Scientific Research

Feasibility Study of Parameter Identification Method Based on Symbolic Time Series Analysis and Adaptive Immune Clonal Selection Algorithm

Rongshuai Li[*], Akira Mita, Jin Zhou
Department of System Design Engineering, Keio University, Yokohama, Japan
Email: [*]lrs0809@gmail.com

ABSTRACT

The feasibility of a parameter identification method based on symbolic time series analysis (STSA) and the adaptive immune clonal selection algorithm (AICSA) is studied. Data symbolization by using STSA alleviates the effects of harmful noise in raw acceleration data. The effect of the parameters in STSA is theoretically evaluated and numerically verified. AICSA is employed to minimize the error between the state sequence histogram (SSH) that is transformed from raw acceleration data by STSA. The proposed methodology is evaluated by comparing it with AICSA using raw acceleration data. AICSA combining STSA is proved to be a powerful tool for identifying unknown parameters of structural systems even when the data is contaminated with relatively large amounts of noise.

1. Introduction

Structural health monitoring (SHM) for predicting the onset of damage and deterioration of building structures is receiving more and more attention because of the rising numbers of aged structures and high costs caused by unpredictable hazards.

Some success has been achieved with various heuristic optimization algorithms such as genetic algorithms (GAs), evolution strategy (ES), simulated annealing (SA), particle swarm optimization (PSO), clonal selection algorithm (CSA), and differential evolution (DE). These heuristic stochastic search techniques seem to be a promising alternative to traditional approaches. The SA and GA methods have been used to accurately describe the dynamic behaviors of structures [1]. Cunha & Smith used GAs to identify the elastic constants of composite materials [2]. PSO has been used to estimate the severity of damage and identify parameters of shear frame building structures [3]. An improved CSA, called adaptive immune CSA (AICSA), has been used for structural damage localization and quantification [4,5]. Moreover, DE has been used to identify induction motor problems [6] and structural systems [7]. These heuristic approaches are very powerful in many applications. However, they are often sensitive to noise.

Symbolic time series analysis (STSA) for anomaly detection in complex systems [8] has the potential to deal with noise. Several case studies [9-11] have shown that STSA is more effective at anomaly detection than pattern recognition techniques such as principal component analysis and neural networks. STSA has also been used for fault detection in electromechanical systems, such as in three-phase induction motors [12] and helical gearboxes in rotorcraft [13].

We studied the feasibility of using the Euclidean distance of a state sequence histogram (statistical features of the symbol series that transformed from time series data) of symbols as an objective function of AICSA for the purpose of identifying structural parameters. We theoretically investigated the effects of parameters in STSA and conducted various numerical tests to show how combining AICSA and STSA improves performance of structural parameter identification. The results show that with the proper parameters, our methodology is a reliable and effective way of identifying structural parameters.

2. Symbolic Time Series Analysis

It may be appropriate to say that, while classical data analysis focuses on individuals, symbolic data analysis deals with concepts, a less specific type of information. Through symbolic conversion, the original time series

---

[*]Corresponding author.

signals are converted into sequences of discrete symbols, and the statistical features of the symbols can be used to describe the dynamic statuses of a system.

Consider a structural system Σ. The response of raw acceleration data can be recorded by using sensors. A section of this data is $\{\ddot{x}_0, \ddot{x}_1, \cdots, \ddot{x}_{T-1}\}$, which can be obtained by sliding a rectangular window with length T along the time series of raw acceleration. The first step is to transform the raw acceleration data into a binary symbol series $\{\sigma_0, \sigma_1, \cdots, \sigma_{T-1}\}$. $\sigma_i (i \in [0, T-1])$ equals "0" or "1" due to a partition line. After that, we select an integer $r \geq 1$ (word length) and define the symbolic state at time t as the vector s_t containing the follow-up r output symbols, namely

$$s_t = [\sigma_t, \sigma_{t+1}, \cdots, \sigma_{t+r-1}], t \in [0, T-r+1] \qquad (1)$$

s_t defines a state series $\{s_0, s_1, \cdots, s_{T-r+1}\}$. A binary coded s_t should be transformed into the decimal domain, and note that s_t can take $Q = 2^r$ possible values (called "states"), which can be listed in a finite set $S = \{0, 1, \cdots, Q-1\}$. We can then derive the statistics of the symbolic state, *i.e.*, compute the vector of the observed state frequencies $D = [d^0, d^1, \cdots, d^{Q-1}]$, where d^i (integer $i \in [0, Q-1]$) is the number of occurrences of $S = i$. Also, since there are $T - r + 1$ states in the state series in total, D can be normalized as $\dfrac{D}{(T-r+1)}$.

