References

Balda, M. An estimation of the residual life of blades. In Proceedings of Colloquium Dynamics of Machines (DM'99), pp. 11–18, Prague, Czech Republic, 1999.

Berndt, D. J. and Clifford, J. Using dynamic time warping to find patterns in time series. In Proceedings of the 1994 AAAI Workshop on Knowledge Discovery in Databases (KDD Workshop'94), pp. 359–370, Seattle, WA, USA, 1994.

B.L.R., D. M. DaISy: Database for the Identification of Systems. [Online]. Available from : `http://www.esat.kuleuven.ac.be/sista/daisy/` , 2010.

Burden, R. L., Faires, J. D., and Reynolds, A. G. Numerical Analysis. Brooks Cole, 1997.

Cai, Y. and Ng, R. T. Indexing spatio-temporal trajectories with chebyshev polynomials. In Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD'04), pp. 599–610, Paris, France, 2004.

Chen, J. Useful clustering outcomes from meaningful time series clustering. In Proceedings of the 6^{th} Australasian Data Mining Conference on Data Mining and Analytics 2007 (AusDM'07), pp. 101–109, Gold Coast, Queensland, Australia, December 3-4 2007a.

Chen, J. R. Making clustering in delay-vector space meaningful. Knowledge Information System, 11,3:369–385, 2007b.

Chiu, B., Keogh, E. J., and Lonardi, S. Probabilistic discovery of time series motifs. In Proceedings of the 9^{th} ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'03), pp. 493–498, Washington, DC, USA, August 24 - 27 2003.

Cotofrei, P. Statistical temporal rules. In Proceedings of the 15^{th} Conference on Computational Statistics (COMPSTAT'02), pp. 24–28, Berlin, Germany, 2002.

Cotofrei, P. and Stoffel, K. Classification rules + time = temporal rules. In Proceedings of the International Conference on Computational Science (ICCS'02), pp. 572–581, Amsterdam, The Netherlands, April 21-24 2002.

Dai, B.-R., Huang, J.-W., Yeh, M.-Y., and Chen, M.-S. Adaptive clustering for multiple evolving streams. IEEE Transactions on Knowledge and Data Engineering, 18,9:1166–1180, 2006.

Das, G., Lin, K.-I., Mannila, H., Renganathan, G., and Smyth, P. Rule discovery from time series. In Proceedings of the 4th International Conference on Knowledge Discovery and Data Mining (KDD'98), pp. 16–22, New York City, NY, USA, August 27-31 1998.

Denton, A. Kernel-density-based clustering of time series subsequences using a continuous random-walk noise model. In Proceedings of the 5th IEEE International Conference on Data Mining (ICDM'05), pp. 122–129, Houston, TX, USA, 2005.

Fletcher, R. A modified marquardt subroutine for nonlinear least squares. Technical report, AERE R6799, Atomic Energy Research Establishment, 1971.

Fu, T.-C., Chung, F.-L., Ng, V., and Luk, R. Pattern discovery from stock time series using self-organizing maps. In Workshop Notes of KDD'01 Workshop on Temporal Data Mining, pp. 27–37, 2001.

Fu, T.-C., Chung, F.-L., Luk, R. W. P., and Ng, C.man . Preventing meaningless stock time series pattern discovery by changing perceptually important point detection. In Proceedings of Second International Conference on Fuzzy Systems and Knowledge Discovery (FSKD'05), pp. 1171–1174, Changsha, China, August 27-29 2005.

Fujimaki, R., Hirose, S., and Nakata, T. Theorectical analysis of subsequence time-series clustering from a frequency-analysis viewpoint. In Proceedings of the SIAM International Conference on Data Mining (SDM'08), pp. 506–517, Atlanta, GA, USA, April 24-26 2008.

Goldin, D. Q., Mardales, R., and Nagy, G. In search of meaning for time series subsequence clustering: matching algorithms based on a new distance measure. In Proceedings of the 2006 ACM CIKM International Conference on Information and Knowledge Management (CIKM'06), pp. 347–356, Arlington, VA, USA, November 6-11 2006.

Guha, S., Rastogi, R., and Shim, K. ROCK: A robust clustering algorithm for categorical attributes. Information Systems, 25,5:345–366, 2000.

Gupta, L., Molfese, D. L., and Tammana, R. Nonlinear alignment and averaging for estimating the evoked potential. IEEE Transactions on Biomedical Engineering, 43,4:348–356, April 1996.

Han, J. and Kamber, M. Data Mining: Concepts and Techniques. Morgan Kaufmann, 2000.

Harms, S. K., Deogun, J. S., and Tadesse, T. Discovering sequential association rules with constraints and time lags in multiple sequences. In Proceedings of the 13th International Symposium Foundations of Intelligent Systems (ISMIS'02), pp. 432–441, Lyon, France, June 27-29 2002a.

Harms, S. K., Li, D., Deogun, J. S., and Tadesse, T. Data mining in a geospatial decision support system for drought risk management. In Proceedings of the 3rd National Conference on Digital Government Research (DG.O'02), pp. 9 – 16, Los Angeles, CA, USA, May 21 - 23 2002b.

Hazewinkel, M., editor. Encyclopaedia of Mathematics, chapter Sinusoid. Springer, 2001.

Hetland, M. L. Temporal rule discovery using genetic programming and specialized hardware. In Proceedings of the 4th International Conference on Recent Advances in Soft Computing (RASC'02), pp. 12–13, Nottingham, United Kingdom, 2002.

Idé, T. Translational symmetry in subsequence time-series clustering. In Proceedings of Conference and Workshops on New Frontiers in Artificial Intelligence (JSAI'06), pp. 5–18, Tokyo, Japan, June 5-9 2006a.

Idé, T. Why does subsequence time-series clustering produce sine waves? In Proceedings of 10th European Conference on Principles and Practice of Knowledge DiscoveryKnowledge Discovery in Databases (PKDD'06), pp. 211–222, Berlin, Germany, September 18-22 2006b.

Jin, X., Lu, Y., and Shi, C. Distribution discovery: Local analysis of temporal rules. In Proceedings of the 6th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD'02), pp. 469–480, Taipei, Taiwan, May 6-8 2002a.

Jin, X., Wang, L., Lu, Y., and Shi, C. Indexing and mining of the local patterns in sequence database. In Proceedings of the Third International Conference on Intelligent Data Engineering and Automated Learning (IDEAL'02), pp. 68–73, Manchester, UK, August 12-14 2002b.

Karypis, G., Han, E.-H., and Kumar, V. Chameleon: Hierarchical clustering using dynamic modeling. IEEE Computer, 32,8:68–75, 1999.

Kasetty, S., Stafford, C., Walker, G. P., Wang, X., and Keogh, E. J. Real-time classification of streaming sensor data. In Proceedings of the 20th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'08), pp. 149–156, Dayton, OH, USA, November 3-5 2008.

Kaufman, L. and Rousseeuw, P. J. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley-Interscience, 2005.

Keogh, E. J. and Folias, T. The UCR Time Series Data Mining Archive. [Online]. Available from : `http://www.cs.ucr.edu/~eamonn/TSDMA/` , 2011.

Keogh, E. J. and Lin, J. Clustering of time-series subsequences is meaningless: implications for previous and future research. Knowledge and Information Systems, 8,2:154–177, 2005.

Keogh, E. J. and Ratanamahatana, C. A. Exact indexing of dynamic time warping. Knowledge and Information Systems, 7,3:358–386, 2005.

Keogh, E. J., Lonardi, S., and Chiu, B. Y.-C. Finding surprising patterns in a time series database in linear time and space. In Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'02), pp. 550–556, Edmonton, Alberta, Canada, July 23-26 2002.

Keogh, E. J., Lin, J., and Truppel, W. Clustering of time series subsequences is meaningless: Implications for previous and future research. In Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM'03), pp. 115–122, Melbourne, FL, USA, 19-22 December 2003.

Keogh, E. J., Palpanas, T., Zordan, V. B., Gunopulos, D., and Cardle, M. Indexing large human-motion databases. In Proceedings of the 13th International Conference on Very Large Data Bases (VLDB'04), pp. 780–791, Toronto, Canada, August 31 - September 3 2004.

Keogh, E. J., Lin, J., and Fu, A. W.-C. Hot SAX: Efficiently finding the most unusual time series subsequence. In Proceedings of the 5th IEEE International Conference on Data Mining (ICDM'05), pp. 226–233, Houston, TX, USA, November 27-30 2005.

Keogh, E. J., Xi, X., Wei, L., and Ratanamahatana, C. A. UCR Time Series Classification/Clustering Page. [Online]. Available from : `http://www.cs.ucr.edu/~eamonn/time_series_data` , 2011.

Kontaki, M., Papadopoulos, A. N., and Manolopoulos, Y. Continuous subspace clustering in streaming time series. Information Systems, 33,2:240–260, 2008.

Kumar, N., Lolla, V. N., Keogh, E. J., Lonardi, S., and Ratanamahatana, C. A. Time-series bitmaps: A practical visualization tool for working with large time series databases. In Proceedings of the 5th SIAM International Conference on Data Mining (SDM'05), pp. 531–535, Newport Beach, CA, USA, 2005.

Kumar, R. P., Nagabhushan, P., and Chouakria, A. D. Wavesim and adaptive wavesim transform for subsequence time-series clustering. In Proceedings of 9th International Conference in Information Technology (ICIT'06), pp. 197–202, Orissa, India, 18-21 December 2006.

Li, C.-S., Yu, P. S., and Castelli, V. MALM: A framework for mining sequence database at multiple abstraction levels. In Proceedings of the 1998 ACM CIKM International Conference on Information and Knowledge Management (CIKM'98), pp. 267–272, Bethesda, MD, USA, November 3-7 1998.

Lin, J., Keogh, E. J., and Truppel, W. Clustering of streaming time series is meaningless. In Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD'03), pp. 56–65, San Diego, CA, USA, June 13 2003.

Lin, J., Keogh, E. J., Lonardi, S., Lankford, J. P., and Nystrom, D. M. Visually mining and monitoring massive time series. In Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'04), pp. 460–469, Seattle, WA, USA, August 22-25 2004a.

Lin, J., Vlachos, M., Keogh, E. J., and Gunopulos, D. Iterative incremental clustering of time series. In Proceedings of 9th International Conference on Extending Database Technology (EDBT'04), pp. 106–122, Crete, Greece, March 14-18 2004b.

Lloyd, S. P. Least squares quantization in PCM. IEEE Transactions on Information Theory, 28,2: 129–136, 1982.

MacQueen, J. Some methods for classification and analysis of multivariate observations. In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, volume 1, pp. 281–297, San Francisco, CA, USA, 1967.

Mori, T. and Kuni. Extraction of primitive motion and discovery of association rules from human motion. In Proceedings of the 10th IEEE International Workshop on Robot and Human Communication (ROMAN'01), pp. 18–21, Paris, France, 2001.

Mueen, A. and Keogh, E. J. Online discovery and maintenance of time series motifs. In Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'10), pp. 1089–1098, Washington, DC, USA, 2010.

Mueen, A., Keogh, E. J., Zhu, Q., Cash, S., and Westover, M. B. Exact discovery of time series motifs. In Proceedings of the SIAM International Conference on Data Mining (SDM'09), pp. 473–484, Sparks, NV, USA, April 30 - May 2 2009.

Niennattrakul, V. and Ratanamahatana, C. A. Inaccuracies of shape averaging method using dynamic time warping for time series data. In Proceedings of 7th International Conference on Computational Science (ICCS'07), pp. 513–520, Beijing, China, May 27-30 2007a.

Niennattrakul, V. and Ratanamahatana, C. A. On clustering multimedia time series data using k-means and dynamic time warping. In Proceedings of the International Conference on Multimedia and Ubiquitous Engineering (MUE'07), pp. 733–738, Seoul, Korea, April 26-28 2007b.

Niennattrakul, V. and Ratanamahatana, C. A. Meaningful subsequence matching under time warping distance for data stream. In Proceedings of 13th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD'09), pp. 1013–1020, Bangkok, Thailand, 2009.

Niennattrakul, V. and Ratanamahatana, C. A. Clustering multimedia data using time series. In Proceedings of the International Conference on Hybrid Information Technology (ICHIT'06), pp. 372–379, Cheju Island, Korea, November 9-11 2006.

Niennattrakul, V. and Ratanamahatana, C. A. Learning DTW global constraint for time series classification. In Time Series Classification Challenges at 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Jose, CA, USA, August 12-15 2007c.

Niennattrakul, V., Wanichsan, D., and Ratanamahatana, C. A. Accurate subsequence matching on data stream under time warping distance. In Proceedings of 13th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining (PAKDD'09) International Workshops, pp. 156–167, Bangkok, Thailand, 2009.

Niennattrakul, V., Ratanamahatana, C. A., and Keogh, E. Data editing techniques to allow the application of distance-based outlier detection to streams. In Proceedings of the 10th IEEE International Conference on Data Mining (ICDM'10), pp. 947–952, Sydney, Australia, 2010a.

Niennattrakul, V., Ruengronghirunya, P., and Ratanamahatana, C. A. Exact indexing for massive time series databases under time warping distance. Data Mining and Knowledge Discovery, 21,3:509–541, 2010b.

Osaki, R., Shimada, M., and Uehara, K. A motion recognition method by using primitive motions. In Proceedings of the 5th Working Conference on Visual Database Systems (VDB'00), pp. 117–128, Fukuoka, Japan, May 10-12 2000.

Peker, K. A. Subsequence time series (STS) clustering techniques for meaningful pattern discovery. In Proceedings of the International Conference on Integration of Knowledge Intensive Multi-Agent Systems (KIMAS'05), pp. 360–365, Waltham, MA, USA, 18-21, 2005.

Protopapas, P., Jimenez, R., and Alcock, C. Fast identification of transits from light-curves. Monthly Notices of the Royal Astronomical Society, 362,2, 2005.

Radhakrishnan, N., Wilson, J., and Loizou, P. An alternate partitioning technique to quantify the regularity of complex time series. International Journal of Bifurcation and Chaos, 10,7: 1773–1780, 2000.

Ratanamahatana, C. A. and Keogh, E. J. Multimedia retrieval using time series representation and relevance feedback. In Proceedings of 8th International Conference on Asian Digital Libraries (ICADL'05), pp. 400–405, Bangkok, Thailand, December 12-15 2005a.

Ratanamahatana, C. A. and Keogh, E. J. Making time-series classification more accurate using learned constraints. In Proceedings of the Fourth SIAM International Conference on Data Mining (SDM'04), pp. 11–22, Lake Buena Vista, FL, USA, April 22-24 2004.

Ratanamahatana, C. A. and Keogh, E. J. Three myths about dynamic time warping data mining. In Proceedings of the 5th SIAM International Conference on Data Mining (SDM'05), pp. 506–510, Newport Beach, CL, USA, April 21-23 2005b.

Rodrigues, P. P., Gama, J., and Pedroso, J. P. ODAC: Hierarchical clustering of time series data streams. In Proceedings of the 6th SIAM International Conference on Data Mining (SDM'06), Bethesda, MD, USA, April 20-22 2006.

Rodrigues, P. P., Gama, J., and Pedroso, J. P. Hierarchical clustering of time-series data streams. IEEE Transactions on Knowledge and Data Engineering, 20,5:615–627, 2008.

Sacchi, L., Larizza, C., Combi, C., and Bellazzi, R. Data mining with temporal abstractions: learning rules from time series. Data Mining and Knowledge Discovery, 15,2:217–247, 2007.

Saito, N. Local feature extraction and its applications using a library of bases. PhD thesis, Yale University, 1994.

Sakurai, Y., Yoshikawa, M., and Faloutsos, C. FTW: Fast similarity search under the time warping distance. In Proceedings of the 24th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS'05), pp. 326–337, Baltimore, MD, USA, 2005.

Salvador, S. and Chan, P. Toward accurate dynamic time warping in linear time and space. Intelligent Data Analysis, 11,5:561–580, 2007.

Sarker, B. K., Mori, T., Hirata, T., and Uehara, K. Parallel algorithms for mining association rules in time series data. In Proceedings of the International Symposium on Parallel and Distributed Processing and Applications (ISPA'03), pp. 273–284, Aizu, Japan, July 2-4 2003.

Schittenkopf, C., Tino, P., and Dorffner, G. The benefit of information reduction for trading strategies. Report Series for Adaptive Information Systems and Management in Economics and Management, 2000.

Shieh, J. and Keogh, E. J. SAX: Disk-aware mining and indexing of massive time series datasets. Data Mining and Knowledge Discovery, 19,1:24–57, 2009.

Simon, G., Lee, J. A., and Verleysen, M. Unfolding preprocessing for meaningful time series clustering. Neural Networks, 19,6-7:877–888, 2006.

Struzik, Z. R. Time series rule discovery: Tough, not meaningless. In Proceedings of 14th International Symposium on Foundations of Intelligent Systems (ISMIS'03), pp. 32–39, Maebashi City, Japan, October 28-31 2003.

Uehara, K. and Shimada, M. Extraction of primitive motion and discovery of association rules from human motion data. In Progress in Discovery Science, pp. 338–348, 2002.

Ueno, K., Xi, X., Keogh, E. J., and Lee, D.-J. Anytime classification using the nearest neighbor algorithm with applications to stream mining. In Proceedings of the 6th IEEE International Conference on Data Mining (ICDM'06), pp. 623–632, Hong Kong, China, 18-22 December 2006.

Wan, D., Zhang, Y., and Li, S. Discovery association rules in time series of hydrology. In Proceedings of IEEE International Conference on Integration Technology (ICIT'07), pp. 653 –657, Shenzhen, China, 2007.

Wikipedia.org. 12 lead ECG. [Online]. Available from : `http://en.wikipedia.org/wiki/Image:12leadECG.jpg` , 2011a.

Wikipedia.org. Comparison of three stock indices after 1975. [Online]. Available from : `http://en.wikipedia.org/wiki/Image:Comparison_of_three_stock_indices_after_1975.svg` , 2011b.

Wikipedia.org. Instrumental Temperature Record. [Online]. Available from : `http://en.wikipedia.org/wiki/Image:Instrumental_Temperature_Record.png` , 2011c.

Wu, W., Au, L., Jordan, B., Stathopoulos, T., Batalin, M., Kaiser, W., Vahdatpour, A., Sarrafzadeh, M., Fang, M., and Chodosh, J. The SmartCane system: An assistive device for geriatrics. In Proceedings of the ICST 3rd International Conference on Body Area Networks (BodyNets'08), pp. 2:1–2:4, Brussels, Belgium, March 13-15 2008.

Yairi, T., Kato, Y., and Hori, K. Fault detection by mining association rules from house-keeping data. In Proceedings of the 4th International Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS 2001), p. 555, Montreal, CA, USA, June 19-21 2001.

Yankov, D. and Keogh, E. J. Manifold clustering of shapes. In Proceedings of the 6th IEEE International Conference on Data Mining (ICDM'06), pp. 1167–1171, Hong Kong, China, December 18-22 2006.

Yankov, D., Keogh, E. J., Medina, J., Chiu, B., and Zordan, V. B. Detecting time series motifs under uniform scaling. In Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'07), pp. 844–853, San Jose, CA, USA, 2007.

Yankov, D., Keogh, E. J., and Rebbapragada, U. Disk aware discord discovery: Finding unusual time series in terabyte sized datasets. Knowledge and Information Systems, 17,2:241–262, 2008a.

Yankov, D., Keogh, E. J., Wei, L., Xi, X., and Hodges, W. L. Fast best-match shape searching in rotation-invariant metric spaces. IEEE Transactions on Multimedia, 10,2:230–239, 2008b.

Ye, L. and Keogh, E. J. Time series shapelets: A novel technique that allows accurate, interpretable and fast classification. Data Mining and Knowledge Discovery, 22,1-2:149–182, 2011.

Zhang, T., Ramakrishnan, R., and Livny, M. BIRCH: An efficient data clustering method for very large databases. In Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data (SIGMOD'96), pp. 103–114, Montreal, Quebec, Canada, June 4-6 1996.

