most similar to the selected point [7, 11, 13]. The weighted averages of the future values of the b other points are used to make predictions about future values of Y_t. The differences between these predictions and the actual future values give the forecast skill ρ. By repeating this process with shadow manifolds of different dimensionalities, the embedding dimension E that optimizes ρ is determined [3].

The (strong) Whitney embedding theorem states [17]:

Theorem 1. Any m-manifold of class C^r ($r \geq 1$ finite or infinite) may be embedded by a regular C^r-map in E^{2m} and by such a map in a one-one manner in E^{2m+1}.

Stated simply, the embedding dimension E of an attractor manifold has an upper bound of $2D + 1$ where D is the true dimension (number of variables) of the system [3, 11]. Thus, simplex projection can be used to definitively identify the optimal E in a finite amount of time.

S-map projection is also an iterative process, but it uses all the neighboring points to create linear regression vectors. Aggregating the regression vectors yields an approximation of an n-dimensional spline. This spline is compared against the shadow manifold attractor to obtain ρ [3, 7, 12]. When generating the regression estimates, a nonlinear tuning parameter θ is used to weigh the neighbors with respect to their distances to the current focal point Y_t. Finally, the time series is determined to belong to a simple linear system if ρ is maximized when $\theta = 0$; otherwise, it is a nonlinear system [3, 11, 12].

Stone et al. [11] claim that this process provides insights into the true dimensionality of the system that generates the observational data

without having a complete understanding of the system itself. Accurate knowledge of E is a prerequisite to effectively applying convergent cross-mapping to multiple time series to detect causality. Alternatively, a proper S-map analysis of time series relationships may indicate whether the relationships correspond to a simple linear system. If so, computationally simpler methods such as Granger causality or autoregressive linear models could replace the more complex convergent cross-mapping technique in order to detect causality [4, 12, 19]. Finally, knowledge of the dimensionality of a system may assist in creating a high quality model of the system. Such a model – and the results of causality analysis – could enable the development of an effective intrusion detection capability for a cyber-physical system.

The analysis was conducted using the rEDM repository on GitHub [18, 19]. The codebase enables empirical dynamic modeling analysis using the R programming language. It includes the following functions (among others):

- Function `simplex`, which corresponds to the first and second empirical dynamic modeling techniques.
- Function `s_map`, which corresponds to the third empirical dynamic modeling technique.
- Functions `ccm` and `ccm_means`, which correspond to the fourth empirical dynamic modeling technique.

These functions, along with some helper functions, facilitate effective empirical dynamic modeling analyses.

Interested readers are referred to [10] for a detailed presentation of empirical dynamic modeling, including the mathematics underlying simplex projection, S-map analysis and convergent cross-mapping.

4. Analysis Results

This section presents the results of the empirical dynamic modeling analyses of the linear and nonlinear datasets.

4.1 Linear Data

Knowledge of the optimal embedding dimension E for each time series of a system is required to effectively apply convergent cross-mapping to make predictions and quantify causality. The optimal value is identified by iteratively utilizing simplex projection to quantify the predictive accuracy at different values of E.

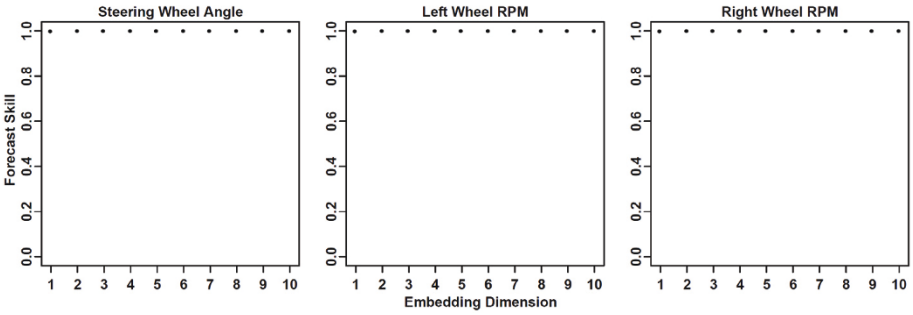

Figure 8. Optimal embedding dimensions for the steering system time series.