In the example shown in **Figure 1**, the window length $T = 26$ and the sampling points of a raw acceleration data series are shown as small circles, which have different values; the *x*-axis is time and the *y*-axis is acceleration data. The partition line is the one with the mean value of the raw acceleration data series. Thus, the whole space is separated into two regions. The acceleration data that falls inside the upper region is symbolized by "1"; otherwise, it is "0". The result of the symbolization is a binary coded symbol series that only contains "0" and "1". In this example, a word length of 3 was used to create words, which means the first three symbols "1 0 0" are chosen as the first word, and the second to fourth

symbols "0 0 0" are chosen as the second word. By repeating this procedure, 24 words can be created from the symbol series. Every binary coded word needs to be transformed into the decimal domain. Take the first word as an example. "1 0 0" can be transformed to $4 (1 \times 2^2 + 0 \times 2^1 + 0 \times 2^0)$, which is called a "state". A state series can be obtained after all the words are transformed from the binary domain to the decimal domain, which constitutes the values 0 - 7.

As shown in **Figure 1**, the occurrence number of certain states in the state series varies. A bar graph used to plot the occurrence number of every state in a state series is called a "state sequence histogram" (SSH). The corresponding SSH for this example is plotted in **Figure 2(a)**. Taking state "5" as an example, the corresponding count number is "3", meaning that state "5" occurs three times in the state series (as marked in the state series of **Figure 1**). Also, the SSH can be normalized, which can be accomplished by dividing the occurrence number of each state by the total number of states in the state series (**Figure 2(b)**).

3. Proposed Method

3.1. Procedure

In the research field of structural parameter identification, the time response of the system is usually compared with that of a parameterized model using a norm or some performance criterion to give us a measure of how well the model explains the system.

We will explain our methodology (**Figure 3**) using a physical system with input u and output y. Let $y(t_i)(i = 1, \cdots, T)$ denote the value of the actual system at the *i*th discrete time step. Suppose that a parameterized model able to capture the behavior of the physical system is developed and this model depends on a set of n parameters, *i.e.*, $x = (x_1, x_2, \cdots, x_n)^T \in R^n$. Given a candidate parameter value x and a guess $\hat{X}_0$ of the initial state, $\hat{y}(t_i)(i = 1, \cdots, T)$, the value of the parameterized model, *i.e.*, the identified system at the *i*th discrete time
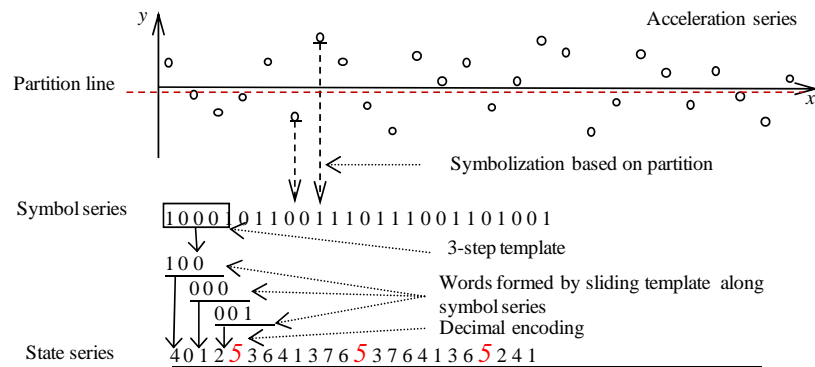