APPENDICES

APPENDIX A

DATASETS

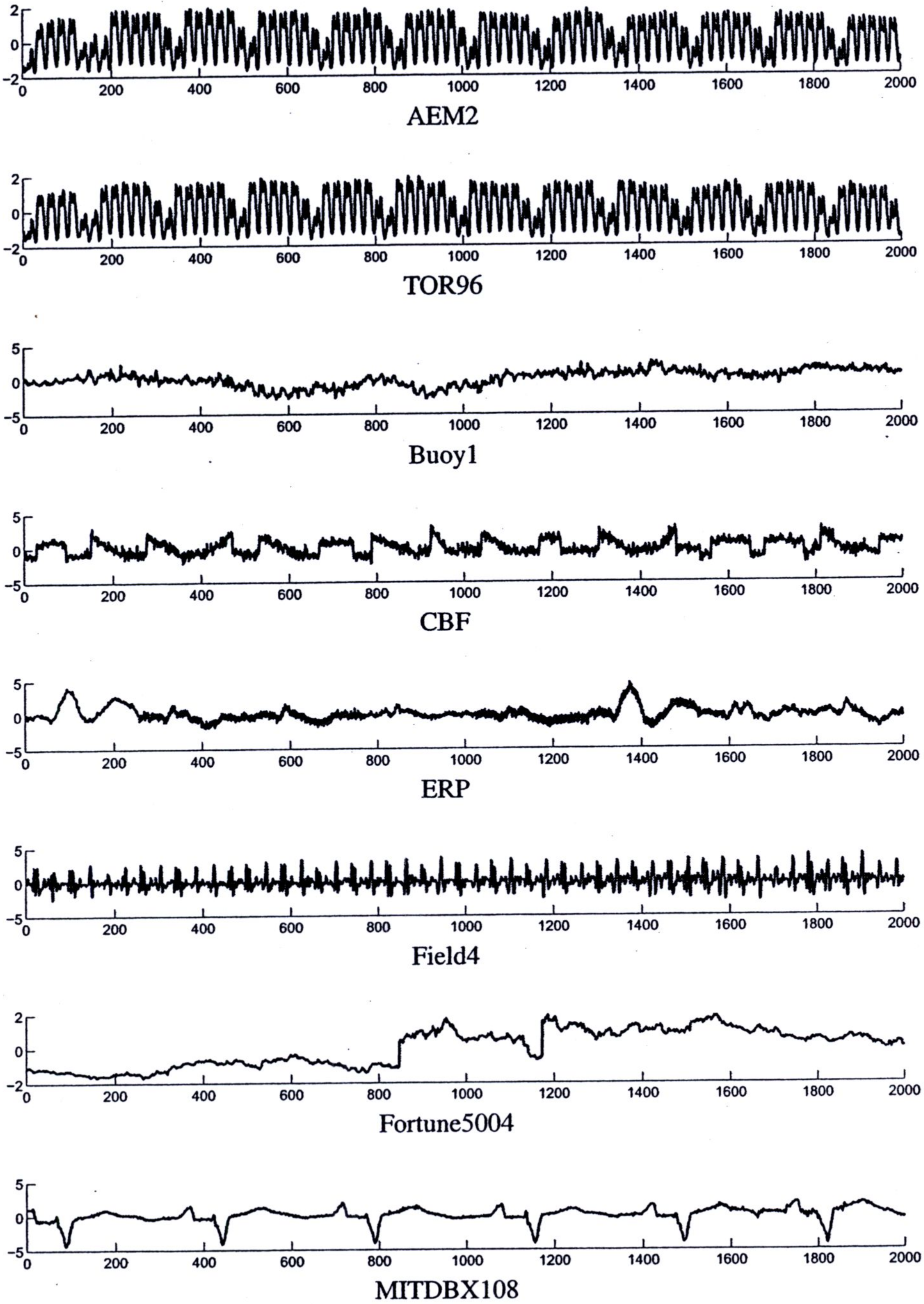


Figure A.1: Datasets from TSDMA used in the experiments of Chapters II, IV, and VI.

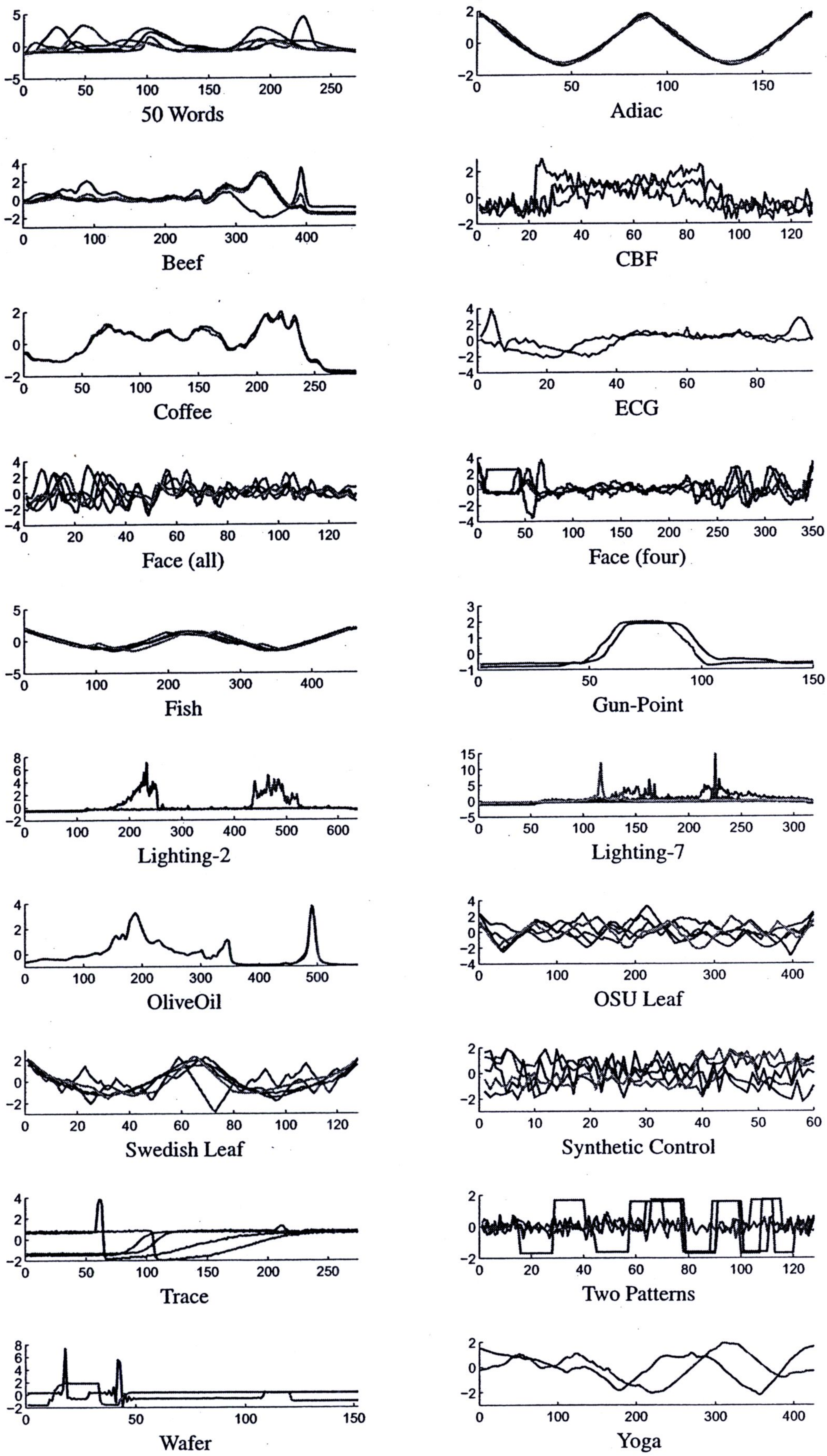


Figure A.2: Examples of some classes of the UCR classification/clustering datasets used in Chapters III and V.

Table A.1: Details of the UCR classification/clustering datasets used in Chapters III and V

Dataset	Number of classes	Length	Size of datasets
50words	50	270	905
Adiac	37	176	781
Beef	5	470	60
CBF	3	128	930
Coffee	2	286	56
ECG	2	96	200
Face (all)	14	131	2250
Face (four)	4	350	112
Fish	7	463	350
Gun-Point	2	150	200
Lighting-2	2	637	121
Lighting-7	7	319	143
Oliveoil	4	570	60
OSULeaf	6	427	442
SwedishLeaf	15	128	1125
Synthetic	6	60	600
Trace	4	275	200
TwoPatterns	4	128	5000
Wafer	2	152	7174
Yoga	2	426	3300

APPENDIX B

COMPLETE EXPERIMENTAL RESULTS OF THE FIRST EXPERIMENT IN CHAPTER II

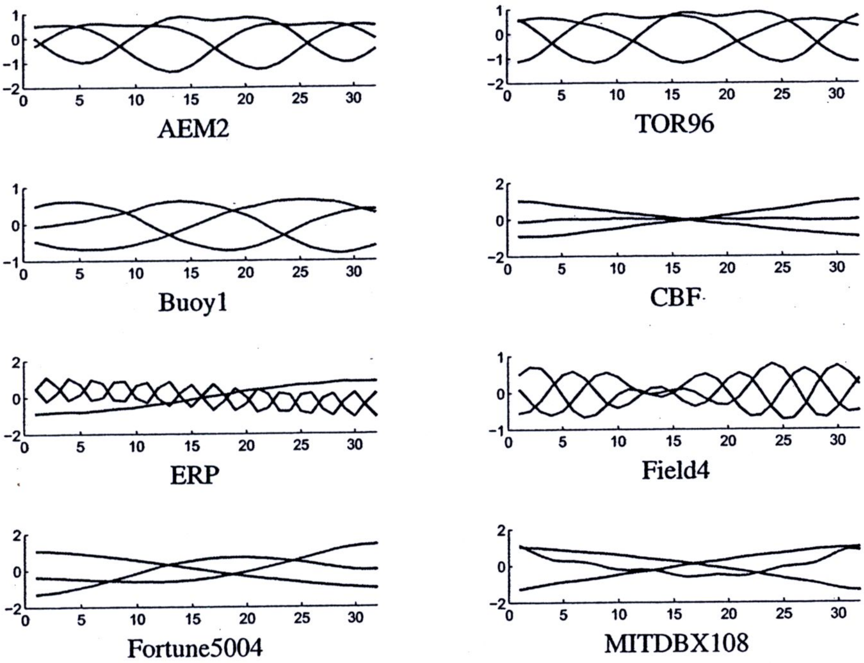


Figure B.1: Cluster representatives generated from STSC using k-means clustering when $k = 3$ and $w = 32$.

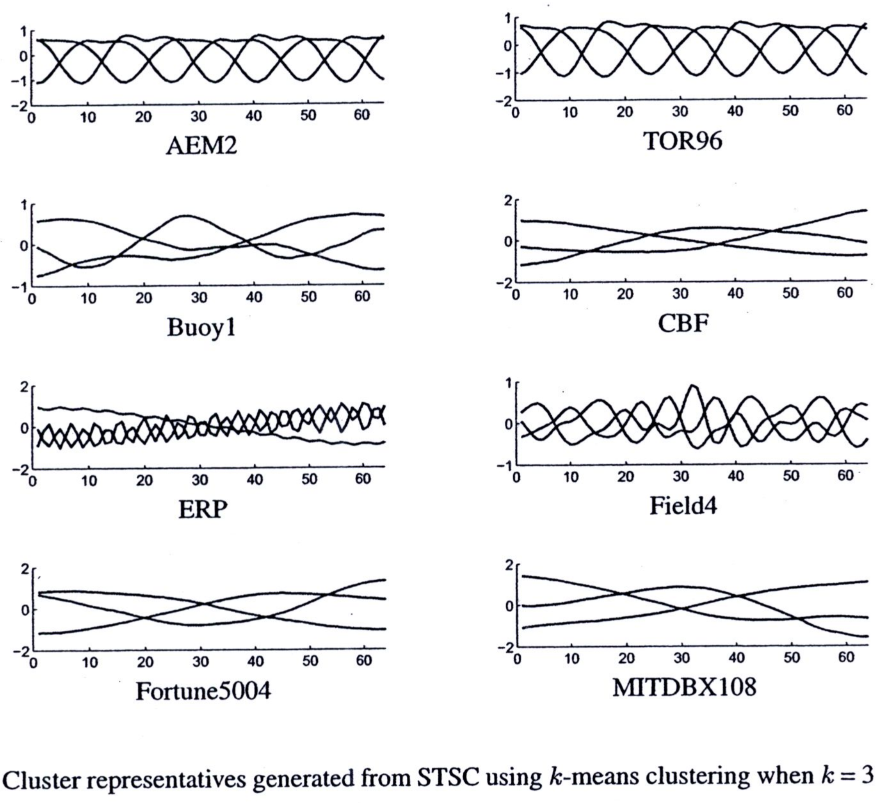


Figure B.2: Cluster representatives generated from STSC using k-means clustering when $k = 3$ and $w = 64$.

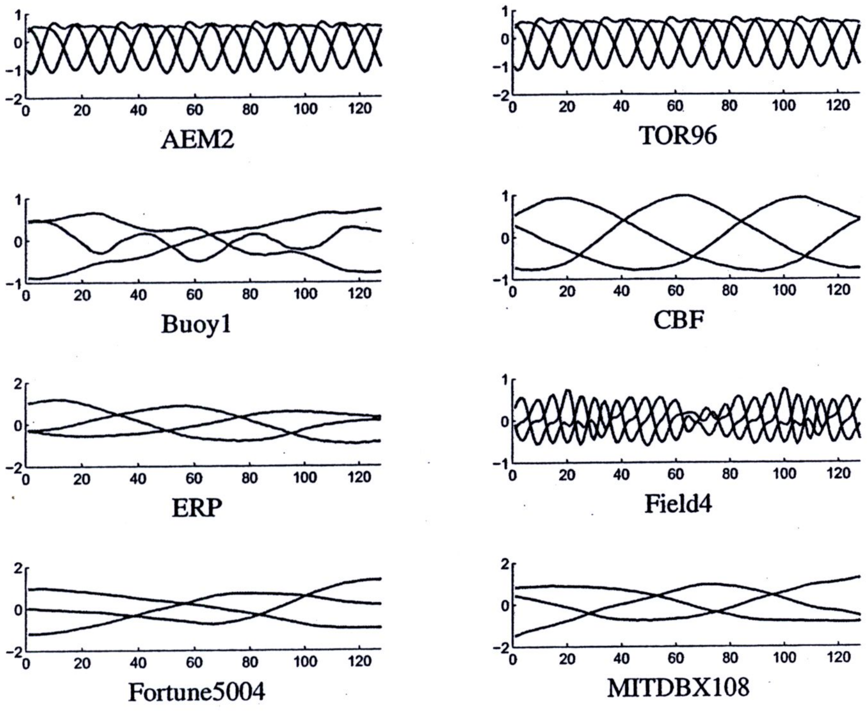


Figure B.3: Cluster representatives generated from STSC using k-means clustering when $k = 3$ and $w = 128$.

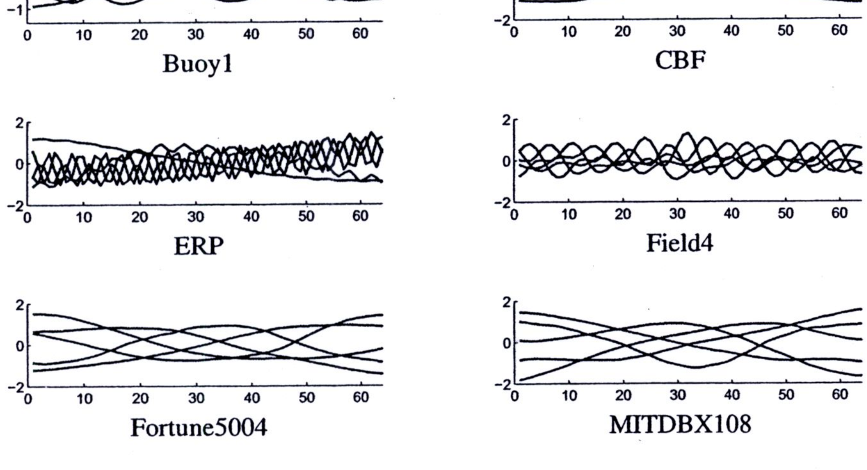


Figure B.4: Cluster representatives generated from STSC using k-means clustering when $k = 5$ and $w = 64$.

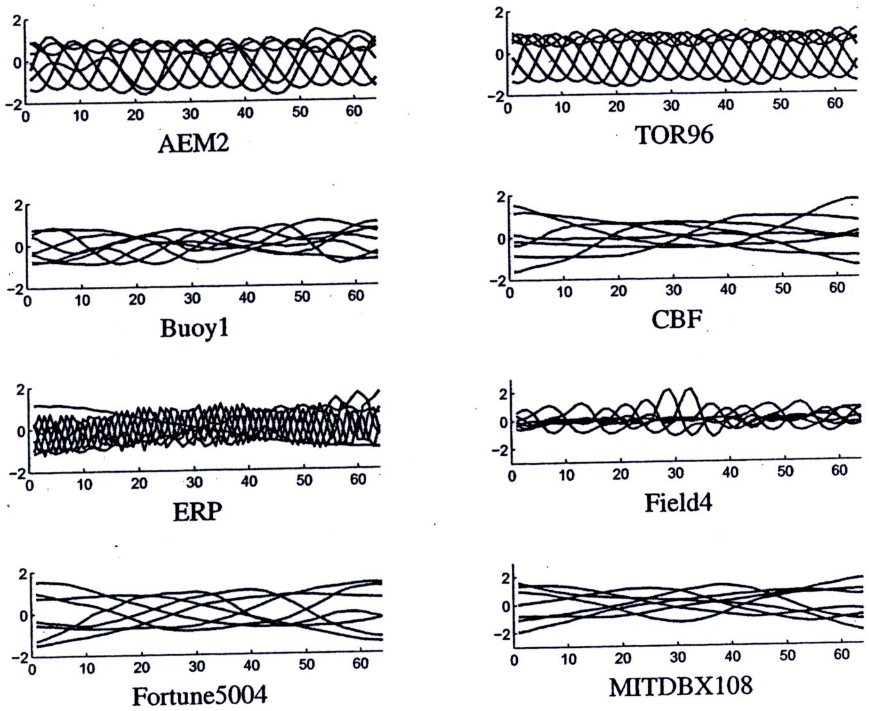


Figure B.5: Cluster representatives generated from STSC using k-means clustering when $k = 7$ and $w = 64$.

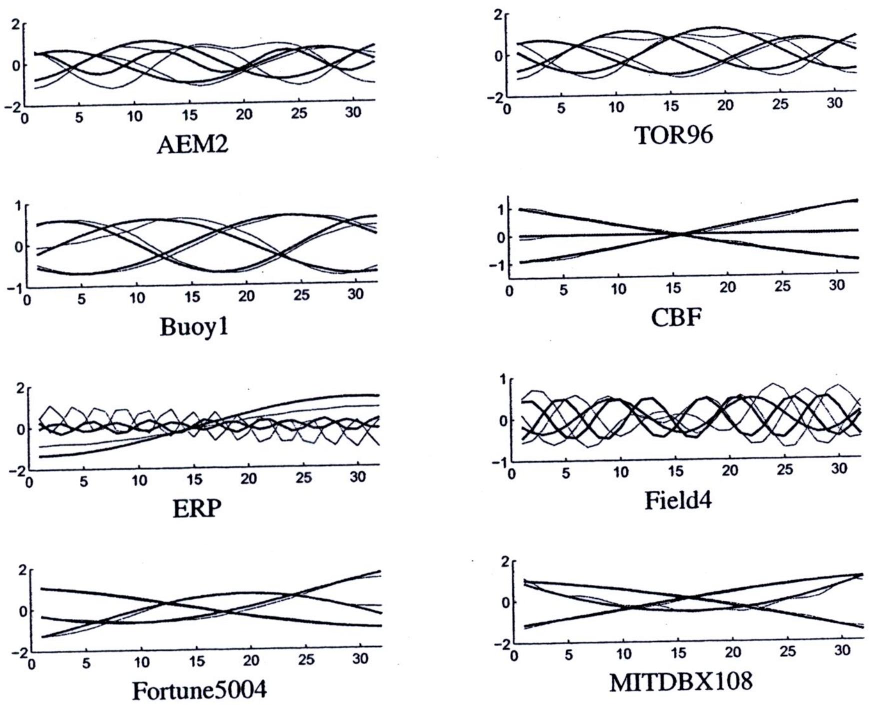


Figure B.6: Constructed sine waves generated from STSC using k-means clustering when $k = 3$ and $w = 32$.

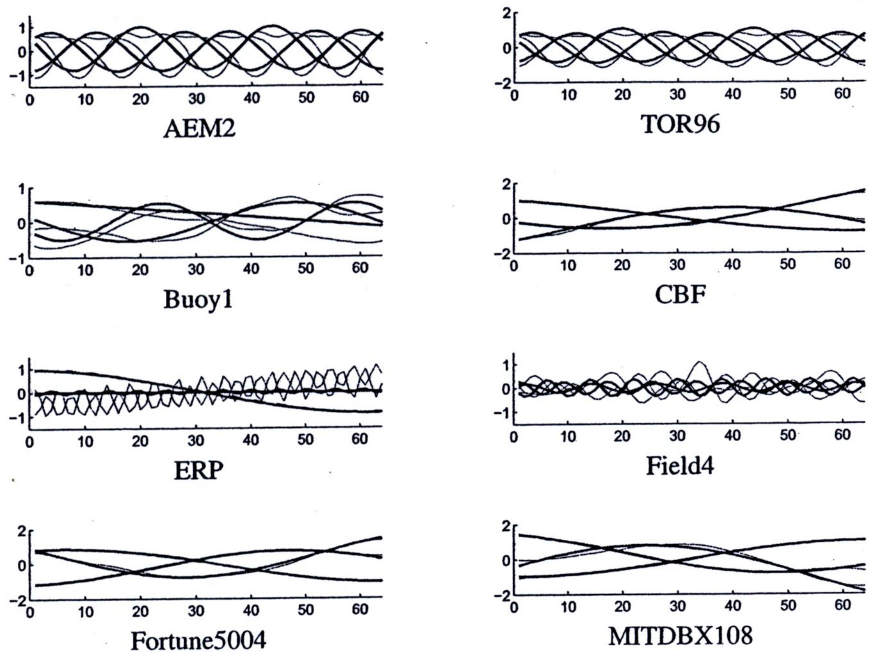


Figure B.7: Constructed sine waves generated from STSC using k-means clustering when k = 3 and w = 64.