Figure 8 illustrates the results of applying this process to each steering system time series. The plots show that the forecast skill (ability to forecast future values of a time series) is maximized when E is greater than one. The value of E is fixed at two for the empirical dynamic modeling techniques in this section because a lower dimensionality reduces complexity and processing time. To be clear, setting $E = 2$ means that the techniques construct a two-dimensional shadow manifold where each dimension is a time series lagged by some multiple of τ. When predicting the steering wheel angle, for example, the technique constructs a shadow manifold using the steering wheel angle and one copy of the steering wheel angle where the copy is lagged by τ. The lag τ is assumed to be equal to one second in the analysis.

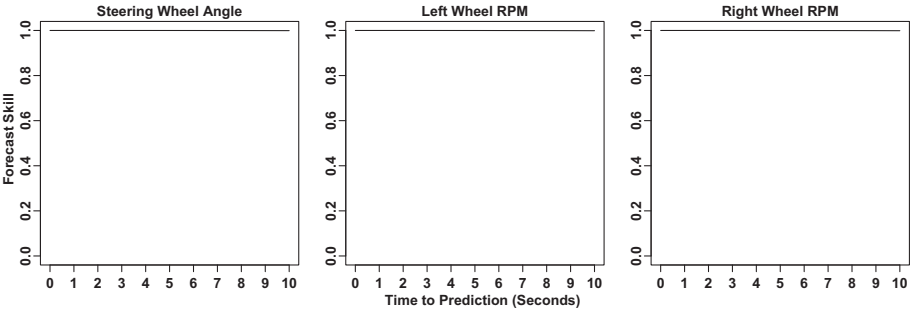

Figure 9. Deterministic chaos present in each steering system time series.

If E is kept constant and the time to prediction tp is varied, simplex projection enables an analysis of the deterministic chaos of the system. Figure 9 shows the deterministic chaos present in each steering system time series. Specifically, it shows how ρ decreases as tp increases for

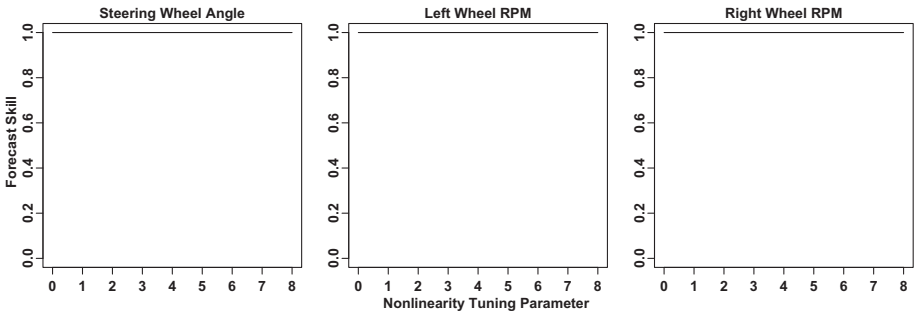

Figure 10. Nonlinearity of each steering system time series.

each of the three time series. In other words, predictions further in the future are much worse than those closer in time, which indicates chaotic behavior of the three variables. This is due to the nature of driving – without knowledge of the route taken by an automobile, it is practically impossible to predict the steering wheel angle. The simulated data conforms to this interpretation of driving behavior. However, the difference in the values of ρ at $tp = 0$ and $tp = 10$ is only about 0.0008; thus, the chaotic behavior in the system is minuscule. Empirical dynamic modeling does not support a deeper analysis of the chaos of the system.

S-map analysis fits local linear maps to a system to describe its nonlinearity. This is different from simplex projection, which analyzes the nearest neighbors of each point. The plots in Figure 10 were obtained by varying the nonlinearity tuning parameter θ in the S-map function call and plotting the value of ρ. When $\theta = 0$, S-map equally weights all the points – as θ increases, the function places more weight on points close to the point under analysis. Thus, when θ is higher, the function assumes that the system has more nonlinearity. For all three time series, ρ is the greatest when θ is high, which indicates the presence of nonlinearity in each time series; however, the trends are miniscule. Indeed, it appears that nonlinearity analysis using empirical dynamic modeling is not particularly useful for a linear system.

Empirical dynamic modeling also enables next-point predictions. Figure 11 overlays the predictions on each time series. Clearly, the predictions are extremely accurate, which indicates that the three variables do not change significantly from one observation to the next. Each plot also shows the prediction variance using a shaded polygon, but the variances are so low that the polygons are all but invisible. In fact, the plots in Figure 9 have already implied this – when tp is small, ρ is very high.