Figure 1. Process of symbolizing a time series of raw acceleration data.
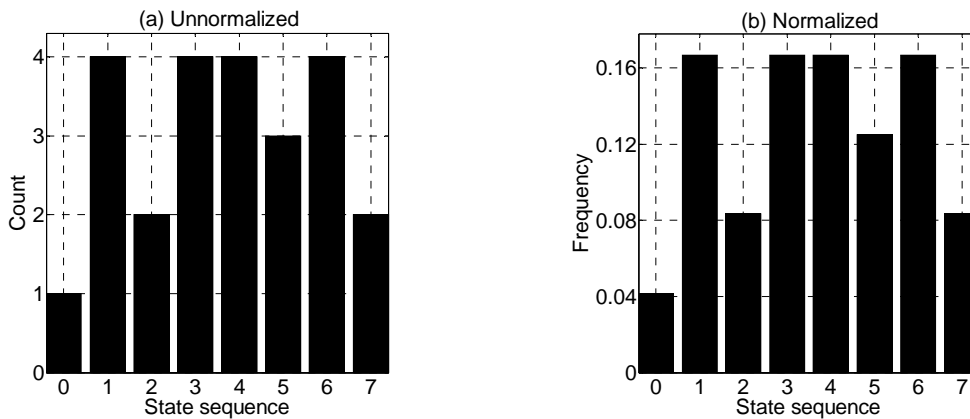
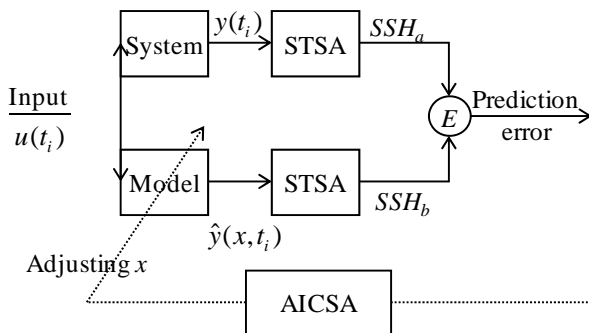**Figure 2. State sequence histogram (SSH).**

Figure 3. Procedure of AICSA combining STSA for identification of structural parameters.

step, can be obtained. Hence, the problem of system identification boils down to finding a set of parameters that minimize the prediction error between the system output $y(t_i)$, which is the measured data, and the model output $\hat{y}(x,t_i)$, which is calculated at each time instant t_i.

Usually, our interest lies in minimizing the predefined error norm of the time series outputs, e.g., the following mean square error (MSE) function,

$$f(x) = \frac{1}{T}\sum_{i=1}^{T} \left\| y(t_i) - \hat{y}(x,t_i) \right\|^2 \qquad (2)$$

where $\|\cdot\|$ represents the Euclidean norm of vectors. Formally, the optimization problem requires one to find a set of n parameters $x^* \in R^n$ so that a certain quality criterion is satisfied, namely, that the error norm $f(\cdot)$ is minimized. The function $f(\cdot)$ is called a fitness function or objective function. Typically, an objective function that reflects the goodness of the solution is chosen.

In our methodology, we introduce an index, the relative state sequence histogram error (RSSHe), to measure the distance between SSH_a and SSH_b (SSH_a and SSH_b are the system output and model output, respectively). The definition is:

$$\begin{cases} SSH_a = \left[d_a^0, d_a^1, \ldots, d_a^{Q-1} \right] \\ SSH_b = \left[d_b^0, d_b^1, \ldots, d_b^{Q-1} \right] \end{cases}, Q = 2^r$$

$$RSSHe = \sqrt{\frac{\sum_{i=0}^{i=Q-1}\left(d_b^i - d_a^i \right)^2}{\sum_{i=0}^{i=Q-1}\left(d_a^i \right)^2}} \qquad (3)$$

where $d_{a/b}^i$ is the frequency of state i in SSH_a or SSH_b.

Inspired by the clonal selection principle (CSP), the clonal selection algorithm (CSA) has been used to deal with optimization problems because of its search capability is superior to those of classical optimization techniques [14].

Although CSA has great advantages over the genetic algorithm (GA), it is still difficult to use it to solve complex problems. To be able to solve complex problems, in AICSA, three strategies, *i.e.*, secondary response, adaptive mutation regulation and vaccination, are used to improve the CSA's convergence speed and global optimum search. For detailed information about AICSA, please refer to [4,15].

3.2. Guideline for Parameter Selection

In STSA, the main parameters are the word length and window length, and they control the resolution of the whole representation space. For a window length T and word length r, two limiting cases of SSH are predefined as:

Case 1: All states in the SSH are distributed uniformly, and the frequency of each state is $\frac{1}{2^r}$.

Case 2: Only one state in the SSH has the frequency of 1; the frequencies of the other states are 0.