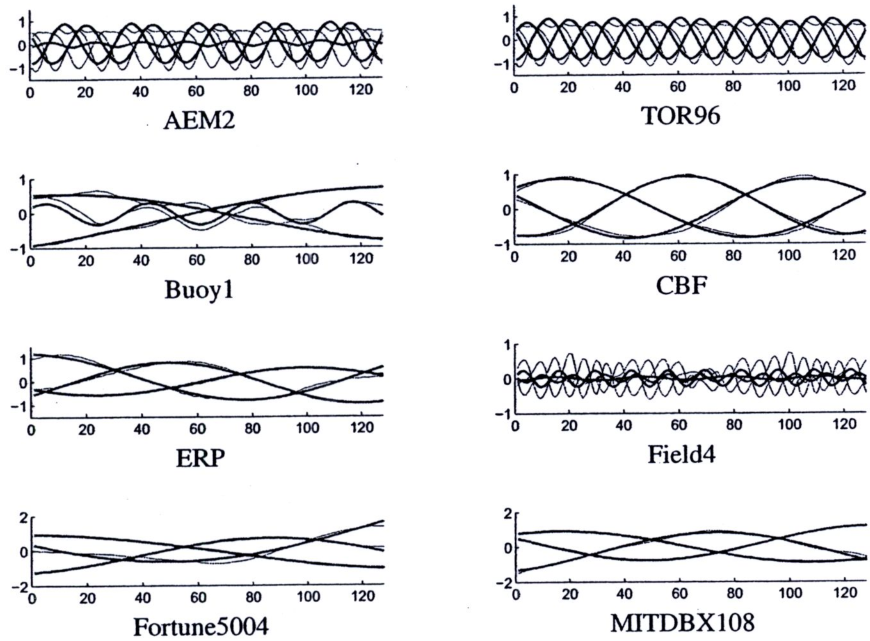


Figure B.8: Constructed sine waves generated from STSC using k-means clustering when k = 3 and w = 128.

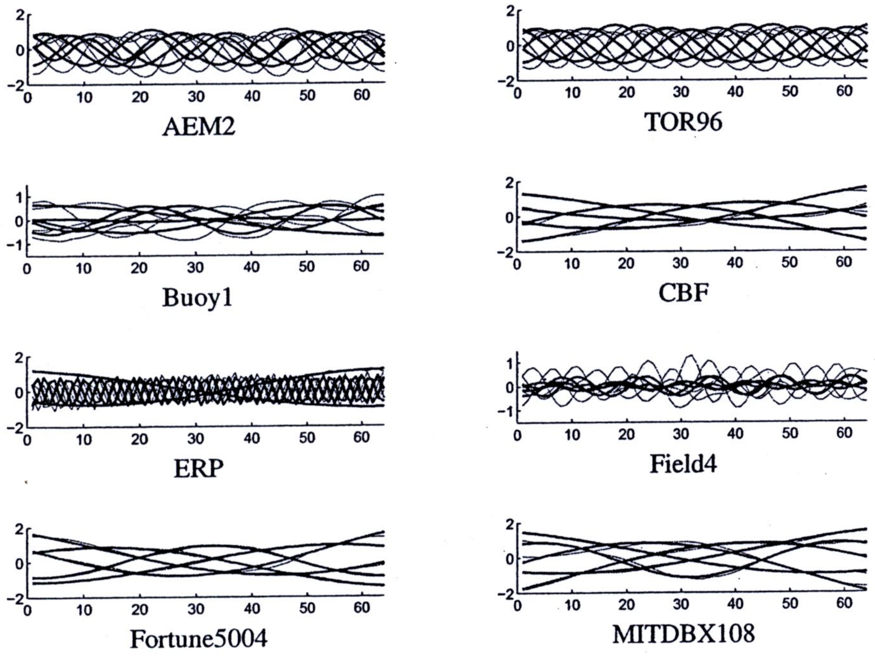


Figure B.9: Constructed sine waves generated from STSC using k-means clustering when k = 5 and w = 64.

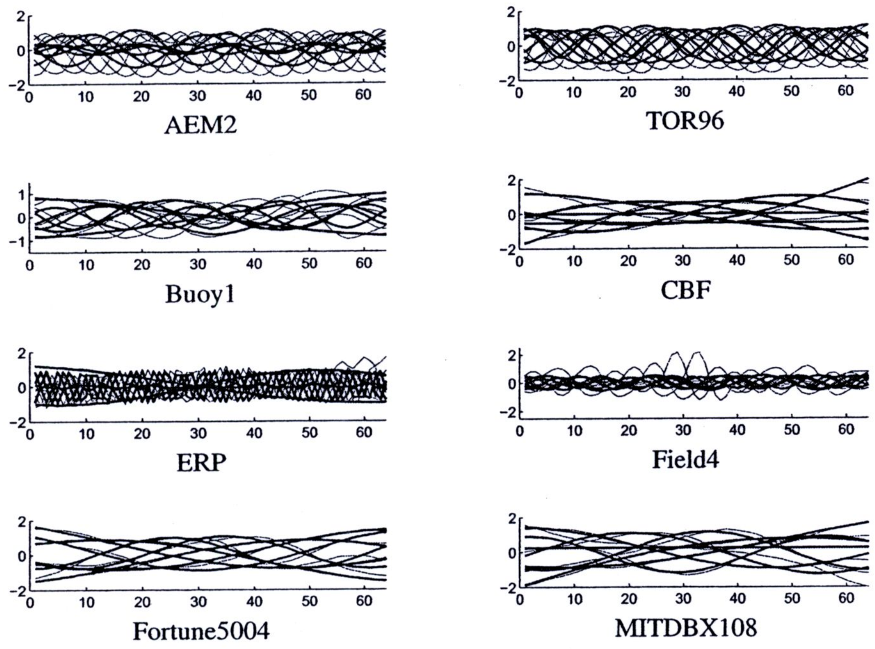


Figure B.10: Constructed sine waves generated from STSC using k-means clustering when k = 7 and w = 64.

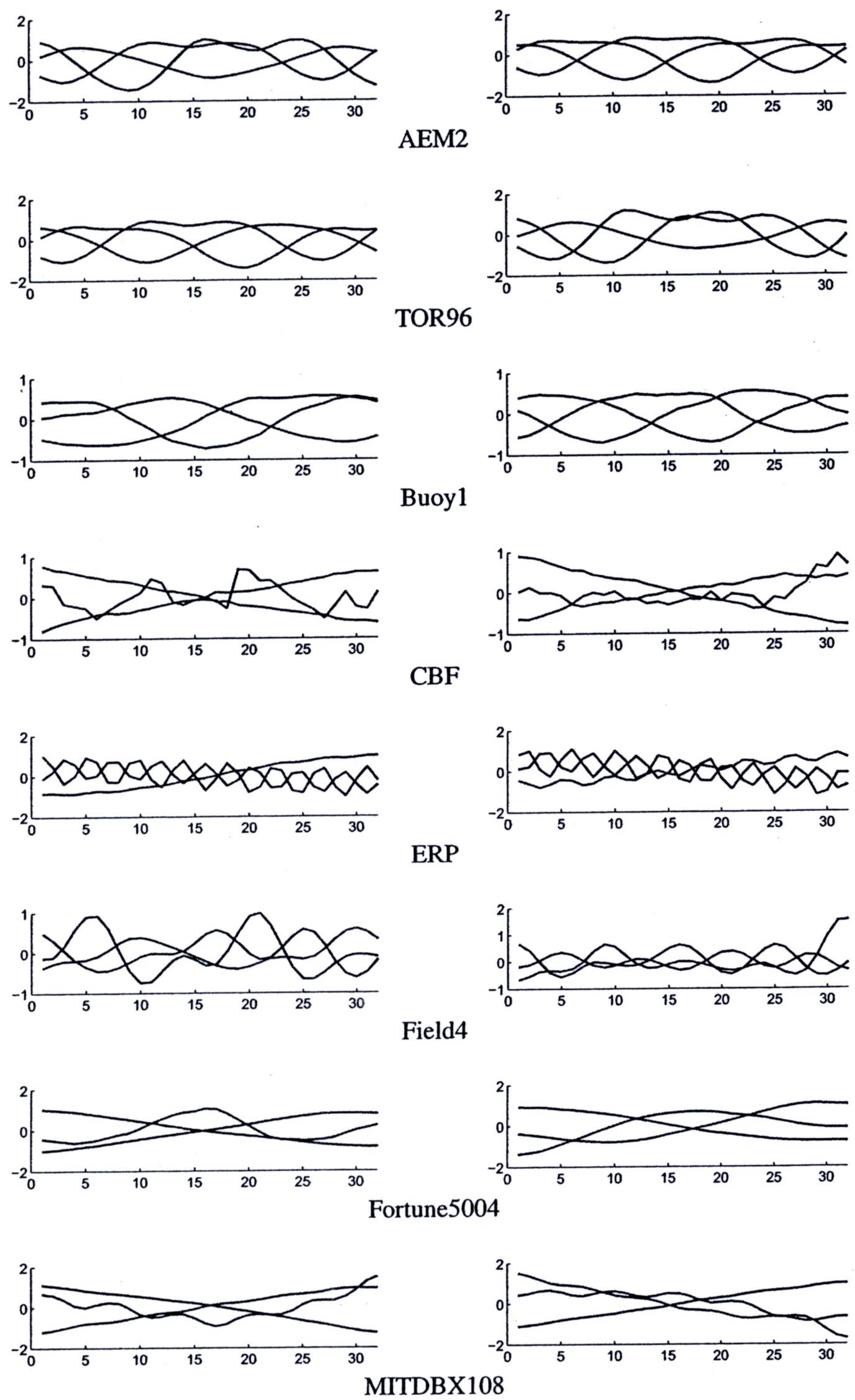


Figure B.11: Cluster representatives generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 3$ and $w = 32$.

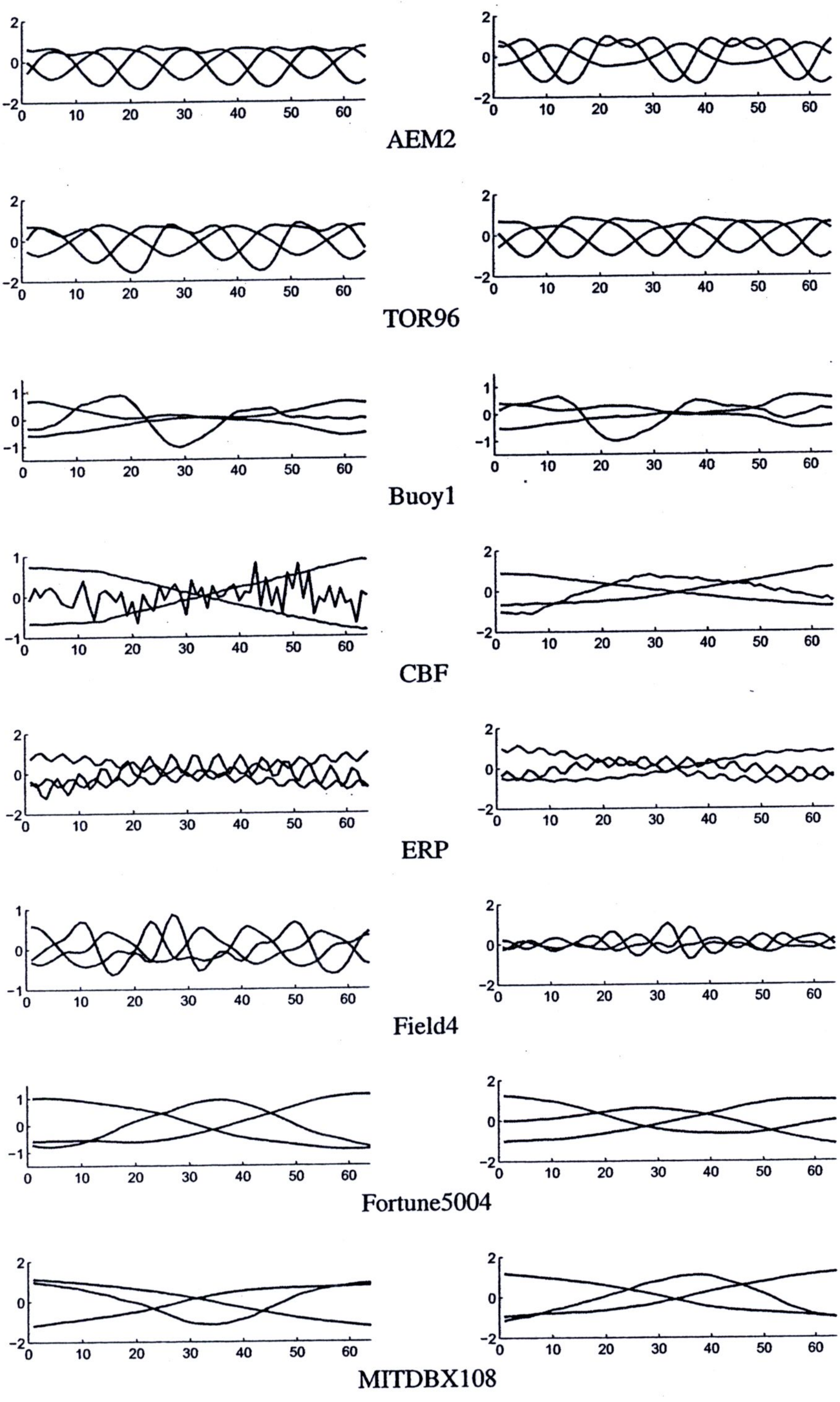


Figure B.12: Cluster representatives generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 3$ and $w = 64$.

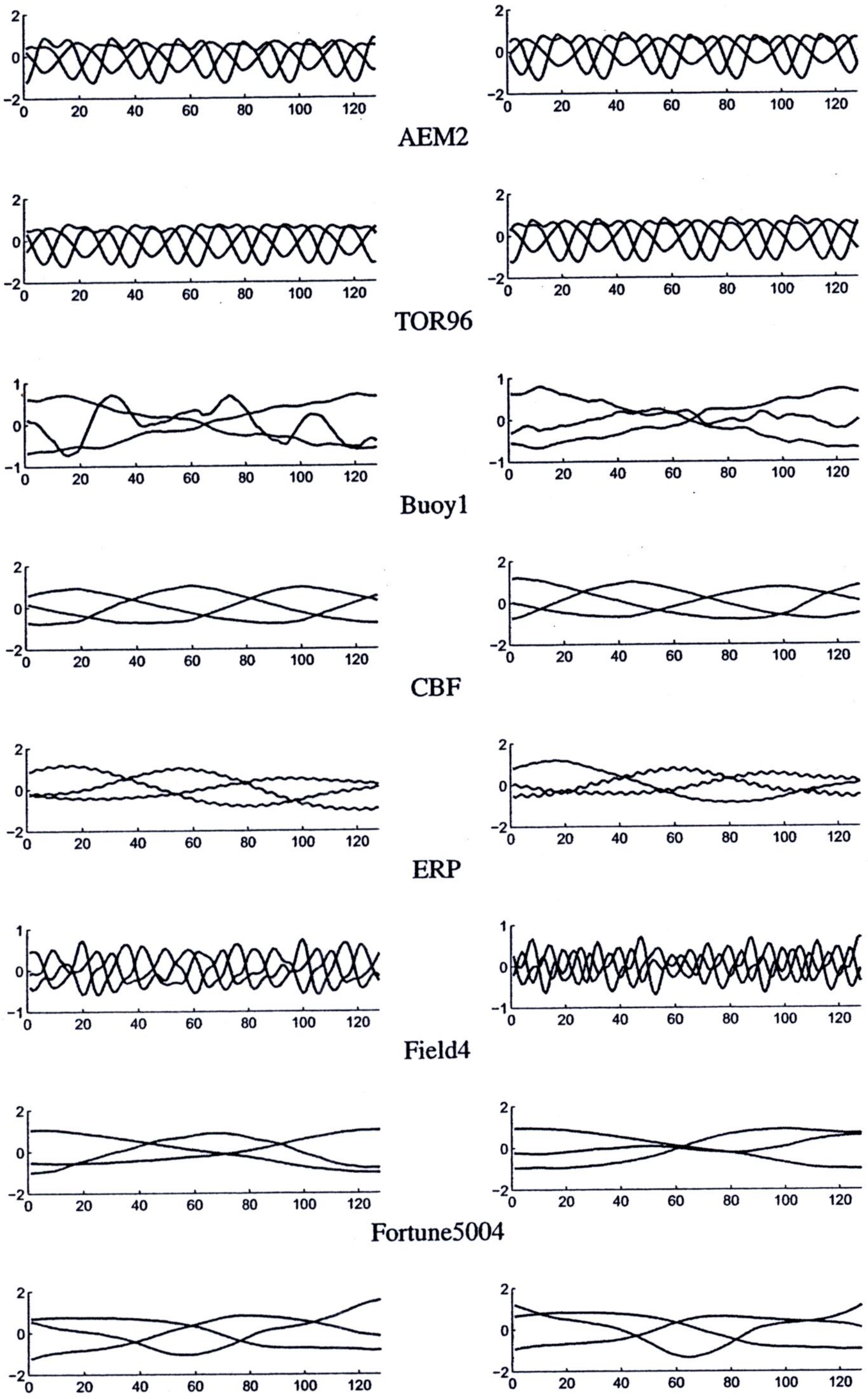


Figure B.13: Cluster representatives generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 3$ and $w = 128$.

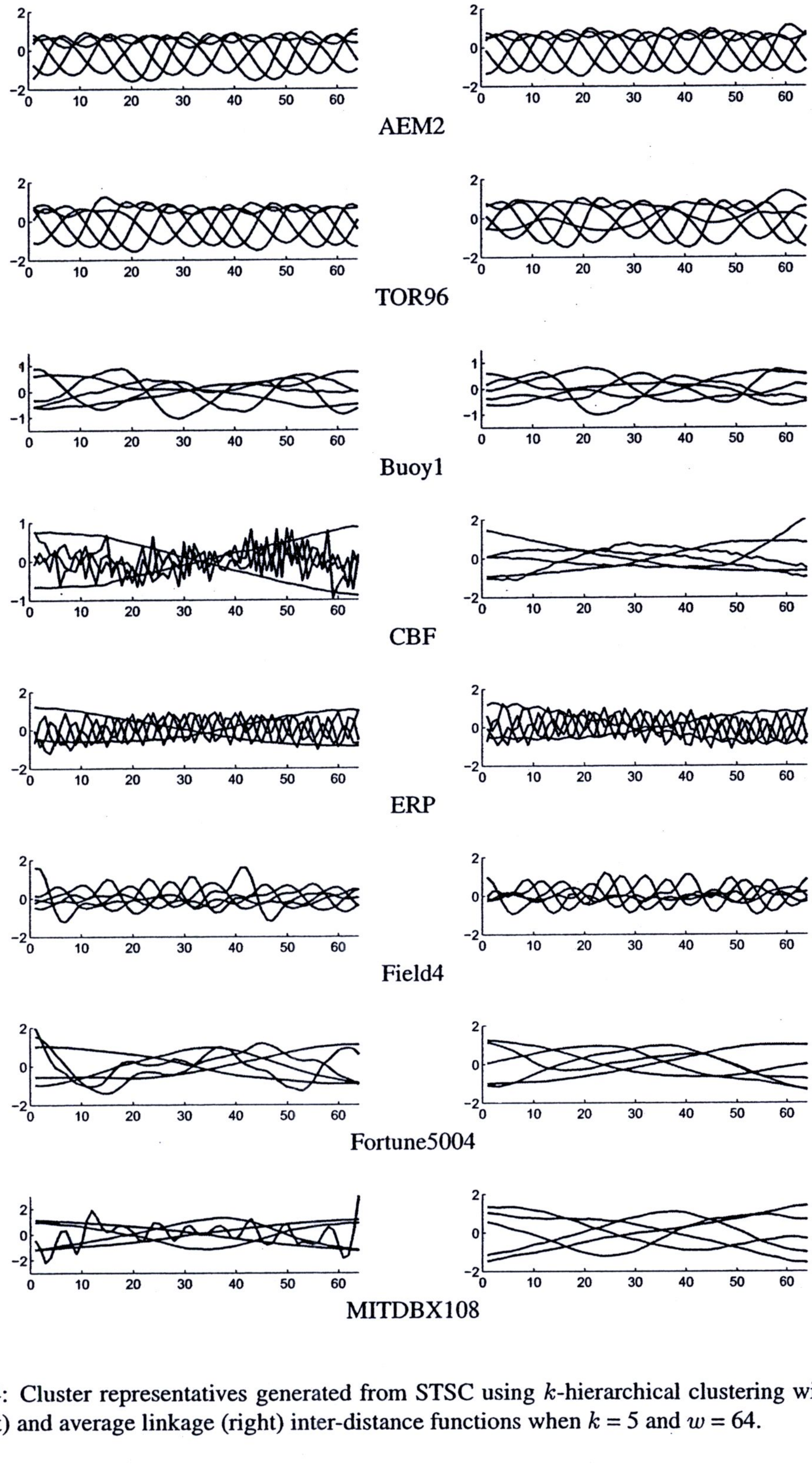


Figure B.14: Cluster representatives generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 5$ and $w = 64$.