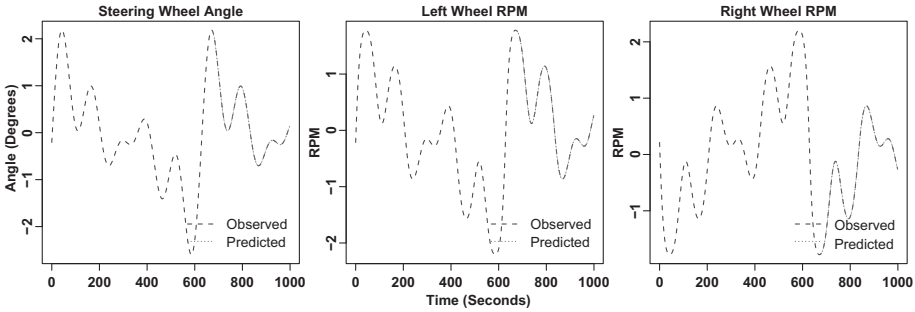

Figure 11. Next-point predictions for each steering system time series.

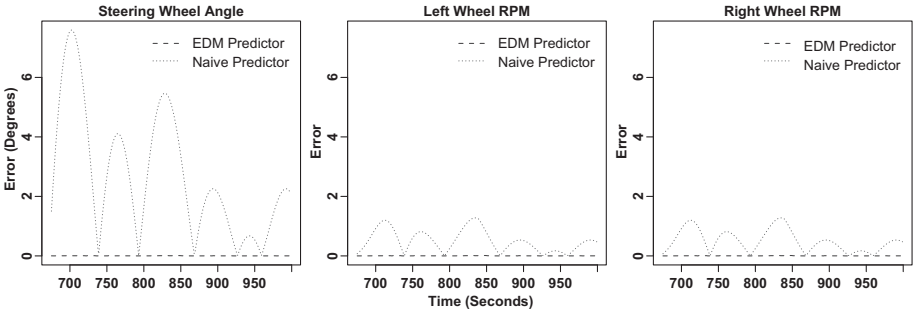

Figure 12. Next-point prediction errors for each steering system time series.

Additionally, a naive prediction model was created that simply predicts that a point at time $t + 1$ has the same value as the point at time t. In other words, the model predicts no change in the next value. Figure 12 shows the next-point prediction errors (residuals) for the naive model and empirical dynamic model. The majority of the errors are small, especially for the empirical dynamic model.

Table 2. Root-mean-square errors for each steering system time series.

Time Series	**Naive Model**	**Empirical Dynamic Model**
Steering wheel angle	0.009424	0.003351
Left wheel rotational velocity	0.005742	0.003893
Right wheel rotational velocity	0.005742	0.003893

Table 2 compares the root-mean-square errors for the naive and empirical dynamic models. Note that the root-mean-square error was used

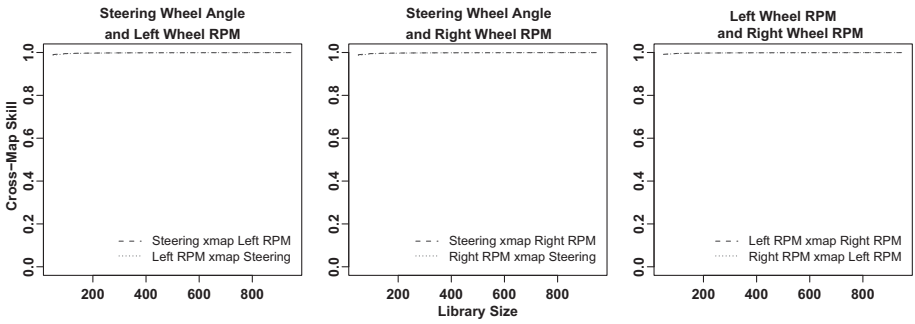

Figure 13. Causality between each pair of steering system time series.

in order to penalize large mispredictions heavily because such errors strongly affect intrusion detection system performance. As the table illustrates, the empirical dynamic model outperforms the baseline predictor for each time series. The time series are incapable of large, instantaneous changes, so accurately predicting the next point is not impressive nor it is very useful in practical applications. However, it could still assist in designing intrusion detection systems with low complexity. Of course, methods other than empirical dynamic modeling would also suffice for linear systems.