Suppose there are two different $SSHs$: SSH_a and SSH_b. From Equation (3), when SSH_a corresponds to limiting case 1 and SSH_b to limiting case 2, the maximum value of $RSSHe$ is:

$$RSSHe_{\max} = \sqrt{\frac{\sum_{i=0}^{i=Q-1}\left(d_b^i - d_a^i\right)^2}{\sum_{i=0}^{i=Q-1}\left(d_a^i\right)^2}}$$

$$= \sqrt{\frac{(1-\frac{1}{2^r})^2 + \left(\frac{1}{2^r}\right)^2\left(2^r-1\right)}{\left(\frac{1}{2^r}\right)^2 * 2^r}} = \sqrt{2^r-1} \qquad (4)$$

when SSH_a and SSH_b are the same, the minimum $RSSHe$ is 0. Then,

$$RSSHe \in \left[RSSHe_{\min}, RSSHe_{\max}\right] = \left[0, \sqrt{2^r-1}\right] \qquad (5)$$

Since the minimum changeable unit in SSH is $\frac{1}{T-r+1}$, the change in frequency of one state in SSH will absolutely be related to the change in frequencies of other states. Supposing that there are only two minimum unit differences between SSH_a and SSH_b, the minimum distinguishable $RSSHe$ is:

$$RSSHe_{\mathrm{dis}} = \sqrt{\frac{\sum_{i=0}^{i=Q-1}\left(d_b^i - d_a^i\right)^2}{\sum_{i=0}^{i=Q-1}\left(d_a^i\right)^2}}$$

$$= \sqrt{\frac{\left(\frac{1}{T-r+1}\right)^2 + \left(\frac{1}{T-r+1}\right)^2}{\sum_{i=0}^{i=Q-1}\left(d_a^i\right)^2}} = \frac{\sqrt{\frac{2}{\sum_{i=0}^{i=Q-1}\left(d_a^i\right)^2}}}{T-r+1} \qquad (6)$$

when SSH_a is limiting case 1, the maximum distinguishable $RSSHe_{\mathrm{dis}}^{\max}$ will be:

$$RSSHe_{\mathrm{dis}}^{\max} = \frac{\sqrt{2^{(r+1)}}}{T-r+1} \qquad (7)$$

when SSH_a is limiting case 2, the minimum distinguishable $RSSHe_{\mathrm{dis}}^{\min}$ will be:

$$RSSHe_{\mathrm{dis}}^{\min} = \frac{\sqrt{2}}{T-r+1} \qquad (8)$$

The resolution is:

$$\left[RSSHe_{\mathrm{dis}}^{\min}, RSSHe_{\mathrm{dis}}^{\max}\right] = \left[\frac{\sqrt{2}}{T-r+1}, \frac{\sqrt{2^{(r+1)}}}{T-r+1}\right] \qquad (9)$$

Note that we also need to consider the number of the possible distributions of states in one SSH. If the number of states in SSH is 2^r and the minimum changeable unit is $\frac{1}{T-r+1}$, finding the total number of possible distributions N_{SSH} of SSH boils down to a classic combination problem, which is "put $T-r+1$ identical balls in 2^r different boxes". The combinatorial number is:

$$N_{SSH} = C_{T-r+2^r}^{2^r-1} \qquad (10)$$

As we can see, longer window and word lengths are related to higher resolution, which means that the self and non-self spaces can be separated much more accurately. This is the key to obtaining accurate structural parameter identification.

So far, our discussion of the effect of the window length and word length has been based on a case in which only one story's output (raw acceleration data) is used, but structures with multiple degrees of freedom (MDOF) may have more outputs than that. Supposing the outputs from N stories can be obtained, the boundary of the solution space is:

$$RSSHe \in \left[RSSHe_{\min}, RSSHe_{\max}\right]^N$$
$$= \left[0, \sqrt{2^r-1}\right]^N \qquad (11)$$

The resolution falls to:

$$\left[RSSHe_{\mathrm{dis}}^{\min}, RSSHe_{\mathrm{dis}}^{\max}\right]^N$$
$$= \left[\frac{\sqrt{2}}{T-r+1}, \frac{\sqrt{2^{(r+1)}}}{T-r+1}\right]^N \qquad (12)$$

Also, the total number of possible distributions increases to:

$$\left[N_{SSH}\right]^N = \left[C_{T-r+2^r}^{2^r-1}\right]^N \qquad (13)$$

From Equations (11) to (13), it is evident that as more story outputs are obtained, the more accurate the identification results will be.