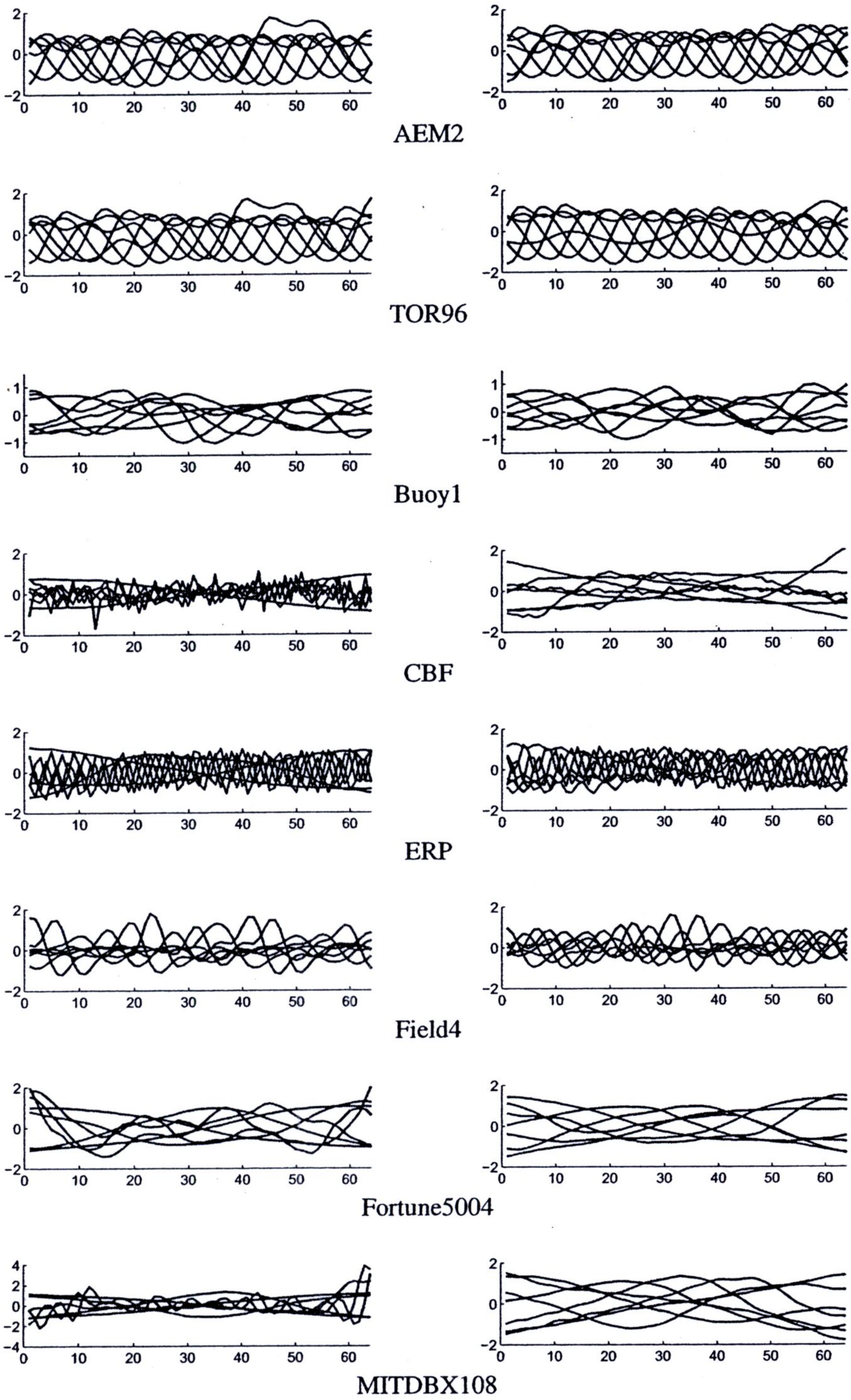


Figure B.15: Cluster representatives generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 7$ and $w = 64$.

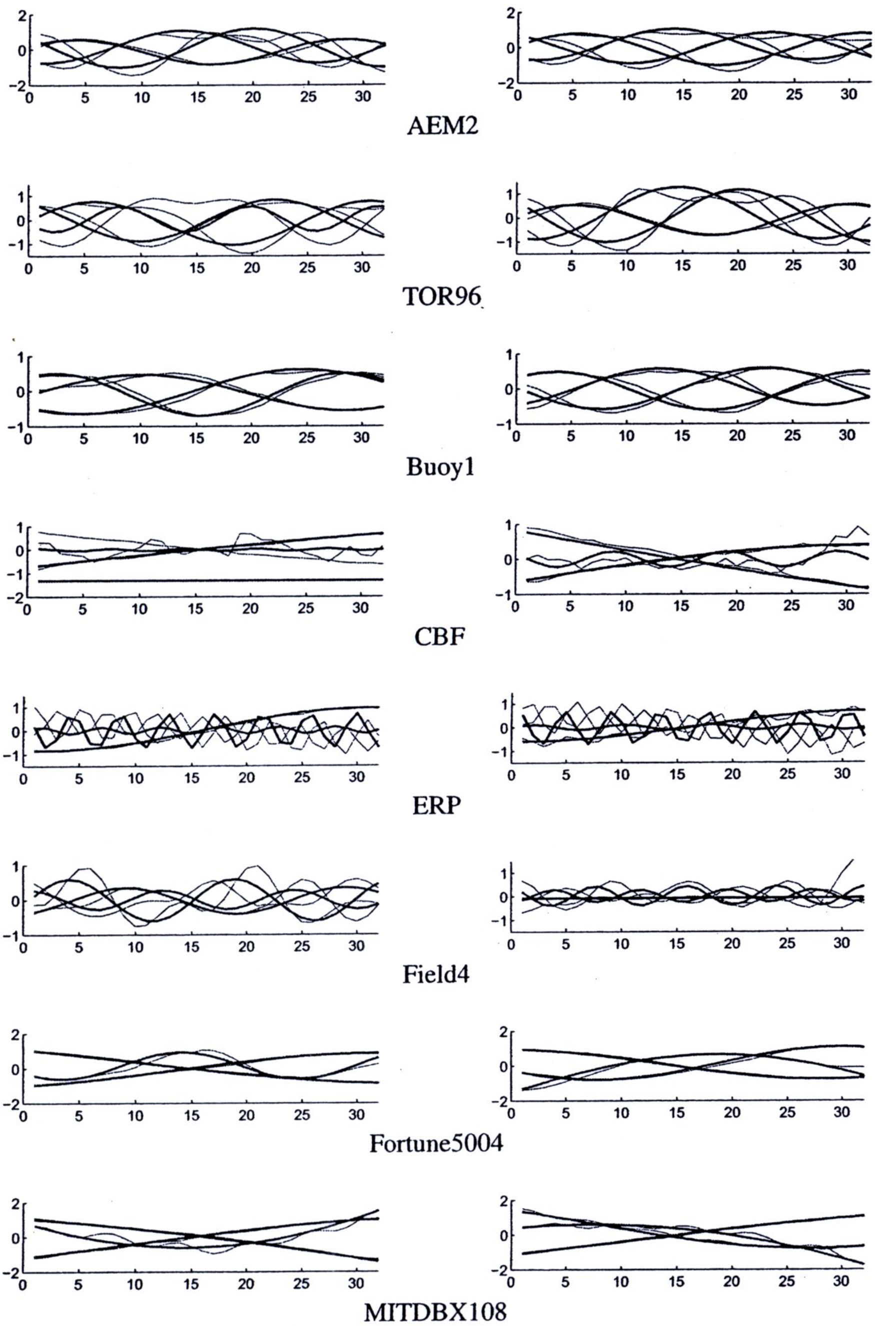


Figure B.16: Constructed sine waves generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 3$ and $w = 32$.

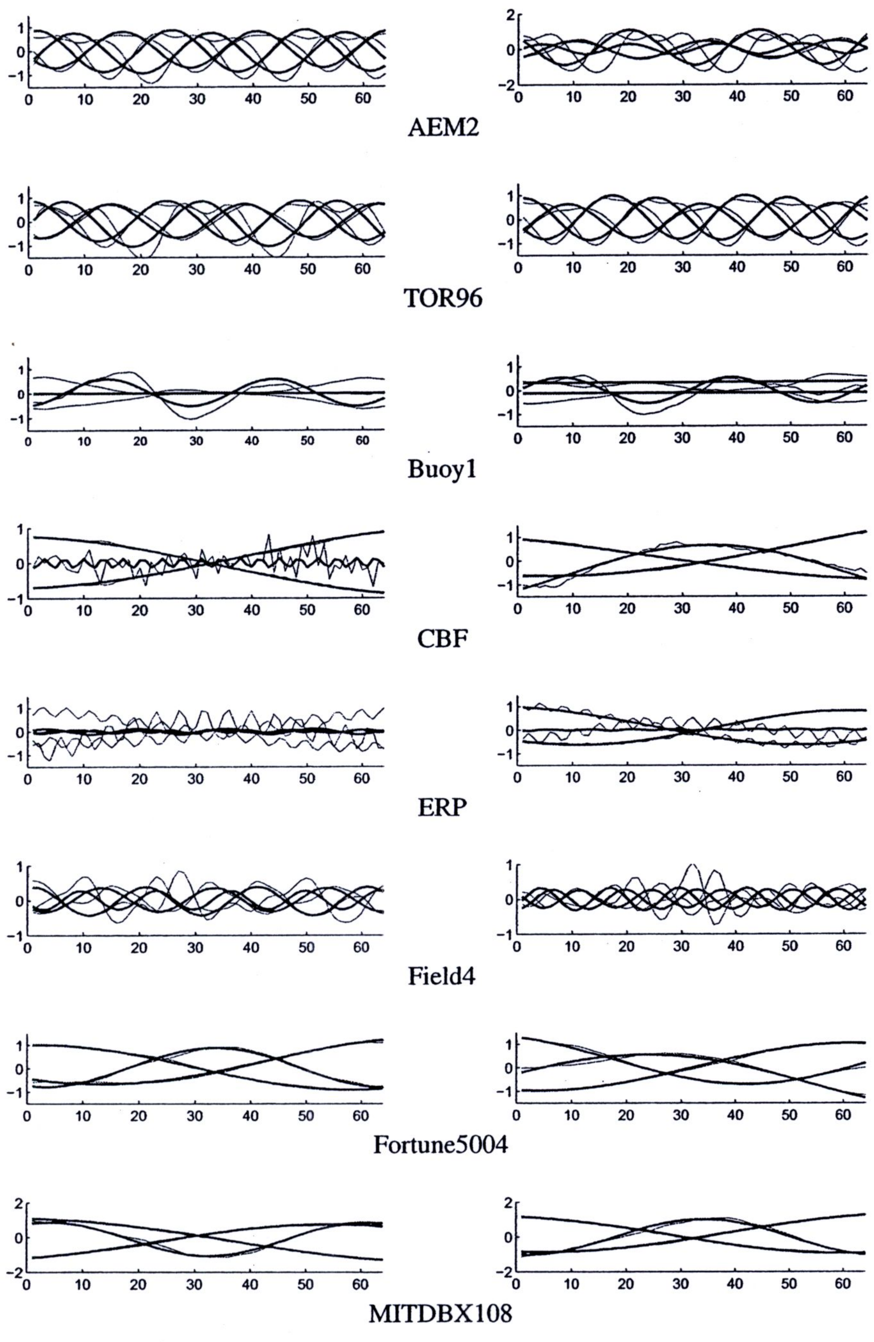


Figure B.17: Constructed sine waves generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 3$ and $w = 64$.

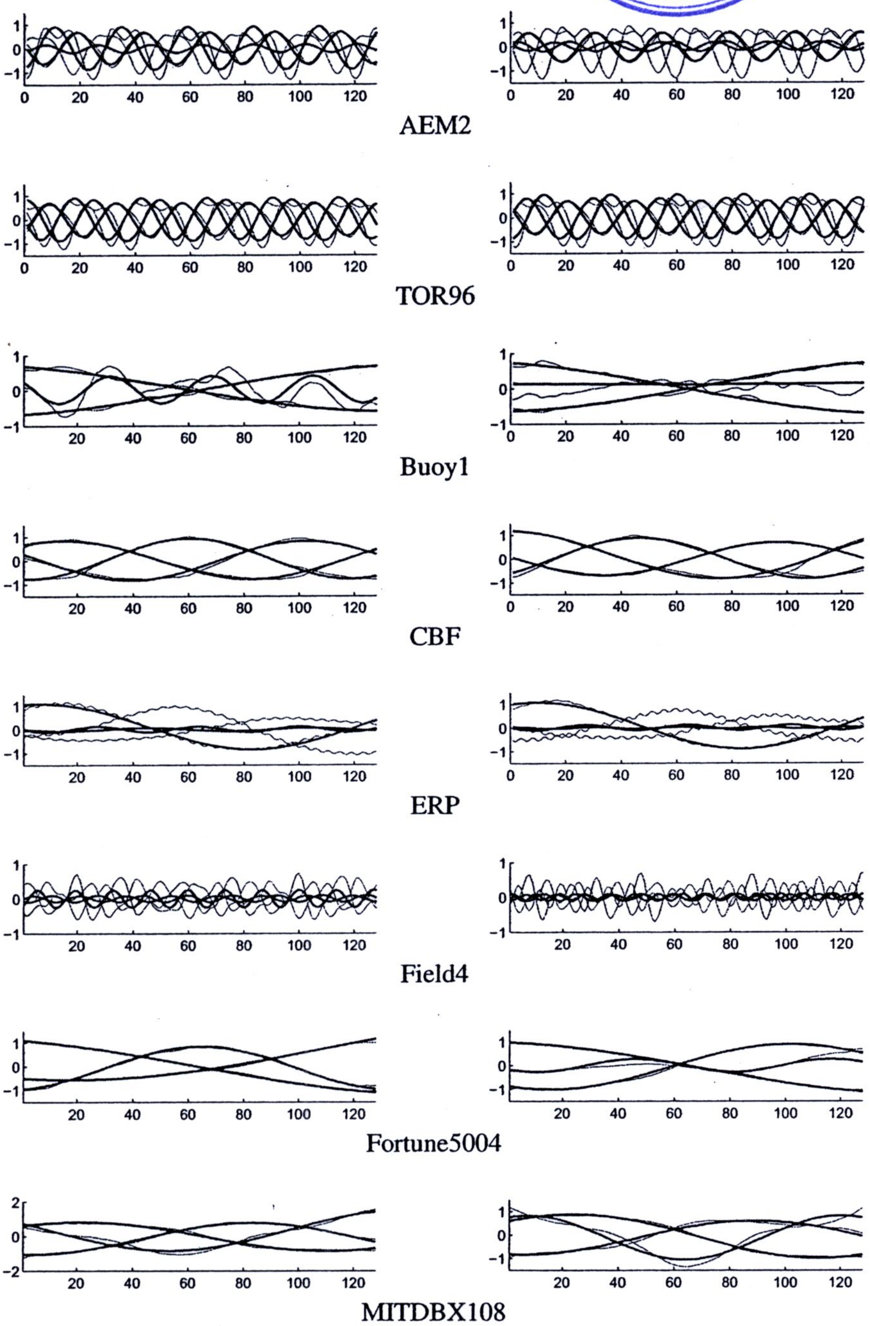


Figure B.18: Constructed sine waves generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 3$ and $w = 128$.

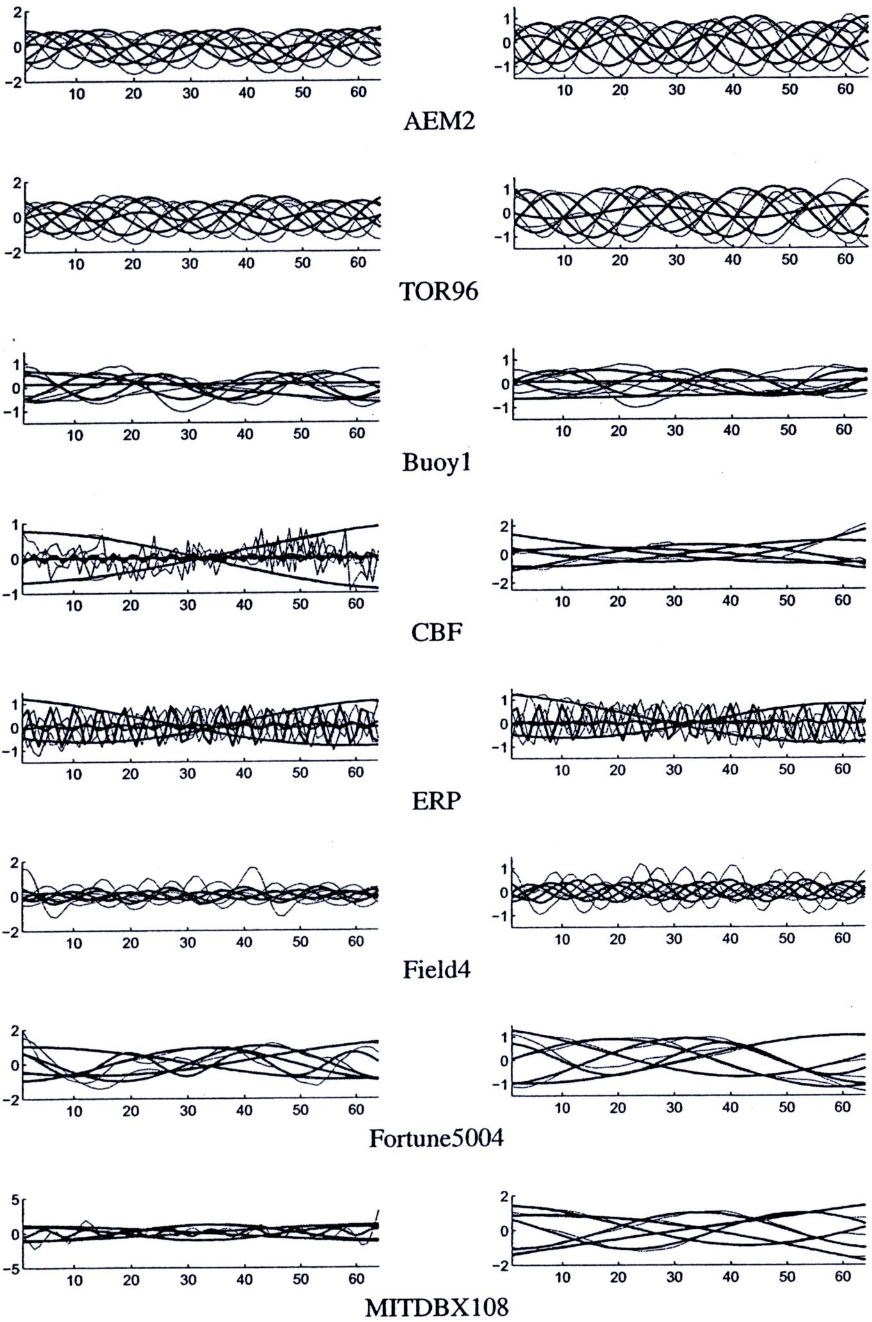


Figure B.19: Constructed sine waves generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 5$ and $w = 64$.