Figure 13 shows the inter-variable dynamics in the automobile steering system. Specifically, it plots the cross-mapping skill ρ against the library size (number of points) used to compute ρ for each pair of variables. The cross-mapping skill quantifies the ability to use one shadow manifold to identify values in another. Each plot has two lines, one for X xmap Y and one for Y xmap X. X xmap Y refers to the convergent cross-mapping analysis technique, which uses the shadow manifold of X to forecast the shadow manifold of Y. The value of ρ obtained for a given library size indicates this predictive capability. The three plots show that ρ is equivalent across library sizes and in both directions for every pair of time series. This means that the steering angle data is encoded in the wheel velocity data and the wheel velocity data is similarly encoded in the steering angle data, which in turn imply an expected causal effect in both directions. Unfortunately, it appears that empirical dynamic modeling does not provide insights about pairwise causality for this linear dataset.

Figure 14 shows the system causality over time. The plots again show the results of using X to forecast Y, but ρ is plotted against tp. According to Ye [18], negative values of tp indicate that past values of Y are best cross-mapped from the reconstructed state of X. Ye also

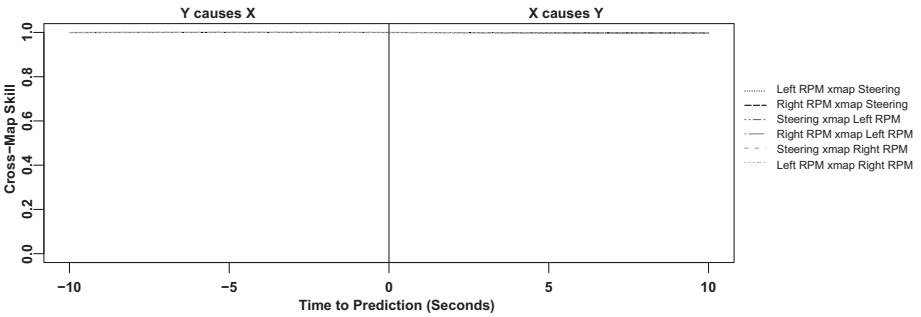

Figure 14. Causality predictions for selected pairs of steering system time series.

suggests that a signal appearing first in Y and later in X is consistent with Y causing X. The opposite is true when tp is positive. In the case of the automobile steering system, regardless of tp and the variables in question, ρ is approximately equal to one. Thus, according to empirical dynamic modeling, each variable has a strong causal effect on every other variable regardless of the time to prediction. This is unlikely and it bolsters the claim that empirical dynamic modeling does not support sophisticated analyses of linear system causality.

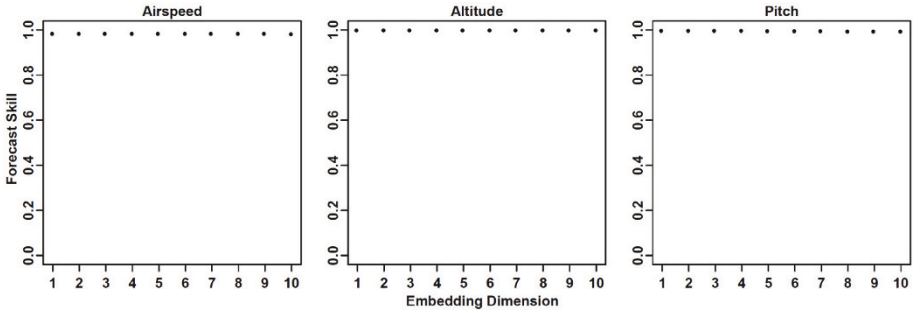

Figure 15. Optimal embedding dimensions for selected AVAS time series.

4.2 Nonlinear Data

Figure 15 presents the forecast skill ρ for various embedding dimensions E for three AVAS time series. The visual differences in ρ are minuscule, but the optimal embedding dimension was two for each series. Thus, $E = 2$ was used in the empirical dynamic modeling techniques in this section. Additionally, the time interval τ between two observations in a given time series was set to one second.

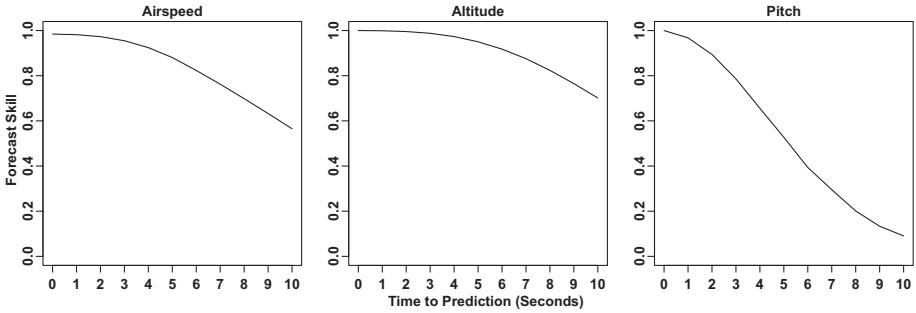

Figure 16. Deterministic chaos present in each selected AVAS time series.