4. Effects of Parameter Selection for SDOF Model

4.1. Description of SDOF Model

For simplicity and generality, we used a single-story shear frame structure as a representative case to verify the effect of the parameters in STSA, and we modeled it as a single degree-of-freedom (SDOF) lumped mass system (**Figure 4**). As for the structure, its mass was 1000 kg, stiffness 1.000 MN/m, and natural frequency 5.032 Hz. The dynamic equation is [16]:

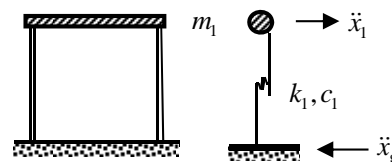$$M\ddot{X}_t + C\dot{X}_t + KX_t = f(t) \qquad (13)$$

Figure 4. Single-story shear frame structure.

where M, C, and K are respectively the mass matrix, damping matrix, and stiffness matrix. $f(t)$ is the force vector linked to the ground acceleration. $\ddot{X}$, $\dot{X}$, and X are respectively relative acceleration, velocity, and displacement response. The sampling frequency was 100 Hz. In the simulation, the input signal was Gaussian white noise. The root-mean-square error (*RMSe*) was used to verify the feasibility and performance of the identification results. *RMSe* is defined as

$$RMSe = \sqrt{\frac{\sum_{i=1}^{n}\left(k_{c,i} - k_{\text{real},i}\right)^2}{\sum_{i=1}^{n} k_{\text{real},i}^2}} \qquad (14)$$

where $k_{c,i}$ and $k_{\text{real},i}$ are the candidate stiffness and real stiffness of the *i*th story, respectively.

To test the noise immunity of our method, noise at levels of 5%, 10% or 20% was added to the raw acceleration data.

4.2. Effect of Varying the Window Length and Word Length

As stated before, the window length and word length are the control parameters in STSA. In the simulation, the mass distribution and damping parameters were assumed to be known and the stiffness of each story was set as the objective parameters that needed to be identified. In the first verification, the word length was varied from 1 to 12 and the window length was 3000. In the second verification, the word length was 9 and the window length was varied from 500 to 6000 at intervals of 500. Each test was run 10 times independently by choosing an initial value of AICSA randomly every time. The parameters of AICSA were the same as in [4,5].

Figures 5 and **6** indicate that the word length and window length greatly affected the performance of the methodology. Larger word and window lengths yielded better performance. The reason is that, as theoretically shown in Section 3.2 (Equations (4)-(10)), a longer word or window can symbolize the raw acceleration data much more accurately than a shorter one. As the word length and window length increase, much more dynamic information about the system is captured, the identified results become more accurate, and the maximum and mean *RMSe* decrease. In the simulation, a word length more than 9 and window length more than 3.0E+03 gave acceptable results.

Table 1 (under the label "STSA") lists the identification results of the SDOF structure using a word length of 9 and a window length of 3000 for different noise levels. For comparison, we estimated the parameters and *RMSe* for an SDOF model using raw acceleration data as input, for a data length of 3000, *i.e.*, the same as the window length in STSA; the results are listed under the label "RAW".
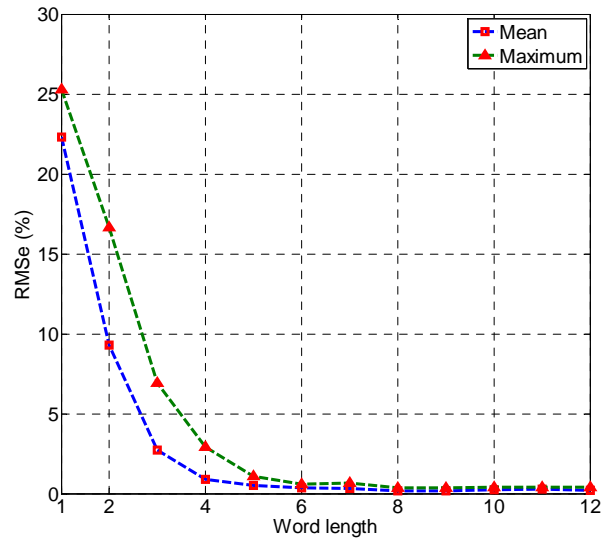

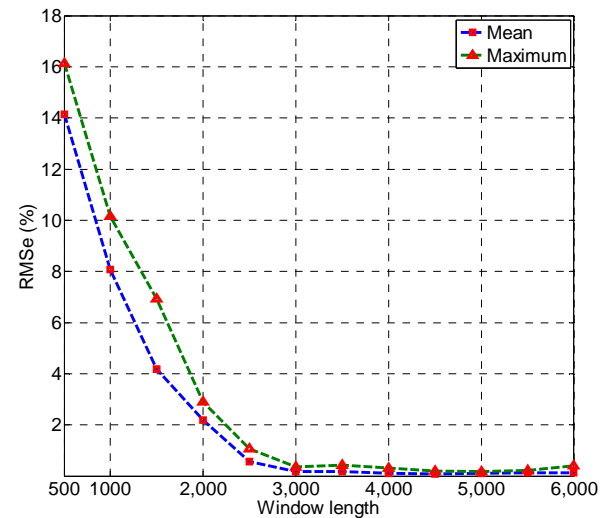**Figure 5. Effect of varying word length for a window length of 3000.**

Figure 6. Effect of varying the window length for a word length of 9.