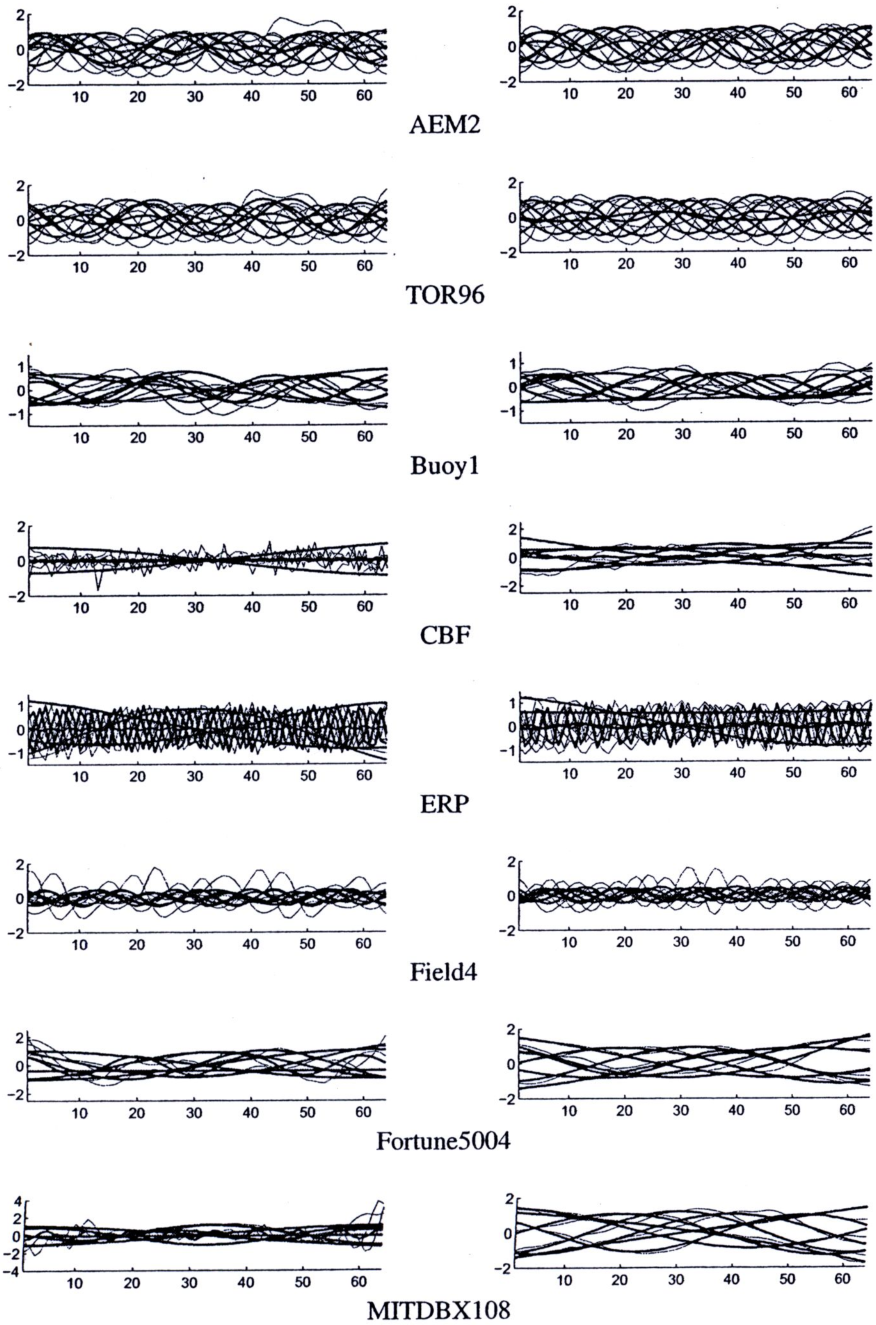


Figure B.20: Constructed sine waves generated from STSC using k-hierarchical clustering with complete linkage (left) and average linkage (right) inter-distance functions when $k = 7$ and $w = 64$.

APPENDIX C

COMPLETE EXPERIMENTAL RESULTS OF THE EXPERIMENT IN CHAPTER III

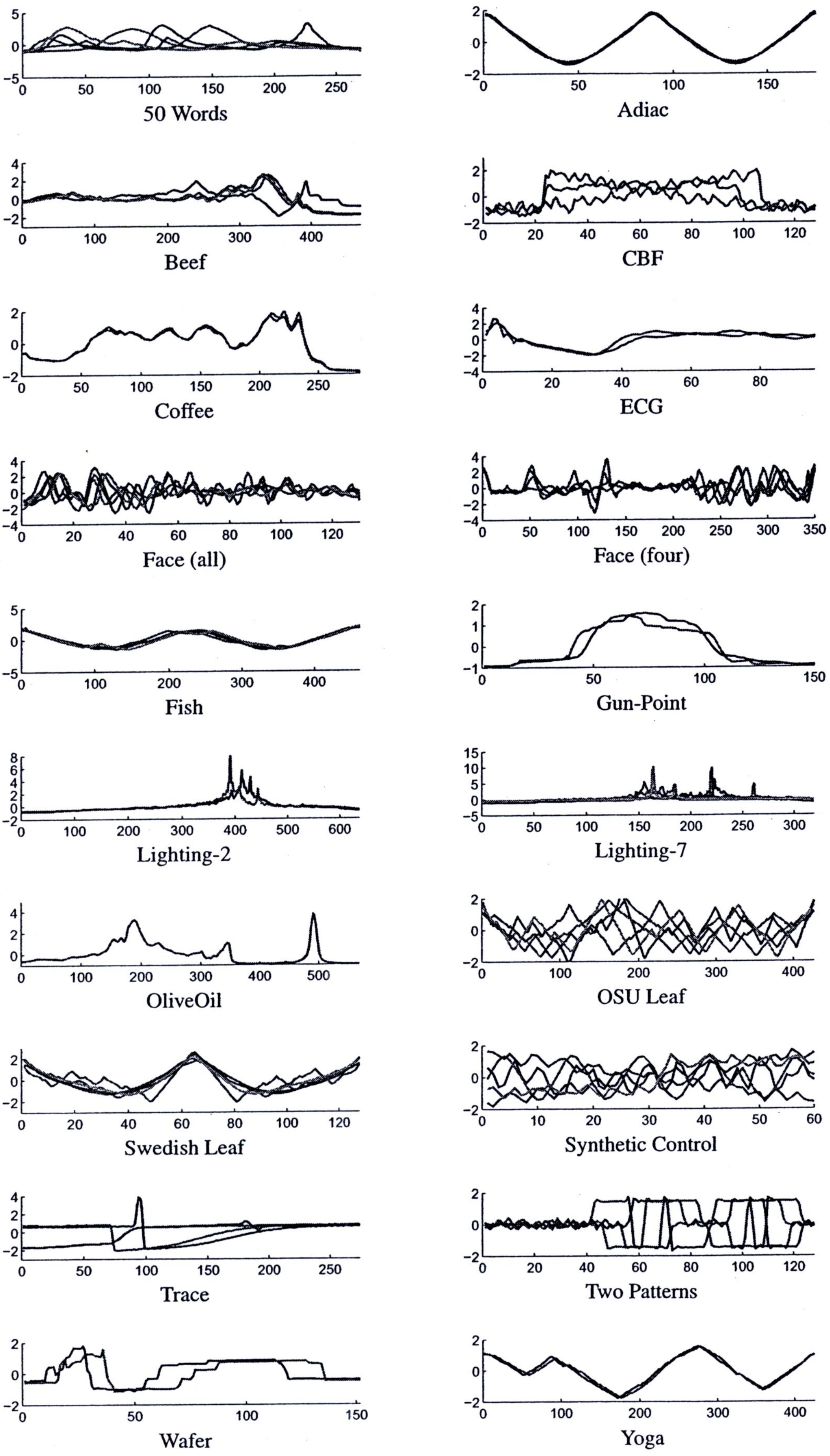


Figure C.1: Averaged results generated from CDTW function of each dataset

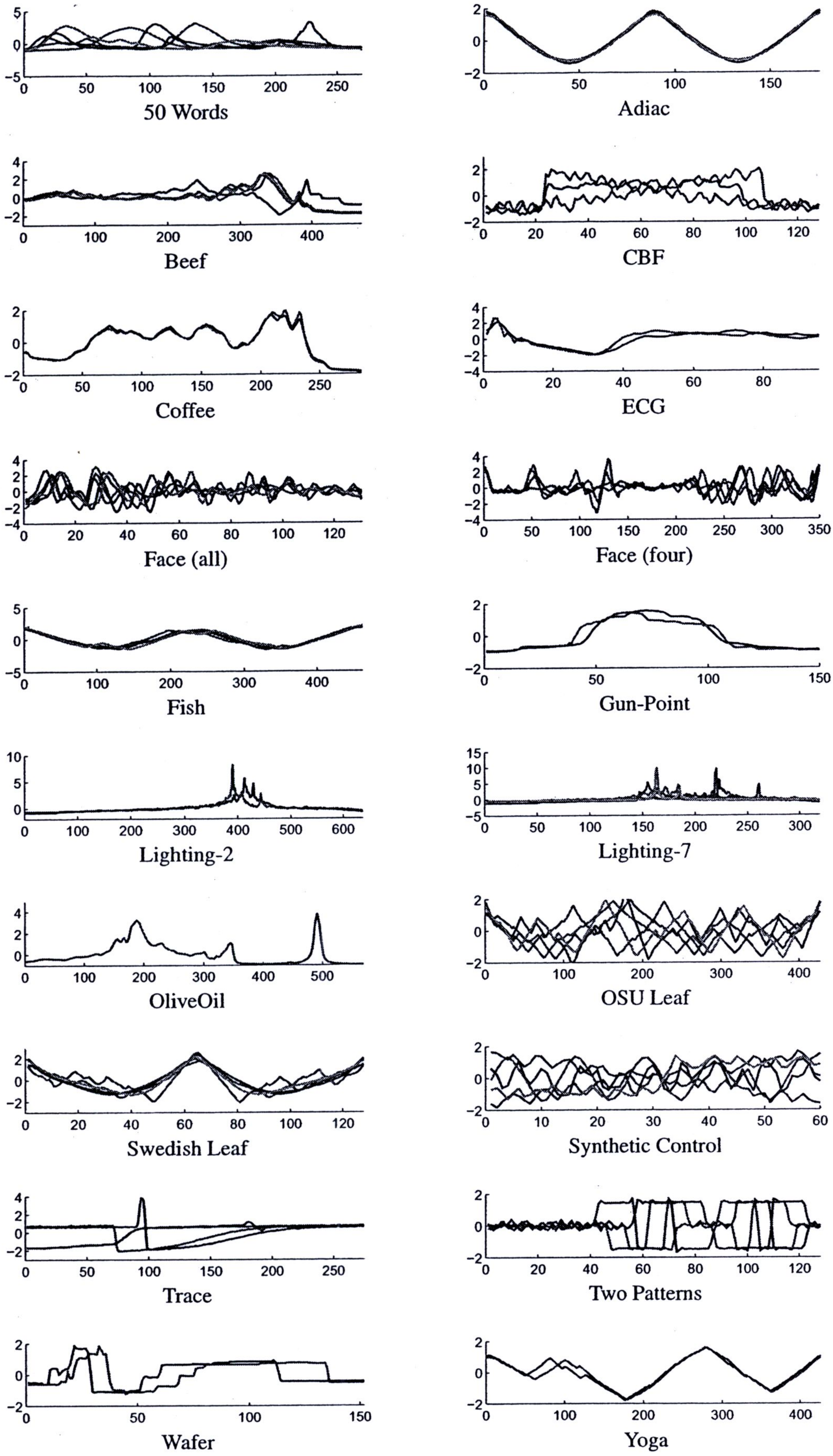


Figure C.2: Averaged results generated from ICDTW function of each dataset

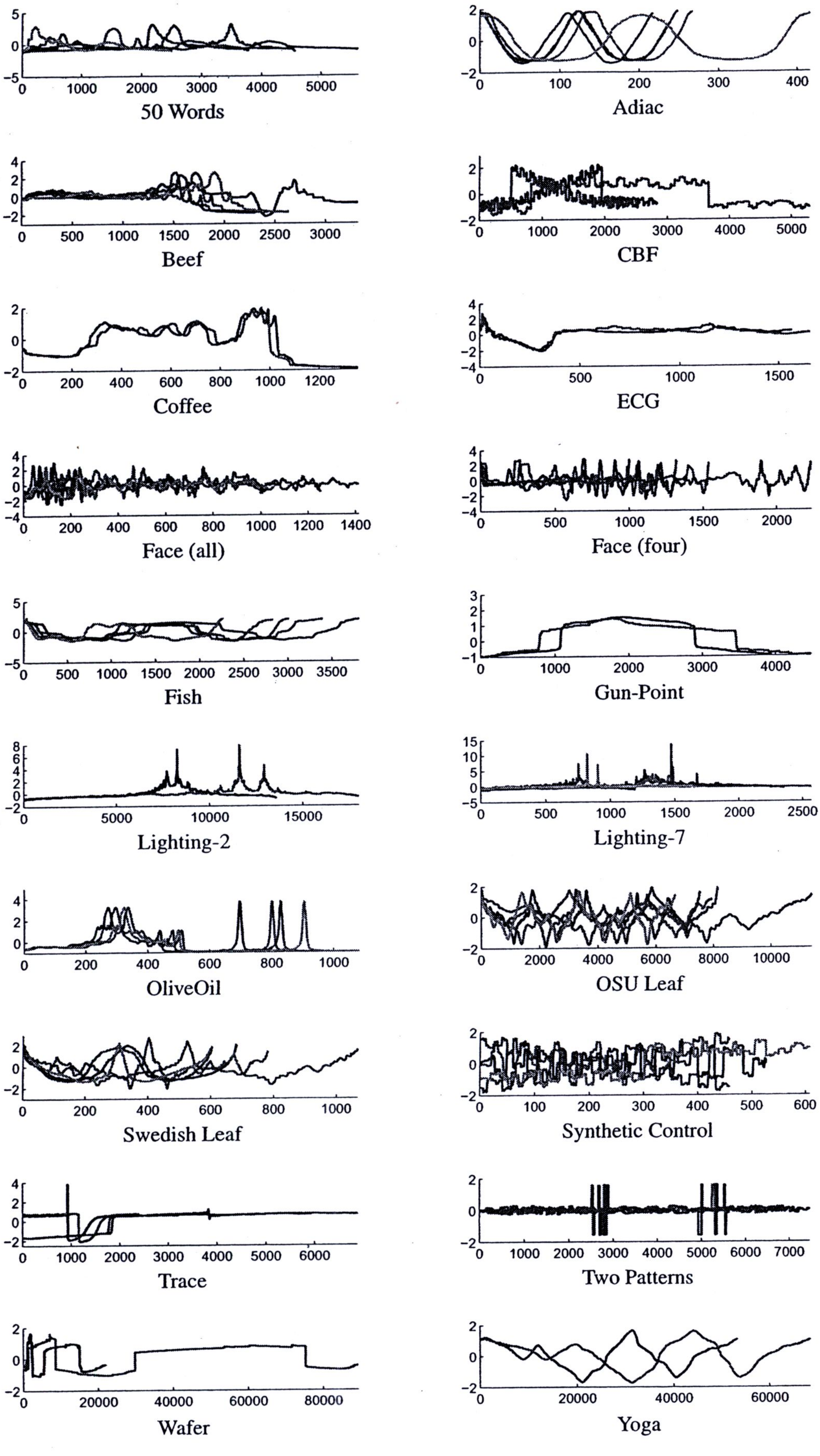


Figure C.3: Averaged results generated from NLAAF of each dataset.

APPENDIX D

COMPLETE EXPERIMENTAL RESULTS OF THE EXPERIMENT IN CHAPTER IV

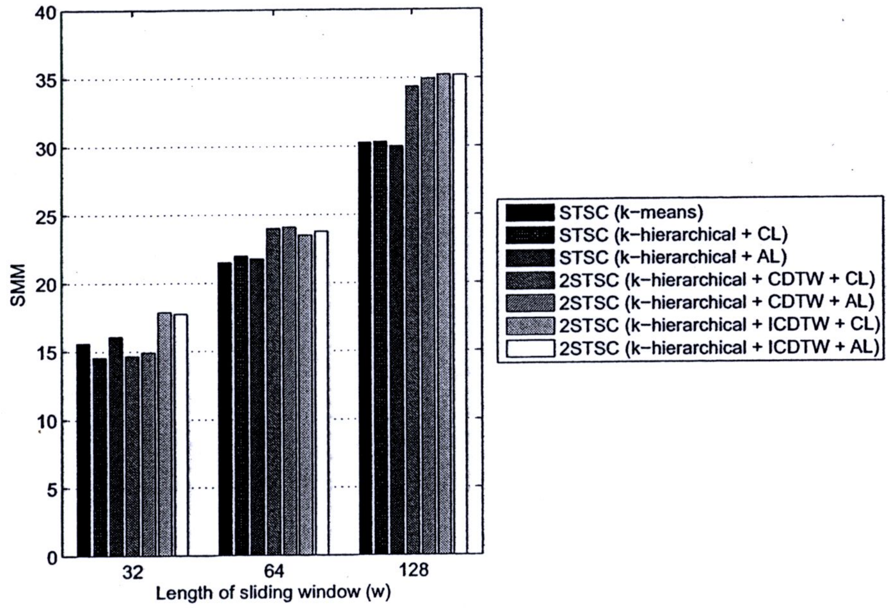


Figure D.1: SMMs of AEM2 when the number of clusters (k) is 3 and the length of sliding window (w) is varied.

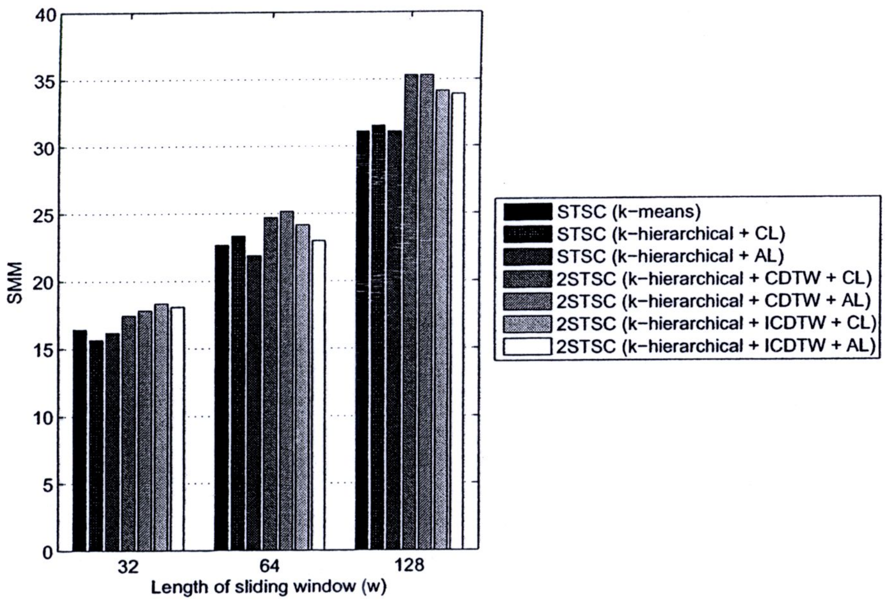


Figure D.2: SMMs of TOR96 when the number of clusters (k) is 3 and the length of sliding window (w) is varied.

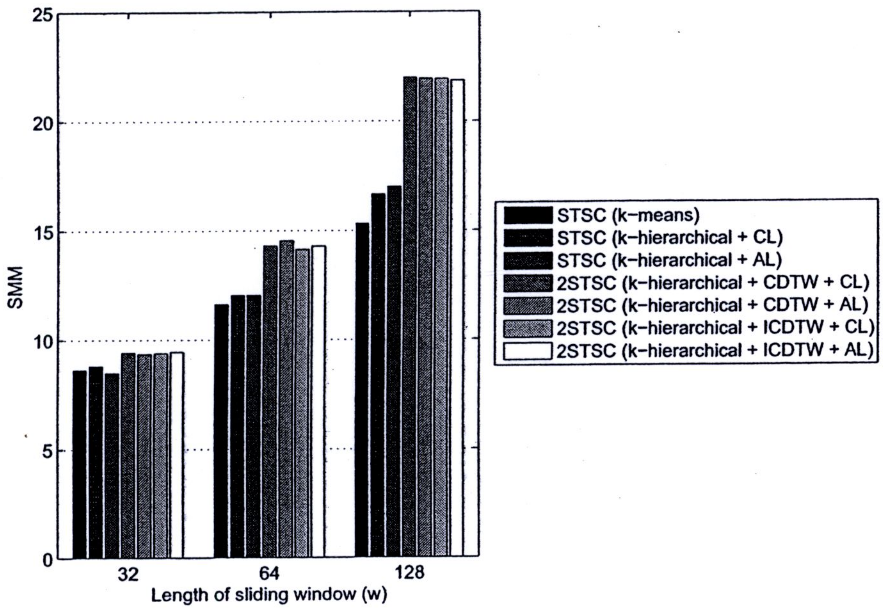


Figure D.3: SMMs of Buoy1 when the number of clusters (k) is 3 and the length of sliding window (w) is varied.