Figure 16 plots the forecast skill ρ against the time to prediction tp to illustrate the deterministic chaos in the system. For each time series, predictions further in the future are much less accurate than earlier predictions. The effects are strongest for pitch and weakest for altitude. Regardless, this is evidence of chaotic behavior for all three variables.

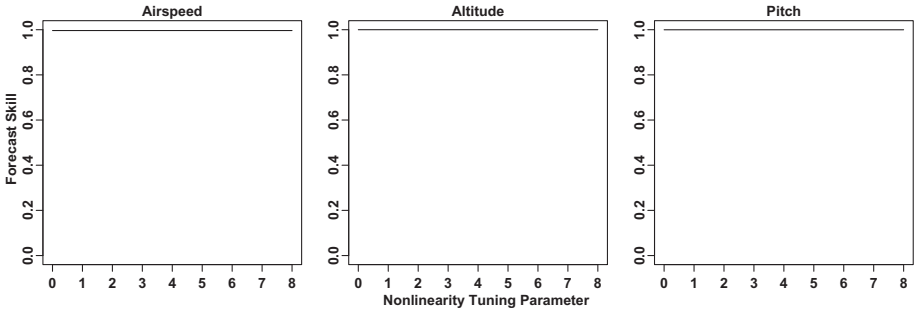

Figure 17. Nonlinearity of each selected AVAS time series.

Figure 17 plots ρ against θ to characterize the nonlinearity of each variable. In the case of airspeed and pitch, ρ is greatest when the function assumes the most nonlinearity; this is indicative of nonlinear dynamics. In the case of altitude, the S-map analysis implies the absence of nonlinear dynamics in the time series, but it is important to note that the change in ρ in all three plots is extremely small regardless of θ. For this reason, it is not possible to definitively claim the presence or absence of nonlinear dynamics.

Figure 18 presents the next-point predictions of empirical dynamic modeling for each time series. Unsurprisingly, the variances of the predictions – as exemplified by the nearly imperceptible shaded polygons –

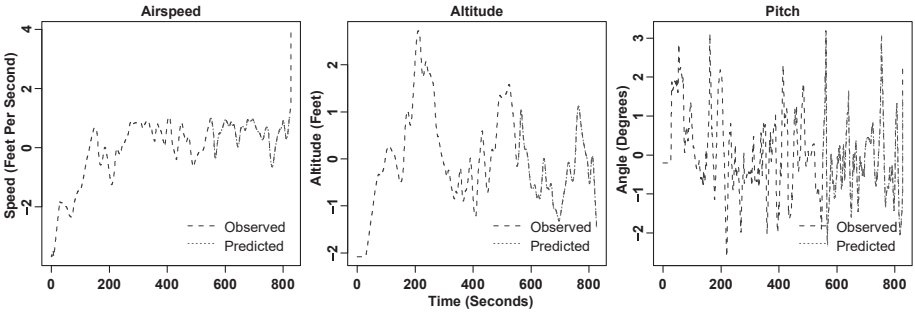

Figure 18. Next-point predictions for each selected AVAS time series.

are small. As Figure 16 strongly indicated, none of the variables change significantly between a pair of observations.

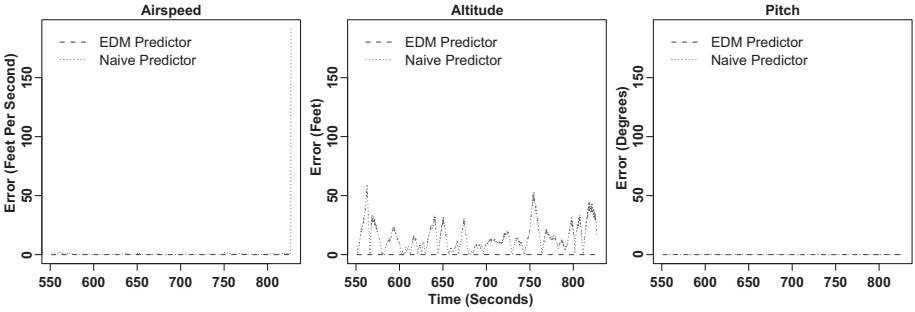

Figure 19. Next-point prediction errors for each selected AVAS time series.