Table 1. Estimated parameters and *RMSe* for SDOF model using STSA (with word length 9 and window length 3000) or raw acceleration data as input.

		True value (MN/m)	Calculated value (MN/m)			
			No noise	5% noise	10% noise	20% noise
STSA	k	1.000	1.000	1.000	1.000	1.000
	RMSe		0.00%	0.00%	0.04%	0.04%
RAW	k	1.000	1.000	0.991	1.031	1.065
	RMSe		0.00%	0.90%	3.12%	6.54%

As we can see, although AICSA using raw acceleration data gives good results for the noise-free case, its *RMSe* greatly increases as the noise level grows. In con-

trast, our method has good noise immunity; the identification results stay accurate as the noise level increases. Even for the high noise level of 20%, the *RMSe* of the results is only 0.04%.

5. Extension to MDOF Models

5.1. Description of MDOF Models

Next, we tried to see if our methodology can be reliably used to identify the parameters of an MDOF system. Here, we choose three cases, a 3-DOF, 5-DOF (shown in **Figure 7** as an example) and 10-DOF structure, as examples. These structures were modeled as multiple degree-of-freedom lumped mass systems. **Table 2** summarizes the structural parameters. In these structures, the mass of each story was 1000 kg, and the stiffness of each story was 2.000 MN/m. The damping ratios of all MDOF structures were the same, *i.e.*, 0.03 and 0.05 for the first and second modes, respectively.

5.2. *RMSe* for MDOF Models

In the input signal of the MDOF simulation, was Gaussian white noise, as was used in the SDOF simulation. The stiffness of each story was unknown and needed to be identified. The methodology of combining AICSA with STSA was compared with AICSA using raw acceleration data. Window and word lengths were the same as before, and the full output of the structure was used.

   Table 3 lists the identified stiffness of each story of the 3, 5 and 10 DOF models. The estimated parameters for the 3, 5 and 10 DOF models using AICSA using raw acceleration are not listed for lack of space, but **Table 4** summarizes the *RMSe* of the identification results of the 3, 5 and 10 DOF models using AICSA combining with STSA as well as AICSA using raw acceleration data.
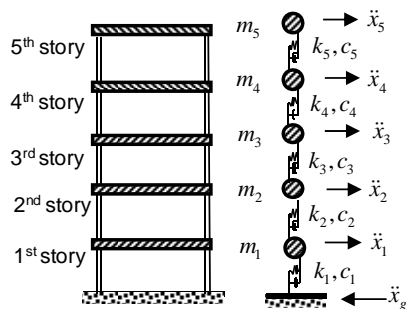

Figure 7. 5-DOF shear frame structure.

Table 2. Parameters of MDOF models.

DOF	1	2	3	4	5
Mass (kg)	1000	1000	1000	1000	1000
Stiffness (MN/m)	2.000	2.000	2.000	2.000	2.000

Table 3. Estimated parameters for 3, 5 and 10 DOF models with a word length of 9 and window length of 3000.

		True value (MN/m)	Calculated value (MN/m)			
			No noise	5% noise	10% noise	20% noise
3-DOF	k_1	2.000	2.000	1.994	2.006	2.028
	k_2	2.000	2.000	1.998	2.025	1.984
	k_3	2.000	2.000	2.017	1.992	1.992
5-DOF	k_1	2.000	2.000	2.001	2.000	2.014
	k_2	2.000	2.000	1.999	2.003	1.989
	k_3	2.000	2.000	2.021	1.987	2.007
	k_4	2.000	2.000	1.997	2.031	1.963
	k_5	2.000	2.000	2.000	1.993	1.996
10-DOF	k_1	2.000	2.000	2.004	1.996	1.995
	k_2	2.000	2.000	1.996	2.013	1.983
	k_3	2.000	2.000	2.009	2.001	2.023
	k_4	2.000	2.000	1.997	2.020	1.998
	k_5	2.000	2.000	2.005	1.993	2.003
	k_6	2.000	2.000	2.009	1.982	2.004
	k_7	2.000	2.000	1.999	1.989	1.969
	k_8	2.000	2.000	2.013	2.001	1.996
	k_9	2.000	2.000	2.005	2.000	1.990
	k_{10}	2.000	2.000	1.991	1.993	2.001

Table 4. *RMSe* for 3, 5 and 10 DOF models using STSA or raw acceleration as input.