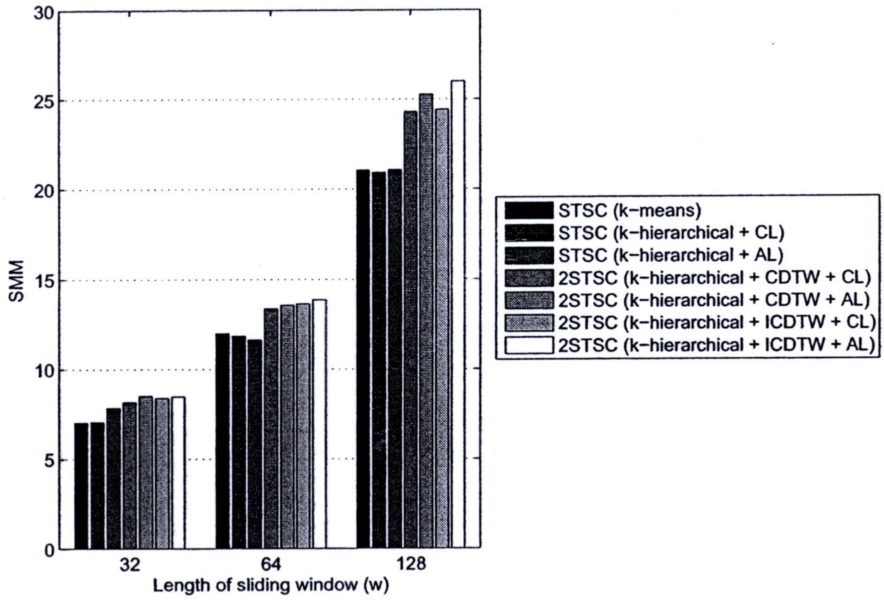


Figure D.4: SMMs of CBF when the number of clusters (k) is 3 and the length of sliding window (w) is varied.

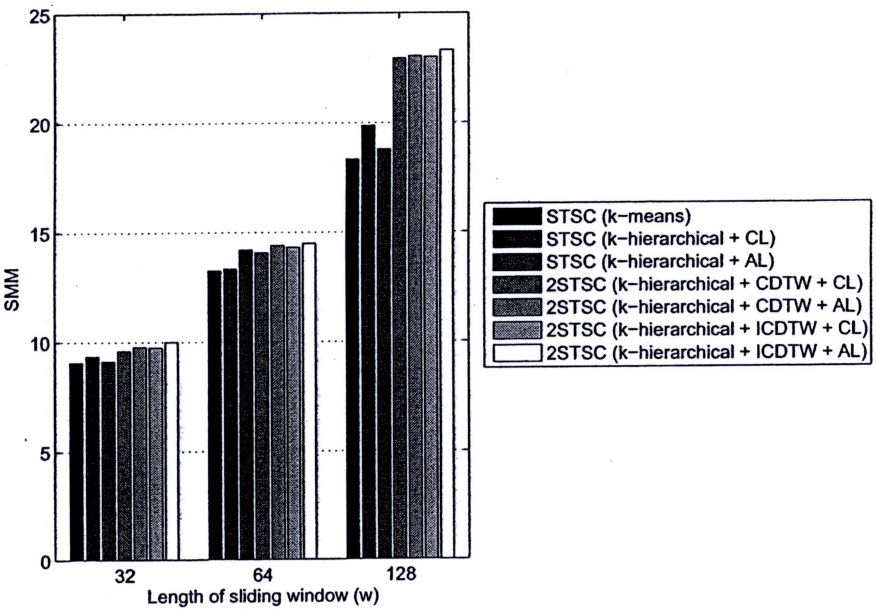


Figure D.5: SMMs of ERP when the number of clusters (k) is 3 and the length of sliding window (w) is varied.

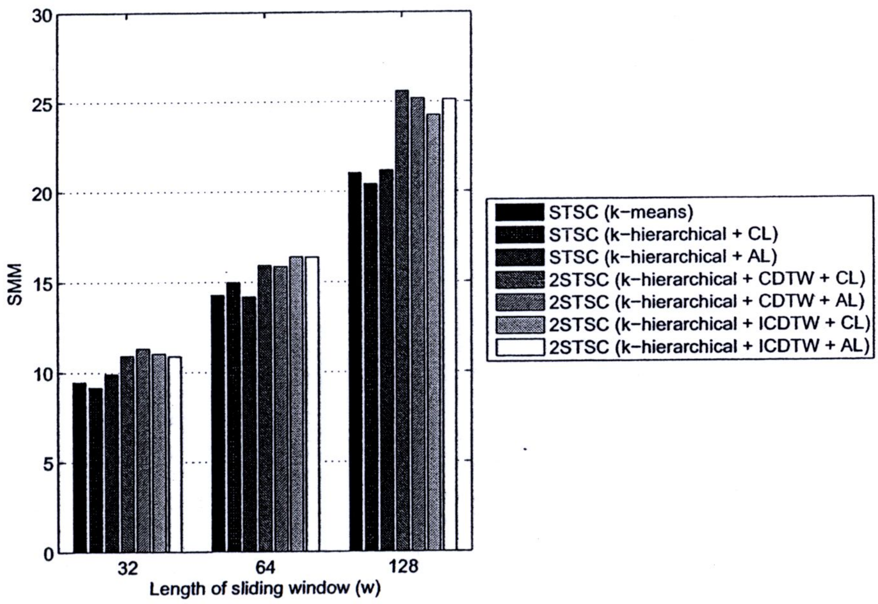


Figure D.6: SMMs of Field4 when the number of clusters (k) is 3 and the length of sliding window (w) is varied.

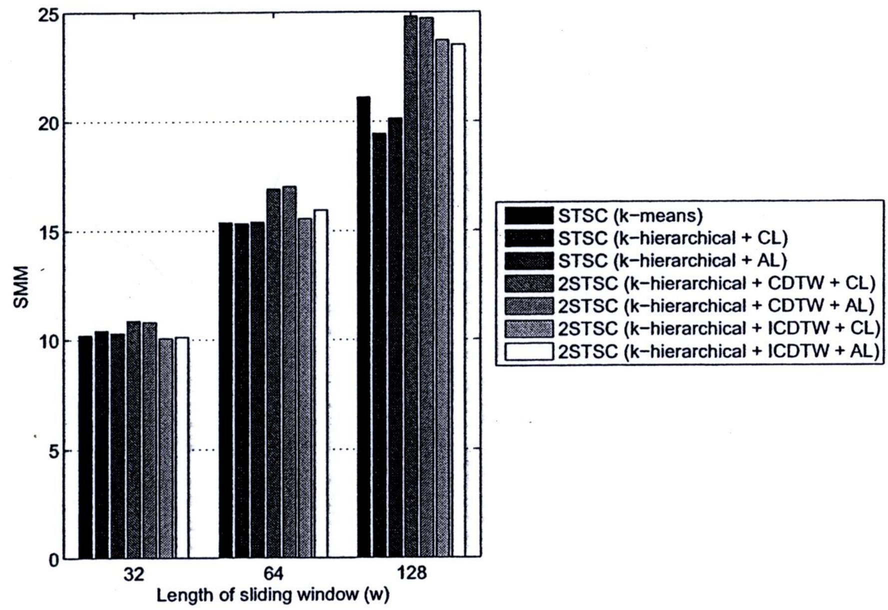


Figure D.7: SMMs of Fortune5004 when the number of clusters (k) is 3 and the length of sliding window (w) is varied.

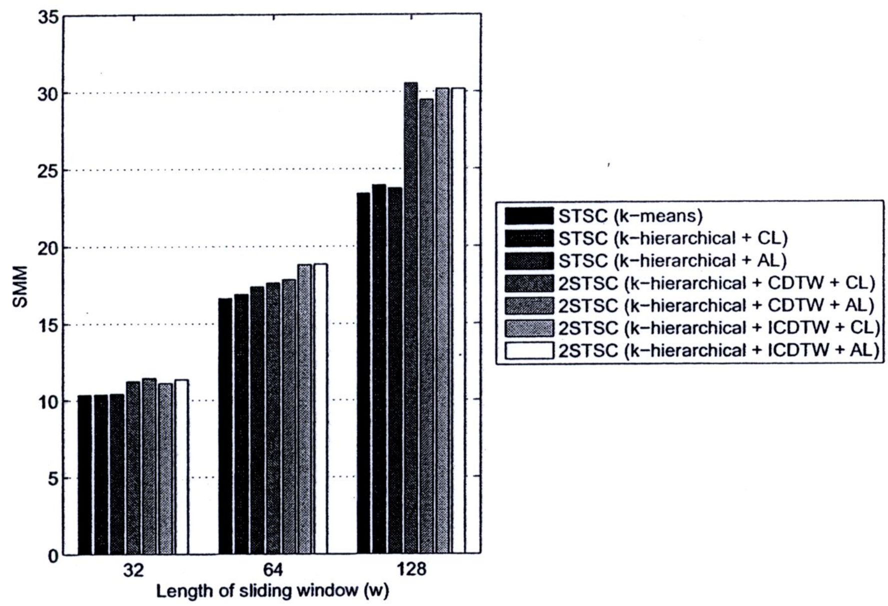


Figure D.8: SMMs of MITDBX108 when the number of clusters (k) is 3 and the length of sliding window (w) is varied.

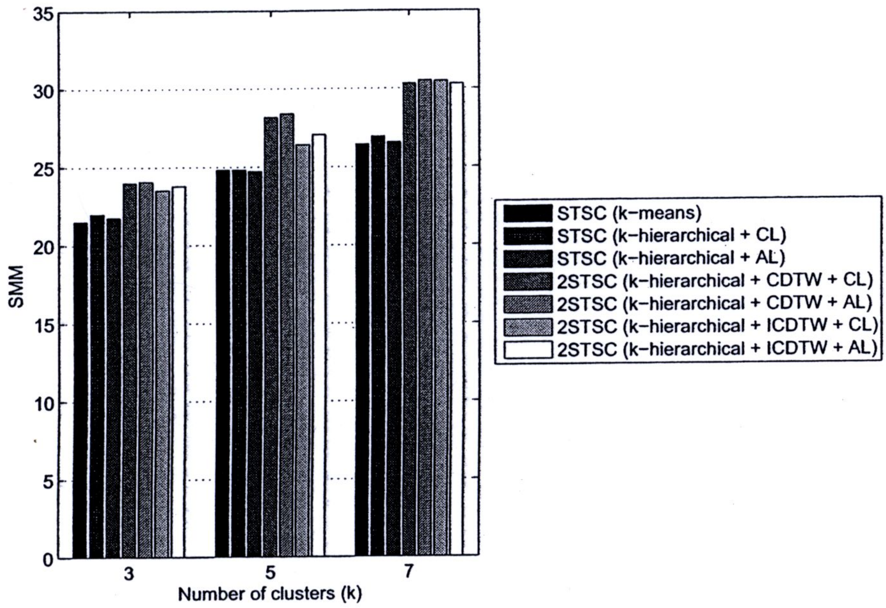


Figure D.9: SMMs of AEM2 when the length of sliding window (w) is 64 and the number of clusters (k) is varied.

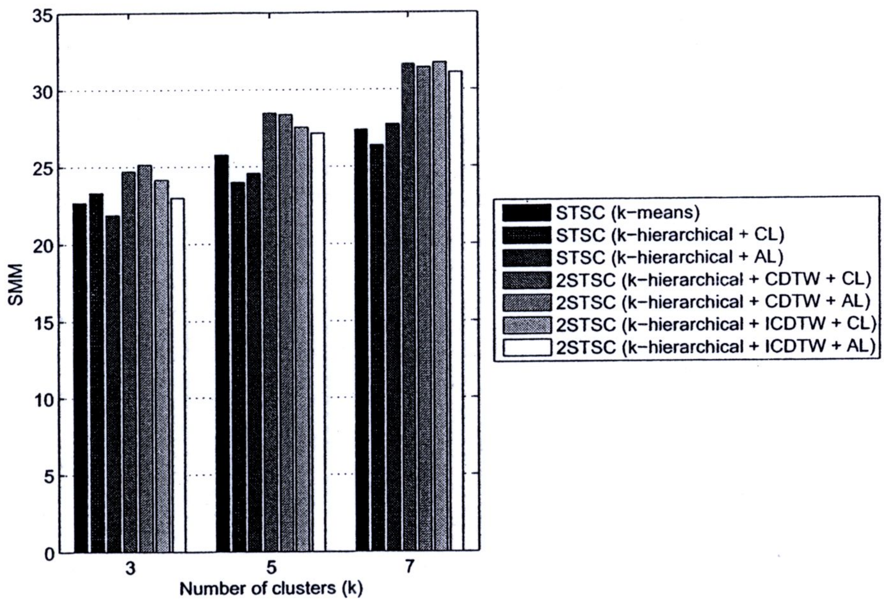


Figure D.10: SMMs of TOR96 when the length of sliding window (w) is 64 and the number of clusters (k) is varied.

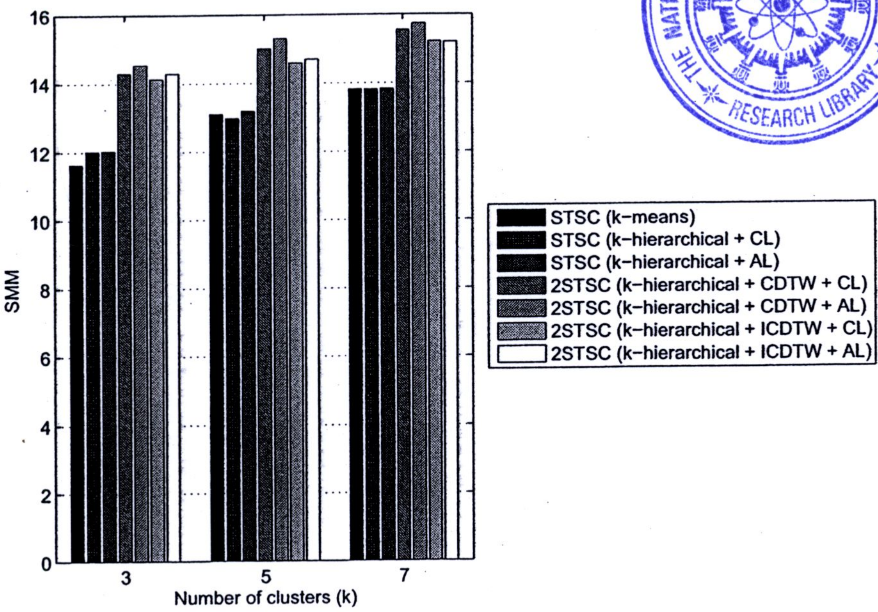


Figure D.11: SMMs of Buoy1 when the length of sliding window (w) is 64 and the number of clusters (k) is varied.

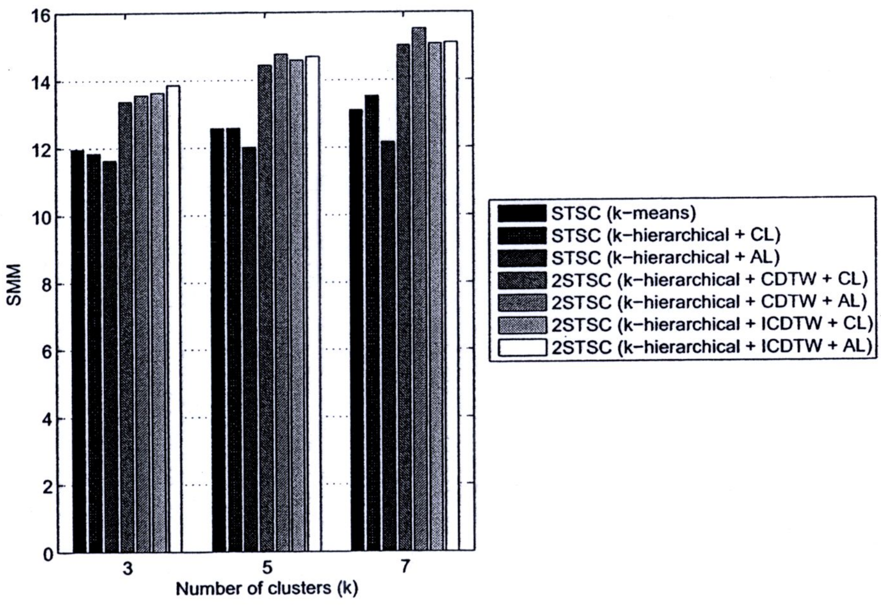


Figure D.12: SMMs of CBF when the length of sliding window (w) is 64 and the number of clusters (k) is varied.

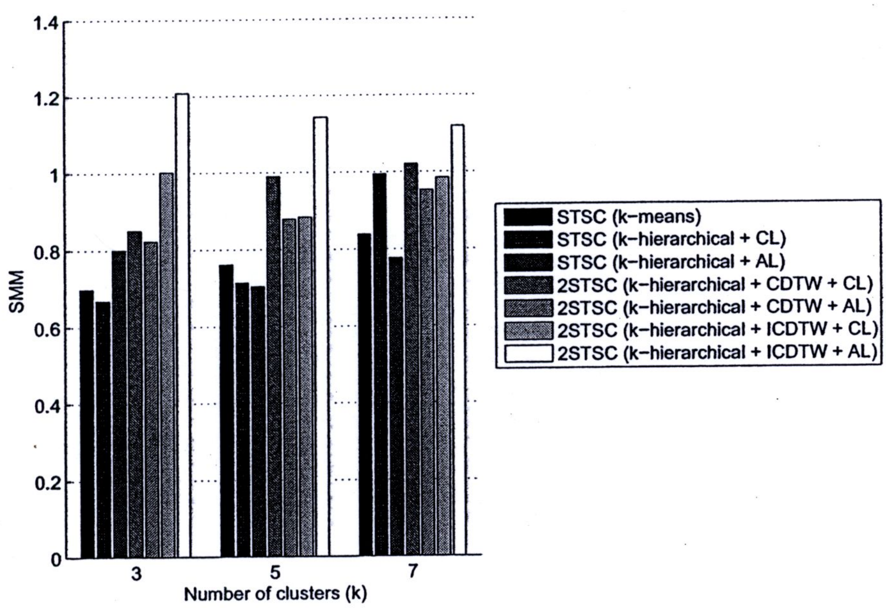


Figure D.13: SMMs of ERP when the length of sliding window (w) is 64 and the number of clusters (k) is varied.

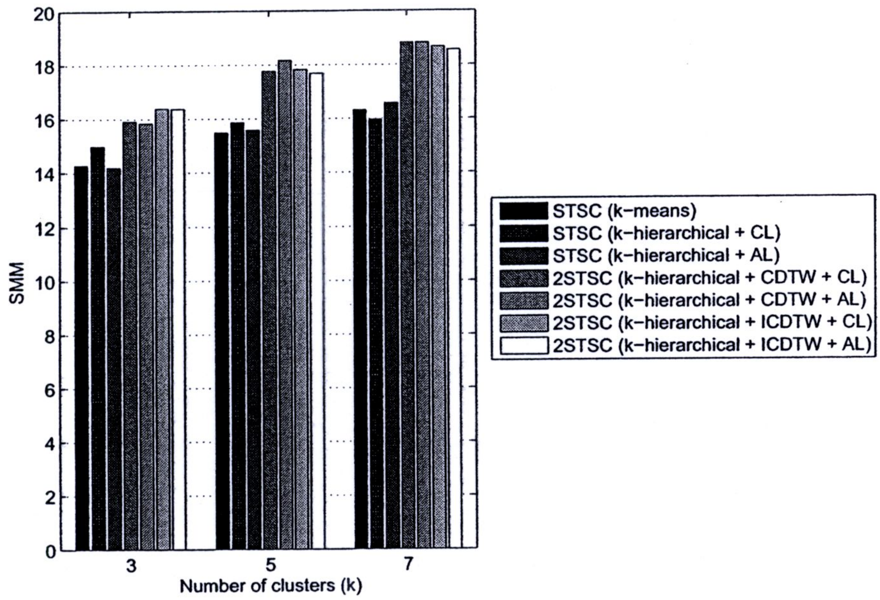


Figure D.14: SMMs of Field4 when the length of sliding window (w) is 64 and the number of clusters (k) is varied.