Figure 19 shows the next-point prediction errors (residuals) for the naive and empirical dynamic models. The results confirm that the empirical dynamic model is highly accurate and once again outperforms the naive model.

Table 3. Root-mean-square errors for each AVAS time series.

Time Series	**Naive Model**	**Empirical Dynamic Model**
Airspeed	3.907005	0.070161
Altitude	18.192201	0.001535
Pitch	0.013749	0.019748

Table 3 compares the root-mean-square errors for the naive and empirical dynamic models. The empirical dynamic model vastly outper-

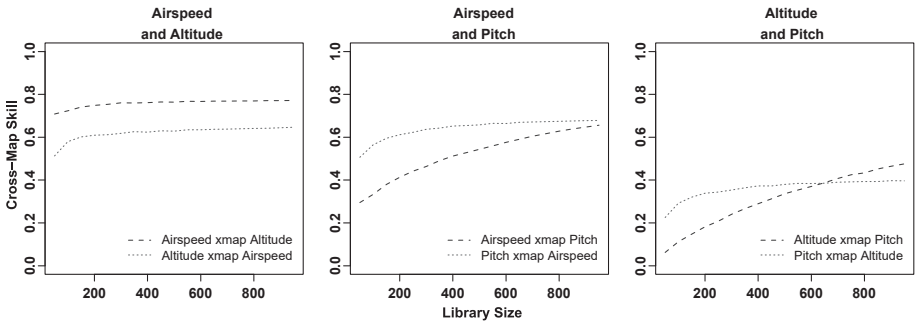

Figure 20. Causality between selected pairs of AVAS time series.

forms the naive model for two of the three time series. Although the naive model is a better predictor of the remaining pitch variable, the difference in root-mean-square errors is insignificant. Once again, it is possible that even these short-term predictions could assist in intrusion detection. However, the empirical dynamic model has a clear limitation – it cannot foresee values that are not in the library. This explains the large outlier predictions.

Figure 20 shows the cross-mapping skill for each pair of time series. The leftmost plot shows that the airspeed manifold can effectively forecast altitude but the opposite is noticeably weaker. The middle plot shows that the difference in cross-mapping skills between airspeed xmap pitch and pitch xmap airspeed decreases as the library size increases. The rightmost plot shows a more extreme case – above a certain library size, an inversion occurs in the difference in cross-mapping skills. In all three cases, the results indicate diminishing returns when attempting to improve ρ by increasing the library size. However, it is still possible that the analysis can improve intrusion detection system design.

Finally, Figure 21 plots the cross-mapping skill against time to prediction. Consider, for example, airspeed xmap pitch. When tp is slightly less than zero, ρ is maximized. This implies that airspeed best predicts pitch when lagged by about one second. In other words, pitch strongly affects airspeed after about one second. This behavior is expected. When tp is positive, ρ quickly decreases, and it can be asserted that airspeed does not have a strong causal effect on pitch. This is consistent with the standard interpretation of airplane mechanics.

5. Conclusions

The study of empirical dynamic modeling demonstrates that it can quantify the behavior of linear systems, but the results are limited and

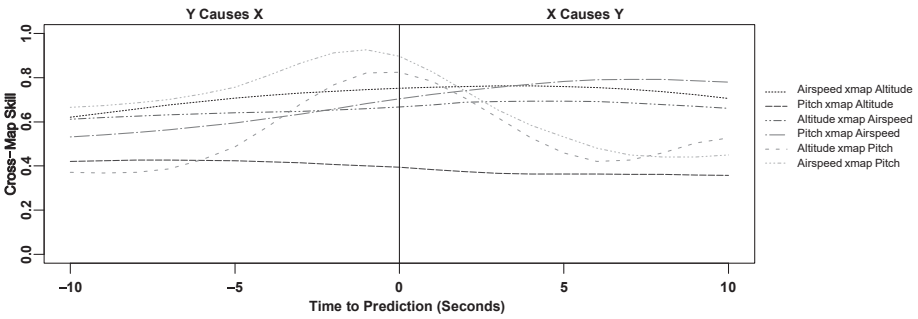

Figure 21. Causality predictions for selected pairs of AVAS time series.

may not assist in developing intrusion detection systems. In contrast, empirical dynamic modeling provides easy-to-use techniques that yield detailed insights about the behavior of nonlinear systems, which could advance intrusion detection efforts.