		RMSe			
		No noise	5% noise	10% noise	20% noise
SATA	3-DOF	0.00%	0.91%	1.35%	1.66%
	5-DOF	0.00%	1.08%	1.74%	2.08%
	10-DOF	0.02%	1.11%	1.68%	2.21%
RAW	3-DOF	0.00 %	3.11%	5.38%	8.39 %
	5-DOF	0.00%	1.47%	3.99%	5.67%
	10-DOF	0.07%	4.26%	5.84%	10.01%

   Figure 8 compares the *RMSes* of the identification results for different structures when using our methodology (the "STSA" column in **Tables 1** and **4**). As we can see, AICSA combining STSA can identify the parameters of a structure accurately regardless of whether the structure is SDOF or MDOF. *RMSe* does increase slightly as the DOF of the structure increase, because the
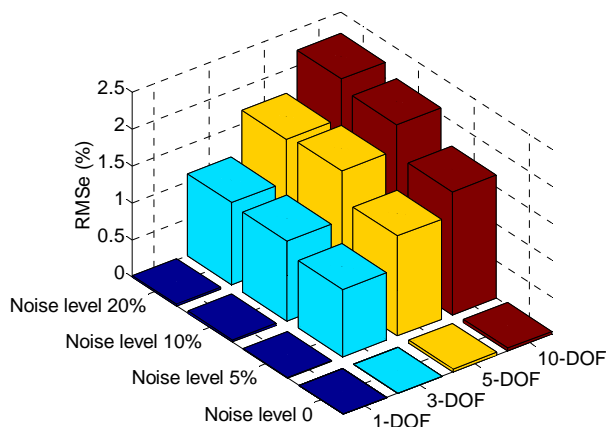
Figure 8. *RMSe* for 1, 3, 5 and 10 DOF models.

solution space of the identification problem becomes much more complex as the DOF go up.

Moreover, AICSA combining STSA outperformed AICSA using raw acceleration data (results under the "RAW" label in **Table 1** and **4**) on the MDOF models. Furthermore, it provided much better estimates when the output data was contaminated with noise. These results clearly show that our methodology has excellent noise immunity.

5.3. Estimation Using Partial Outputs

The simulations results of the MDOF structures are based on the full output of the structural acceleration data. For an SDOF structure, only one output can be used, but for MDOF structures, it may be the case that not all outputs are available. Therefore, to verify the methodology on only partial output, a 5-DOF structure was simulated and data from some of its stories (randomly chosen) were used. The simulated output data was noise-free, with 5% noise, 10% noise, or 20% noise. The window length was 3000, and the word length was 9.

The results in **Figure 9** illustrate that even using partial data, the proposed method gets acceptable results. Note that for a certain noise level, the *RMSe* of the identification results increase as the number of outputs decreases. Moreover, Equations (11)-(13) can be numerically proved to be right as t more outputs are obtained.

6. Conclusion

We conducted a feasibility study of a parameter identification method based on symbolic time series analysis (STSA) and adaptive immune clonal selection algorithm (AICSA). Harmful noise in the raw acceleration data was alleviated by employing STSA. The effect of varying the parameters of word length and window length in STSA was evaluated theoretically and verified numerically. A comparison with AICSA using raw acceleration data
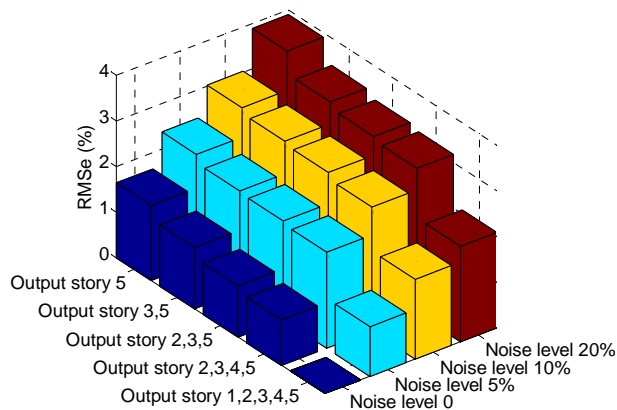

Figure 9. Comparison of *RMSe* due to partial output.

revealed that our methodology provided better estimates of structural parameters when the data was contaminated by noise. The results show that with the proper parameters, our methodology is a reliable and effective method for structural parameter identification.