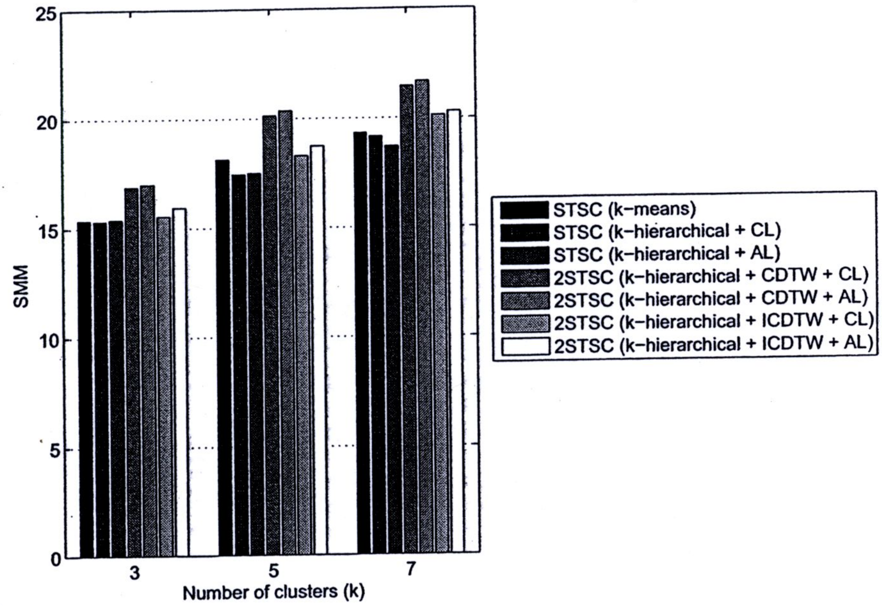


Figure D.15: SMMs of Fortune5004 when the length of sliding window (w) is 64 and the number of clusters (k) is varied.

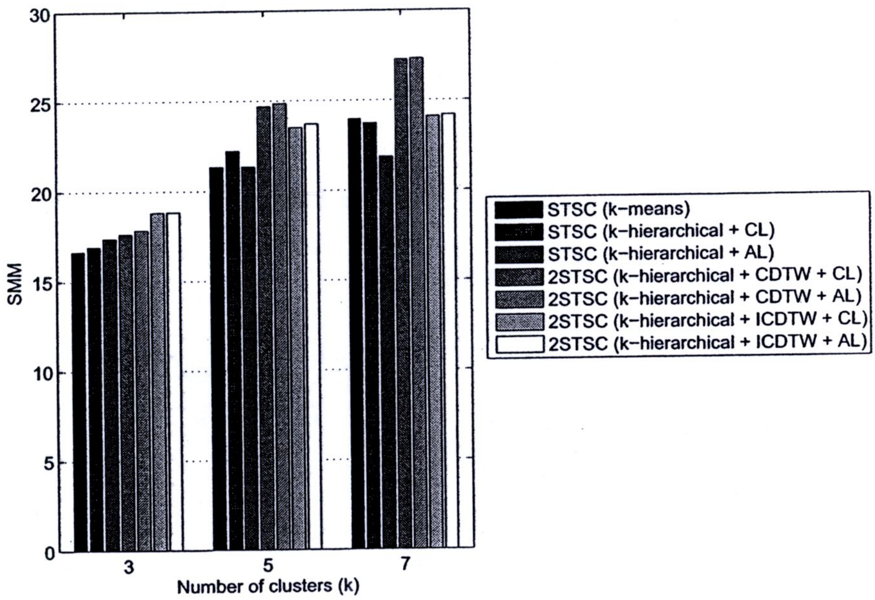


Figure D.16: SMMs of MITDBX108 when the length of sliding window (w) is 64 and the number of clusters (k) is varied.

AEM2

TOR96

Buoy1

CBF

ERP

Field4

Fortune5004

MITDBX108

Figure D.17: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using CDTW function when $k = 3$ and $w = 32$.

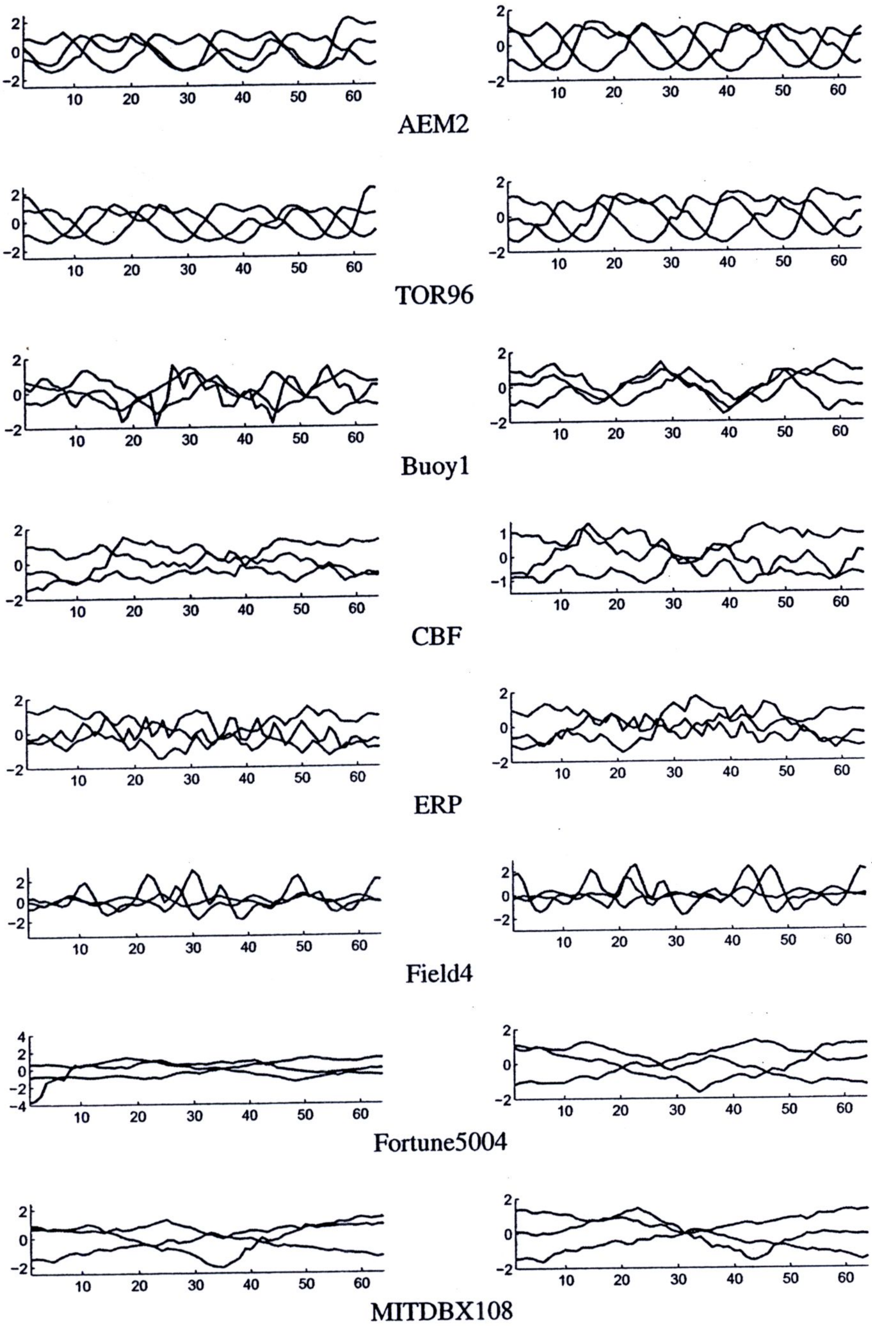


Figure D.18: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using CDTW function when $k = 3$ and $w = 64$.

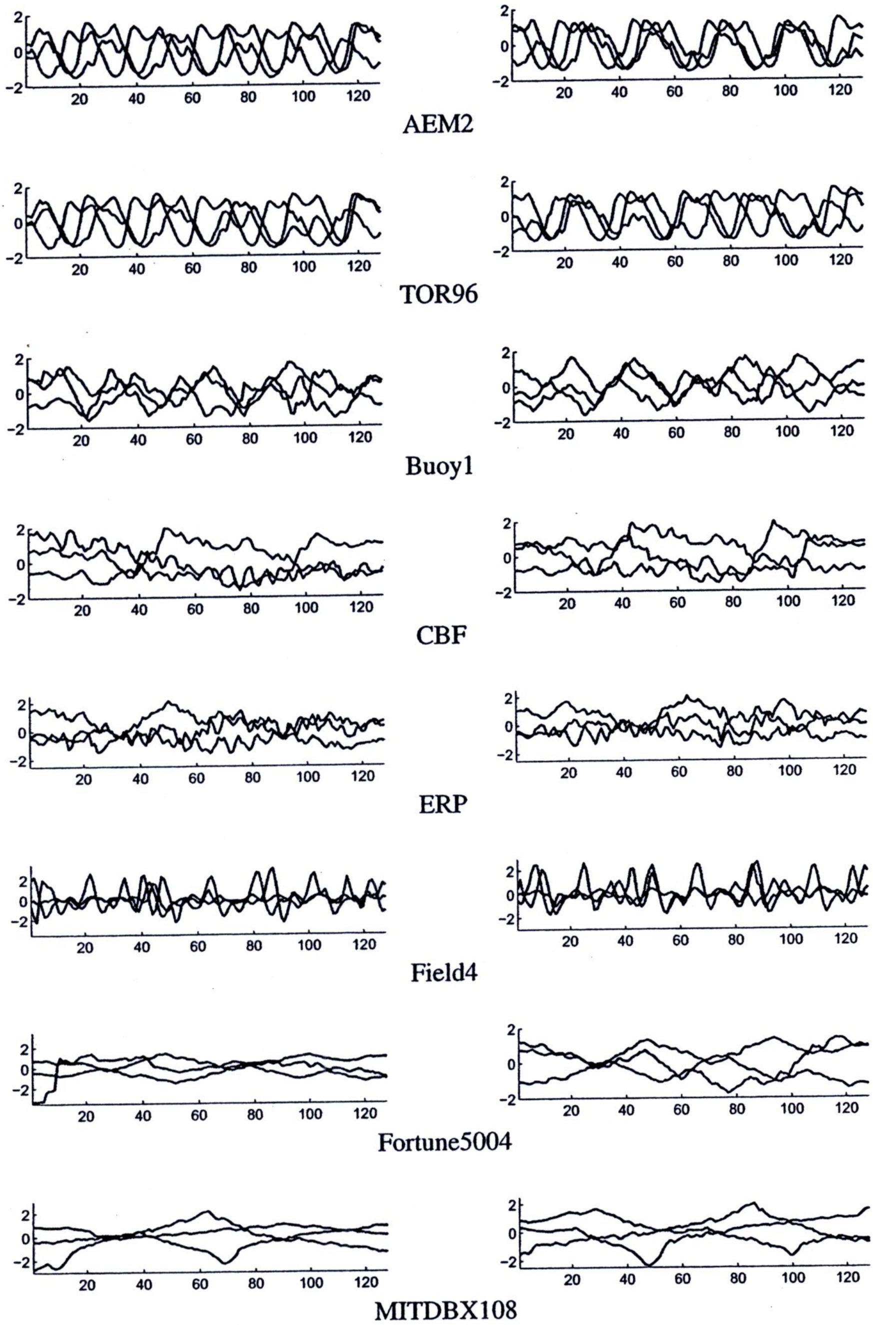


Figure D.19: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using CDTW function when $k = 3$ and $w = 128$.

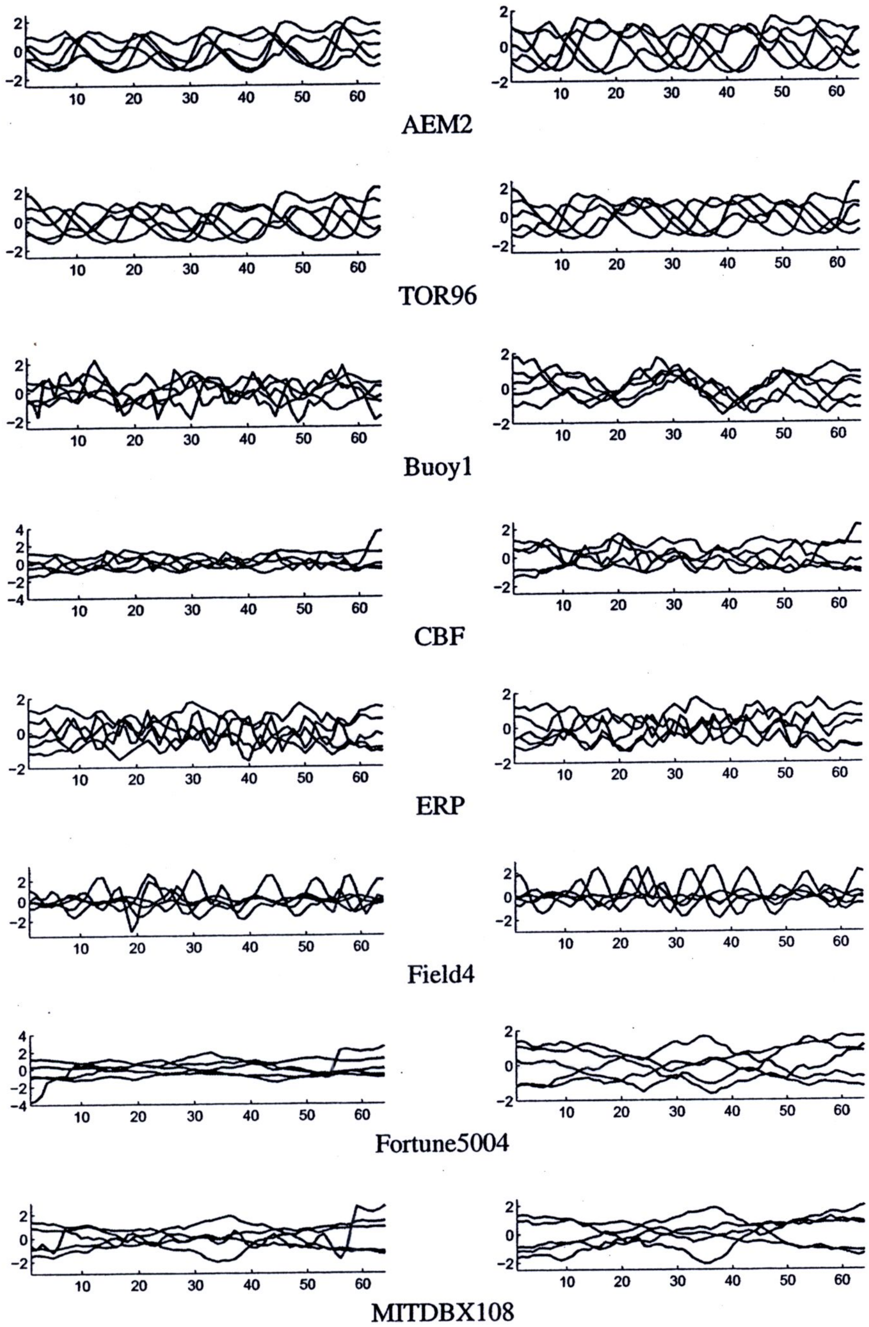


Figure D.20: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using ICDTW function when $k = 5$ and $w = 64$.

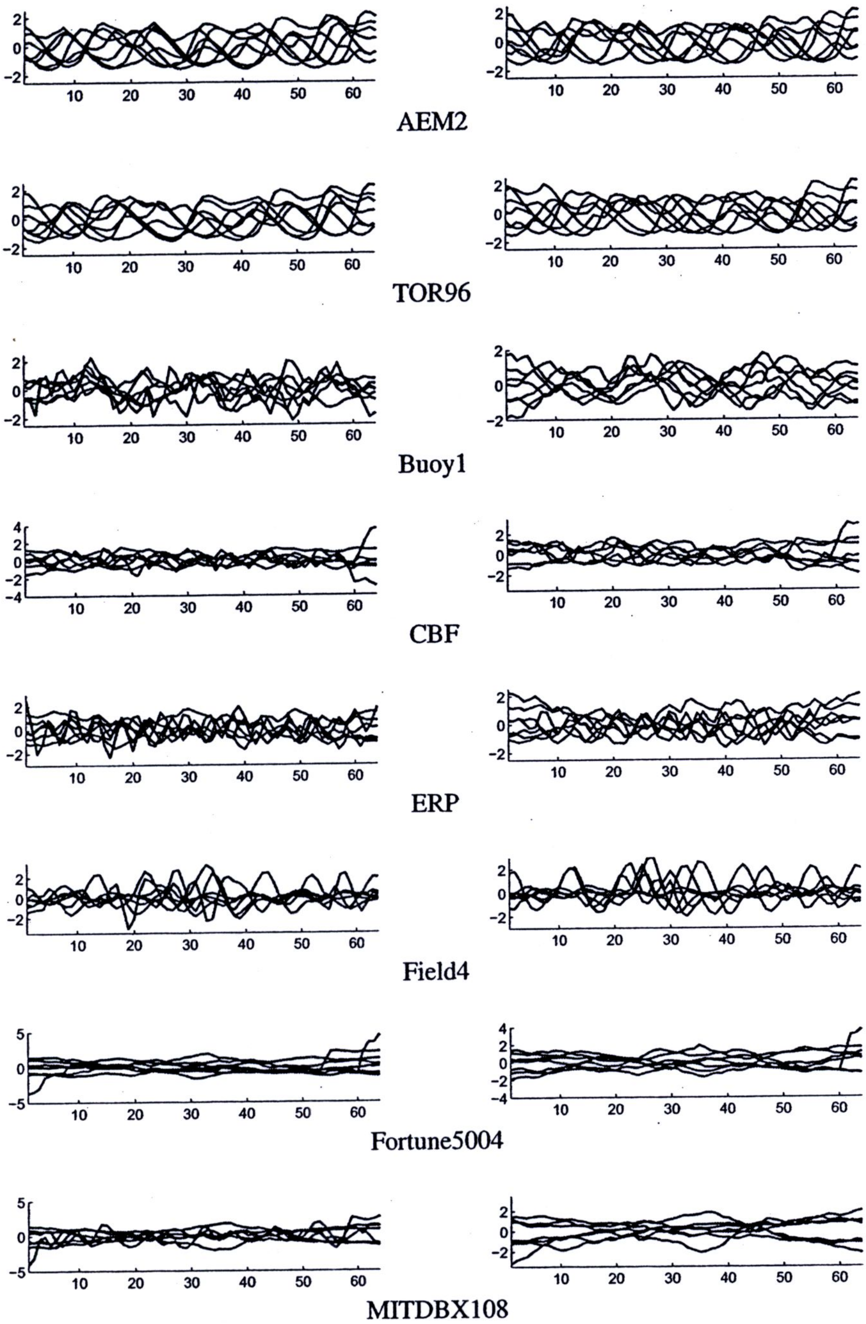


Figure D.21: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using ICDTW function when $k = 7$ and $w = 64$.

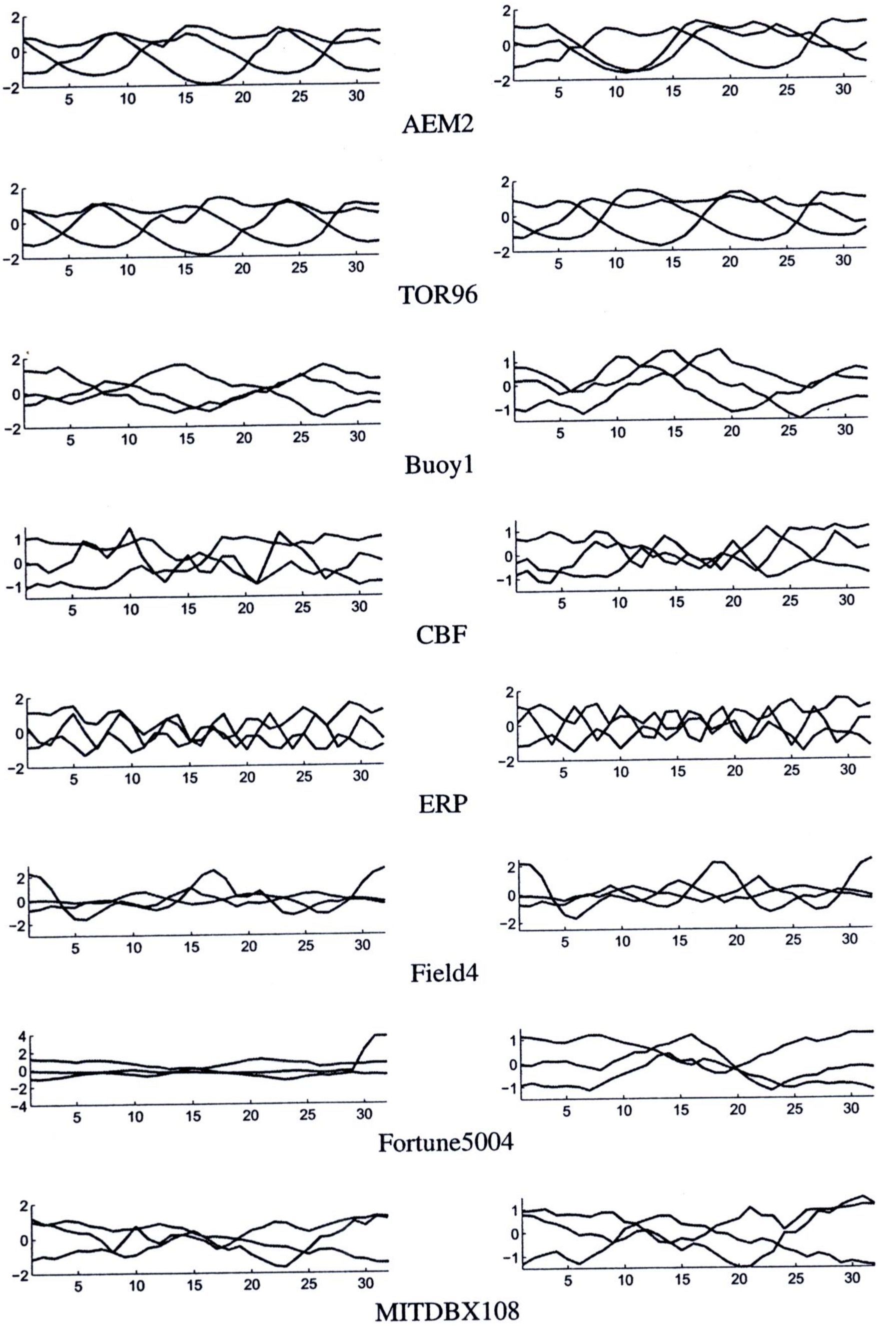


Figure D.22: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using ICDTW function when $k = 3$ and $w = 32$.