While empirical dynamic modeling may not be well suited to linear systems, it is important to note that cyber-physical systems are often highly nonlinear. Moreover, the linear system considered in this work is not fully representative of a real-world linear system. For this reason, future research should verify the applicability of empirical dynamic modeling to robust linear systems.

The nonlinear system analysis provided by empirical dynamic modeling is clearly useful for intrusion system design. In particular, it effectively quantifies causality in nonlinear systems. However, realizing the true potential of empirical dynamic modeling requires analyses of more realistic and complex datasets covering a variety of cyber-physical systems. It is hoped that this research will stimulate further investigations into the applicability of empirical dynamic modeling to intrusion detection and cyber security problems in general.

The views expressed in this chapter are those of the authors, and do not reflect the official policy or position of the U.S. Air Force, U.S. Department of Defense or U.S. Government. This document has been approved for public release, distribution unlimited (Case #88ABW-2020-049).

References

[1] Association for Computing Machinery, ACM Transactions on Cyber-Physical Systems, New York (`tcps.acm.org/about.cfm`), 2020.

[2] G. Boeing, Visual analysis of nonlinear dynamical systems: Chaos, fractals, self-similarity and the limits of prediction, *Systems*, vol. 4(4), article no. 37, 2016.

[3] C. Chang, M. Ushio and C. Hsieh, Empirical dynamic modeling for beginners, *Ecological Research*, vol. 32(6), pp. 785–796, 2017.

[4] C. Granger, Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, vol. 37(3), pp. 424–438, 1969.

[5] R. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, OTexts, Melbourne, Australia, 2018.

[6] V. Kotu and B. Deshpande, *Data Science: Concepts and Practice*, Morgan Kaufmann, Cambridge, Massachusetts, 2019.

[7] J. Lee, *Introduction to Topological Manifolds*, Springer-Verlag, New York, 2011.

[8] E. Lorenz, Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences*, vol. 20(2), pp. 130–141, 1963.

[9] National Institute of Standards and Technology, Introduction to time series analysis, in *NIST/SEMATECH e-Handbook of Statistical Methods*, Gaithersburg, Maryland (`www.itl.nist.gov/div898/handbook/pmc/section4/pmc4.htm`), 2012.

[10] N. Rennie, Empirical Dynamic Models: A Method for Detecting Causality in Complex Deterministic Systems (`docplayer.net/156079632-Empirical-dynamic-models-a-method-for-detecting-causality-in-complex-deterministic-systems.html`), 2018.

[11] B. Stone, Enabling Auditing and Intrusion Detection of Proprietary Controller Area Networks, Ph.D. Dissertation, Department of Computer Science, Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, 2018.

[12] G. Sugihara, Nonlinear forecasting for the classification of natural time series, *Philosophical Transactions of the Royal Society of London, Series A: Physical and Engineering Sciences*, vol. 348(1688), pp. 477–495, 1994.

[13] G. Sugihara and R. May, Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, *Nature*, vol. 344(6268), pp. 734–741, 1990.

[14] G. Sugihara, R. May, H. Ye, C. Hsieh, E. Deyle, M. Fogarty and S. Munch, Detecting causality in complex ecosystems, *Science*, vol. 338(6106), pp. 496–500, 2012.

[15] Sugihara Lab, Empirical Dynamic Modeling, Scripps Institution of Oceanography, University of California at San Diego, La Jolla, California (deepecoweb.ucsd.edu/nonlinear-dynamics-research/edm), 2020.

[16] F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence*, D. Rand and L. Young (Eds.), Springer, Berlin Heidelberg, Germany, pp. 366–381, 1981.

[17] H. Whitney, Differentiable manifolds in Euclidean spaces, *Proceedings of the National Academy of Sciences*, vol. 21(7), pp. 462–464, 1935.

[18] H. Ye, Using rEDM to Quantify Time Delays in Causation (ha0ye.github.io/rEDM/articles/rEDM-time-delay-ccm.html), 2019.

[19] H. Ye, A. Clark, E. Deyle and G. Sugihara, rEDM: An R Package for Empirical Dynamic Modeling and Convergent Cross-Mapping (ha0ye.github.io/rEDM/articles/rEDM.html), 2019.