7. Acknowledgements

REFERENCES

[1] R. I. Levin and N. A. J. Lieven, "Dynamic Finite Element Model Updating Using Simulated Annealing and Genetic Algorithms," *Mechanical Systems and Signal Processing*, Vol. 12, No. 1, 1998, pp. 91-120. doi:10.1006/mssp.1996.0136

[2] K. Cunha and V. V. Smith, "A Determination of the Solar Photospheric Boron Abundance," *Astrophysical Journal*, Vol. 512, No. 2, 1999, pp. 1006-1013. doi:10.1086/306796

[3] S. T. Xue, H. S. Tang and J. Zhou, "Identification of Structural Systems Using Particle Swarm Optimization," *Journal of Asian Architecture and Building Engineering*, Vol. 8, No. 2, 2009, pp. 517-524. doi:10.3130/jaabe.8.517

[4] R. Li and A. Mita, "Structural Damage Identification Using Adaptive Immune Clonal Selection Algorithm and Acceleration Data," *SPIE Smart Structures*, Vol. 7981, 2011, Article ID: 79815A.

[5] R. Li and A. Mita, "Hybrid Immune Algorithm for Structural Health Monitoring Using Acceleration Data," *8th International Workshop on Structural Health Monitoring*, Stanford, September 2011, pp. 1095-1102.

[6] R. K. Ursem and P. Vadstrup, "Parameter Identification of Induction Motors Using Differential Evolution," *Proceedings of the 5th Congress on Evolutionary Com-*

putation, Vol. 2, 2003, pp. 790-796.

[7] H. S. Tang, S. T. Xue and C. X. Fan, "Differential Evolution Strategy for Structural Parameter Identification," *Computers & Structures*, Vol. 86, No. 21-22, 2008, pp. 2004-2012. doi:10.1016/j.compstruc.2008.05.001

[8] A. Ray, "Symbolic Dynamic Analysis of Complex Systems for Anomaly Detection," *Signal Processing*, Vol. 84, No. 7, 2004, pp. 1115-1130. doi:10.1016/j.sigpro.2004.03.011

[9] S. Chin, A. Ray and V. Rajagopalan, "Symbolic Time Series Analysis for Anomaly Detection: A Comparative Evaluation," *Signal Processing*, Vol. 85, No. 9, 2005, pp. 1859-1868. doi:10.1016/j.sigpro.2005.03.014

[10] A. Khatkhate, A. Ray, S. Chin, V. Rajagopalan and E. Keller, "Detection of Fatigue Crack Anomaly: A Symbolic Dynamic Approach," *Proceedings of American Control Conference*, Boston, June-July 2004, pp. 3741-3746.

[11] D. Tolani, M. Yasar, A. Ray and V. Yang, "Anomaly Detection in Aircraft Gas Turbine Engines," *AIAA Journal of Aerospace Computing Information*, *and Communication*, Vol. 3, No. 2, 2006, pp. 44-51. doi:10.2514/1.15768

[12] S. Bhatnagar, V. Rajagopalan and A. Ray, "Incipient Fault Detection in Mechanical Power Transmission Systems," *Proceedings of American Control Conference*, Portland, 8-10 June 2005, pp. 472-477. doi:10.1109/ACC.2005.1469980

[13] R. Samsi, V. Rajagopalan, J. Mayer and A. Ray, "Early Detection of Voltage Imbalances in Induction Machines," *Proceedings of American Control Conference*, Portland, 8-10 June 2005, pp. 478-483. doi:10.1109/ACC.2005.1469981

[14] X. Wang, X. Z. Gao and S. J. Ovaska, "Artificial Immune Optimization Methods and Applications—A Survey," *IEEE International Conference on Systems*, *Man and Cybernetics*, Vol. 4, 2004, pp. 3415-3420.

[15] L. Zhang, X. R. Meng, W. J. Wu and H. Zhou, "Network Fault Feature Selection Based on Adaptive Immune Clonal Selection Algorithm," *International Joint Conference on Computational Sciences and Optimization*, Sanya, 24-26 April 2009, pp. 969-713. doi:10.1109/CSO.2009.342

[16] A. Mita, "Structural Dynamics for Health Monitoring," Sankeisha Co. Ltd., Fairport, 2003.