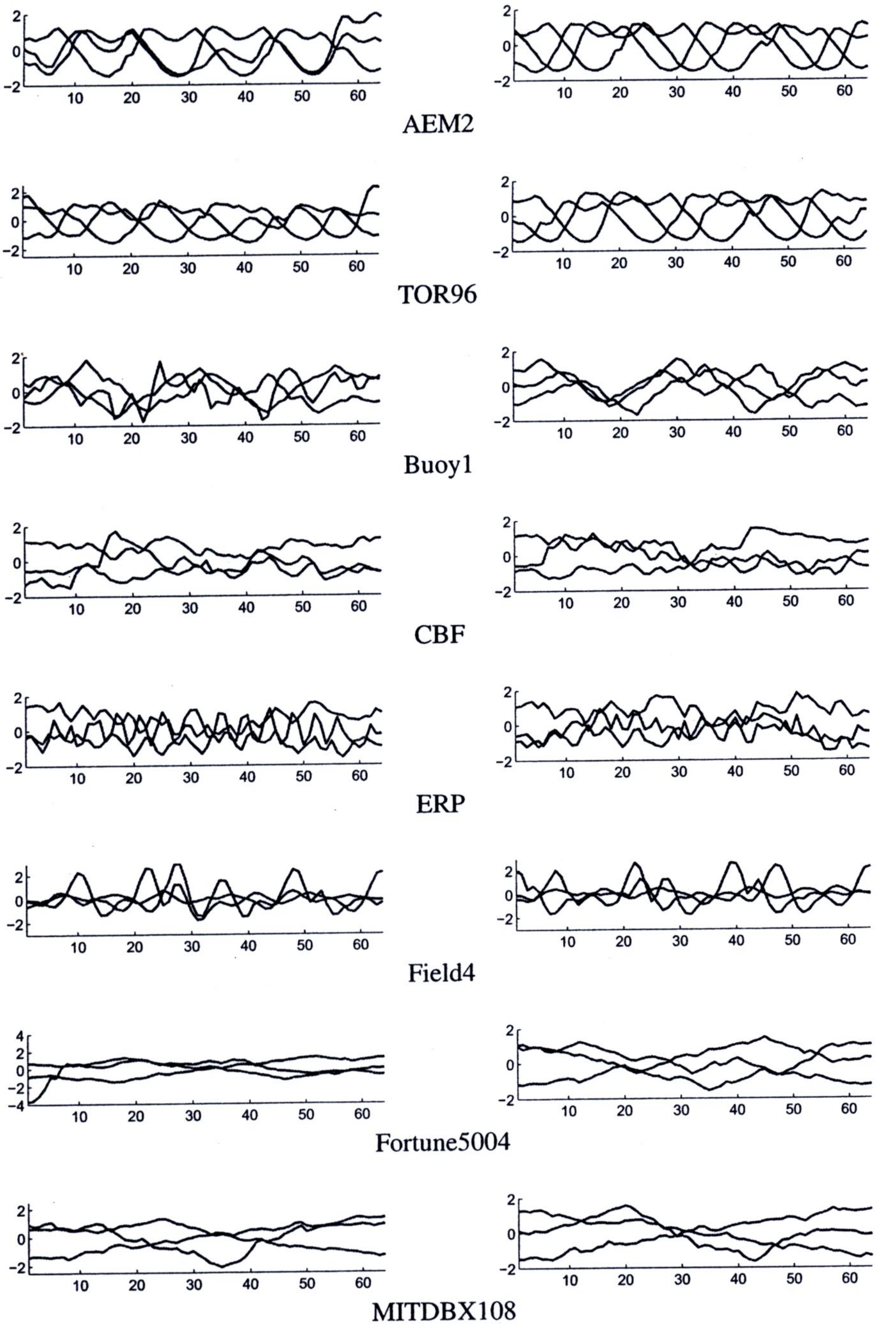


Figure D.23: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using ICDTW function when $k = 3$ and $w = 64$.

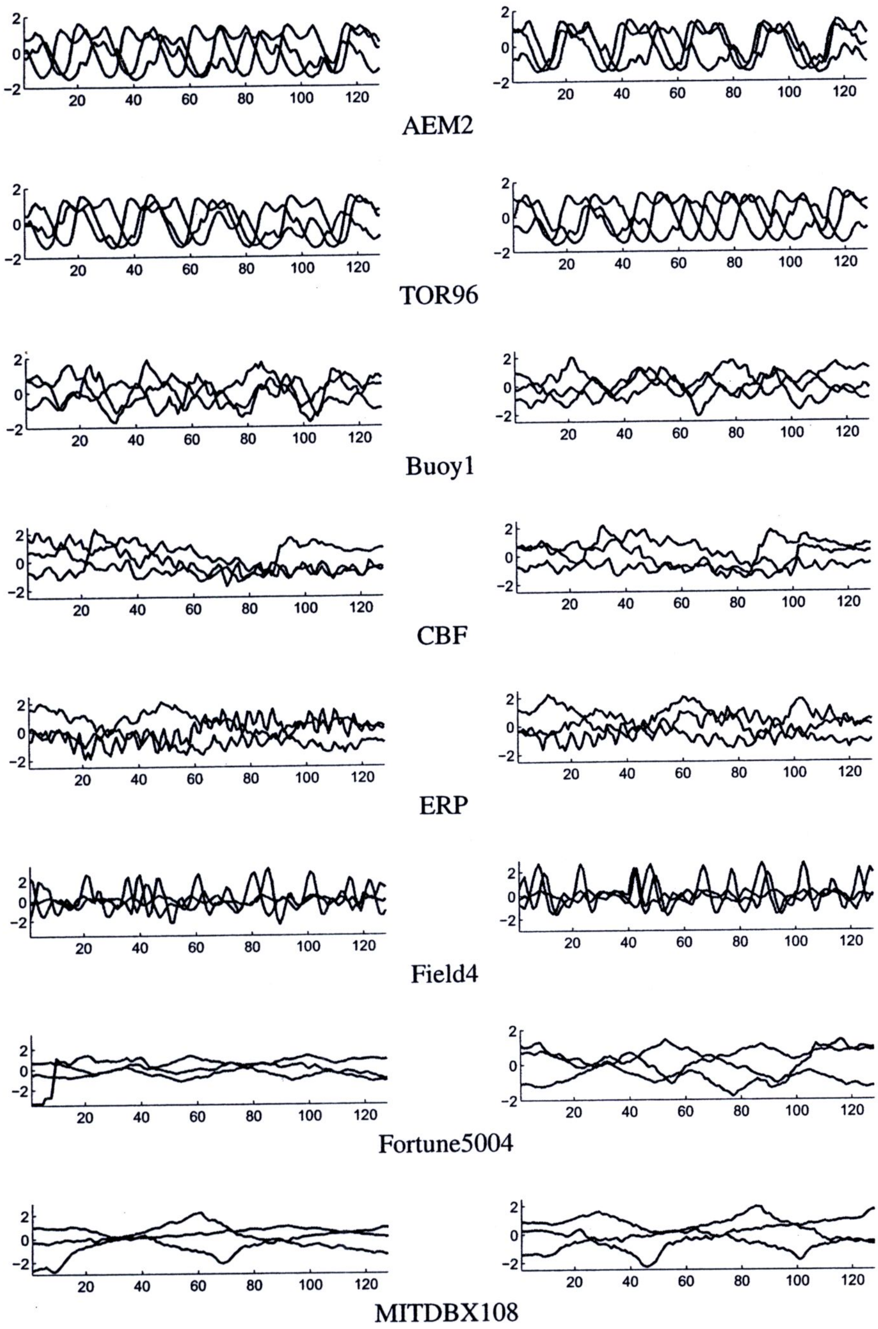


Figure D.24: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using ICDTW function when $k = 3$ and $w = 128$.

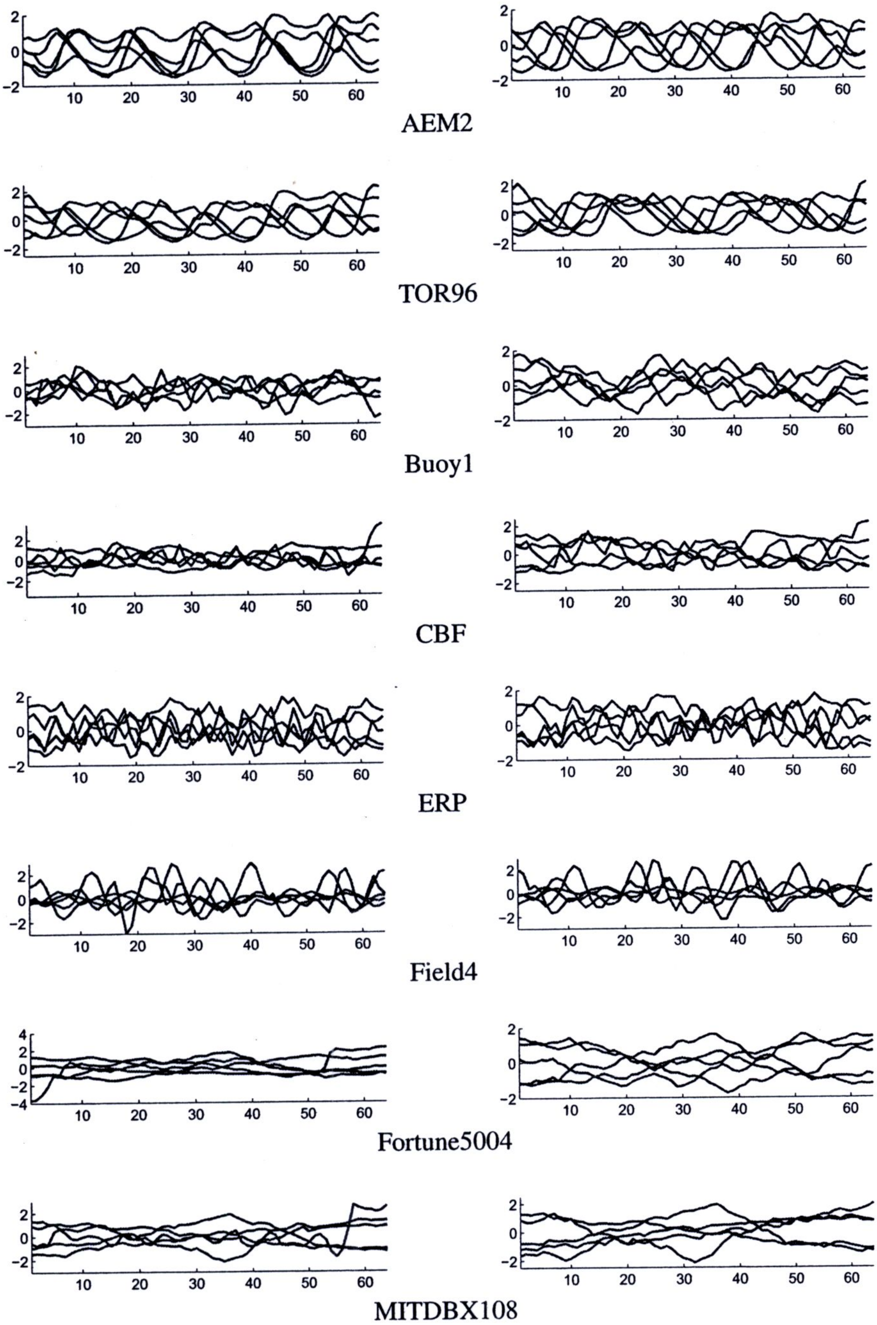


Figure D.25: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using ICDTW function when $k = 5$ and $w = 64$.

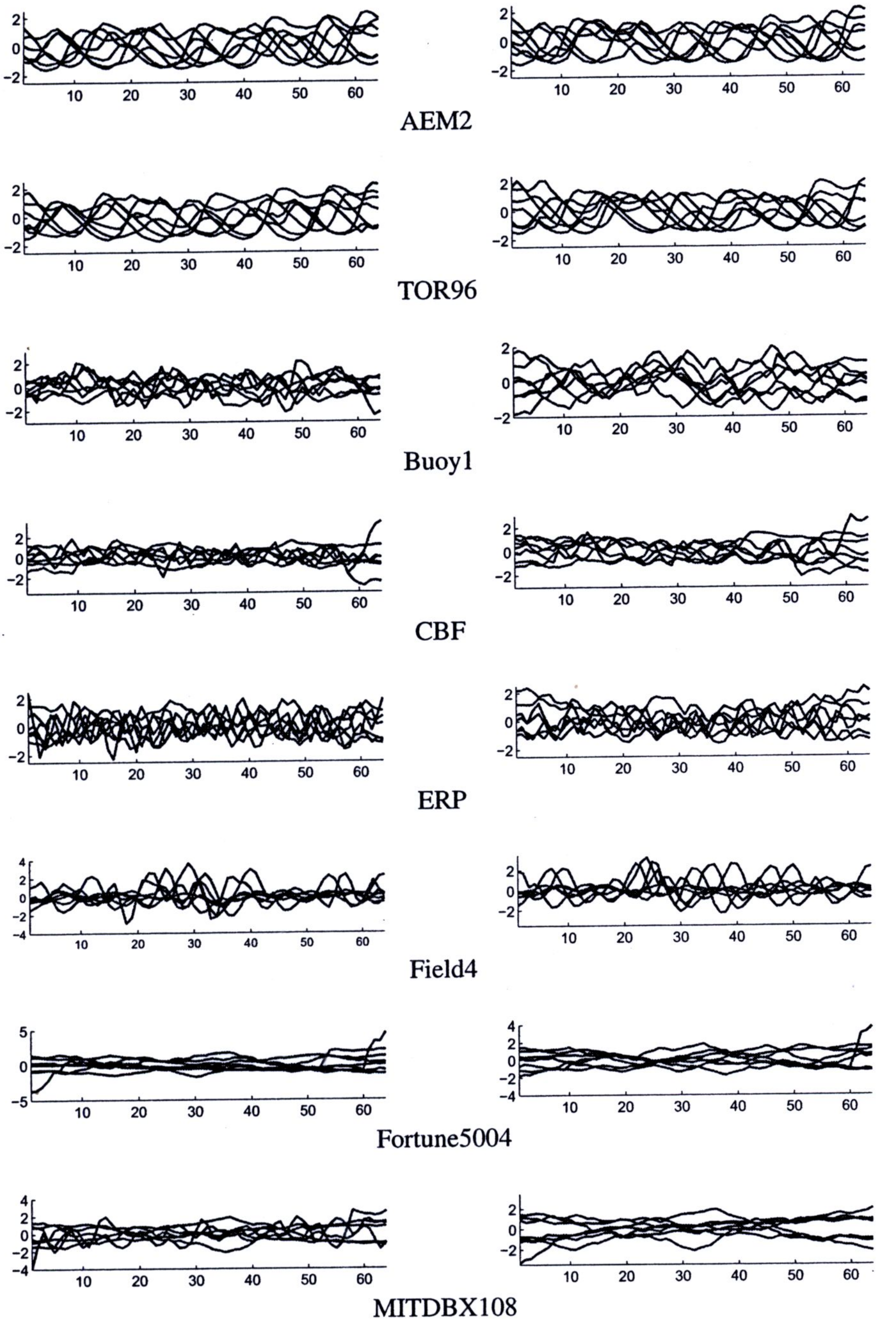


Figure D.26: Cluster representatives generated from 2STSC with complete linkage (left) and average linkage (right) using ICDTW function when $k = 7$ and $w = 64$.

APPENDIX E

COMPLETE EXPERIMENTAL RESULTS OF THE FIRST EXPERIMENT IN CHAPTER V

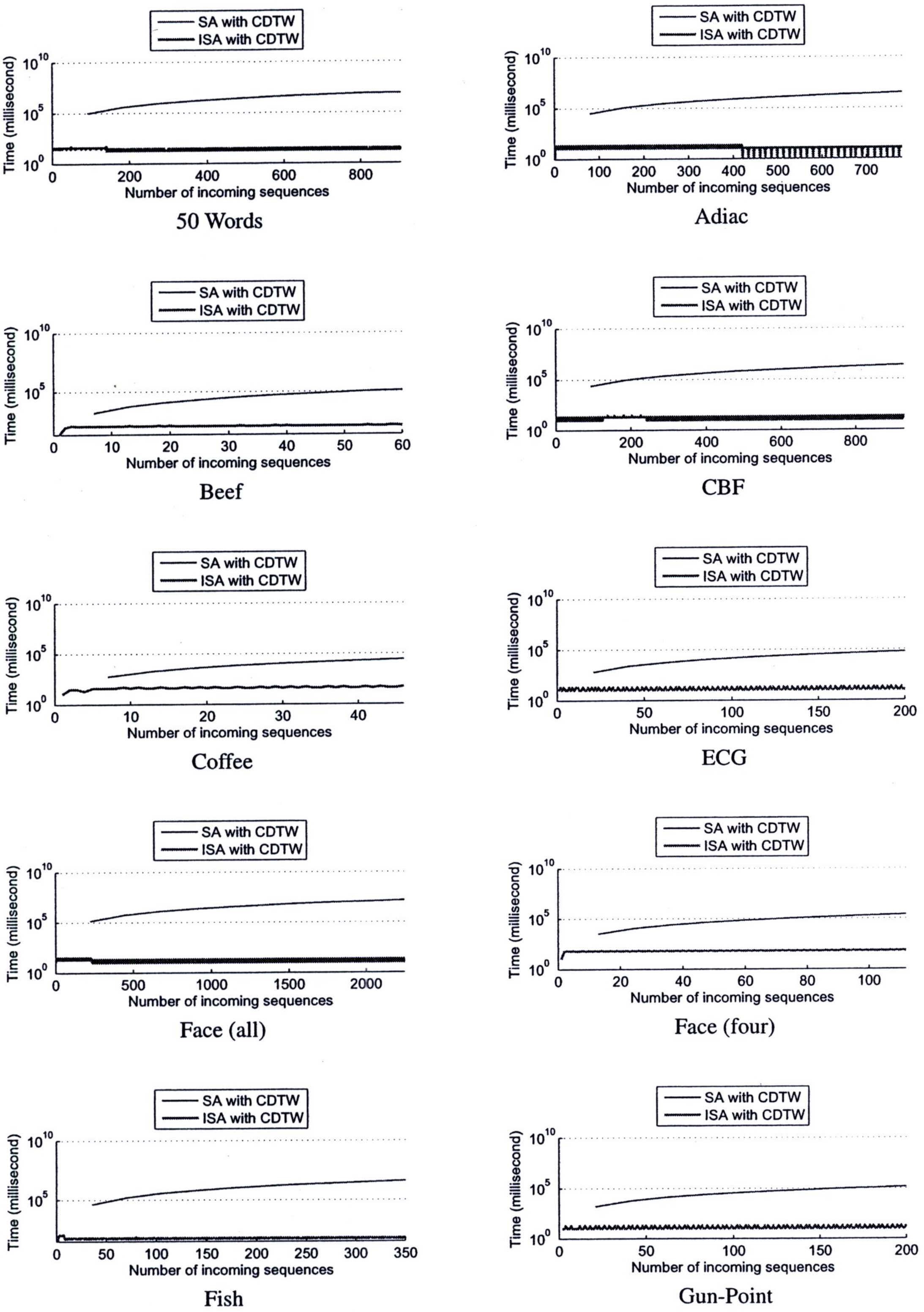


Figure E.1: Computational time of Incremental Shape-based Averaging and Shape-based Averaging with CDTW function when a new incoming sequence arrives.

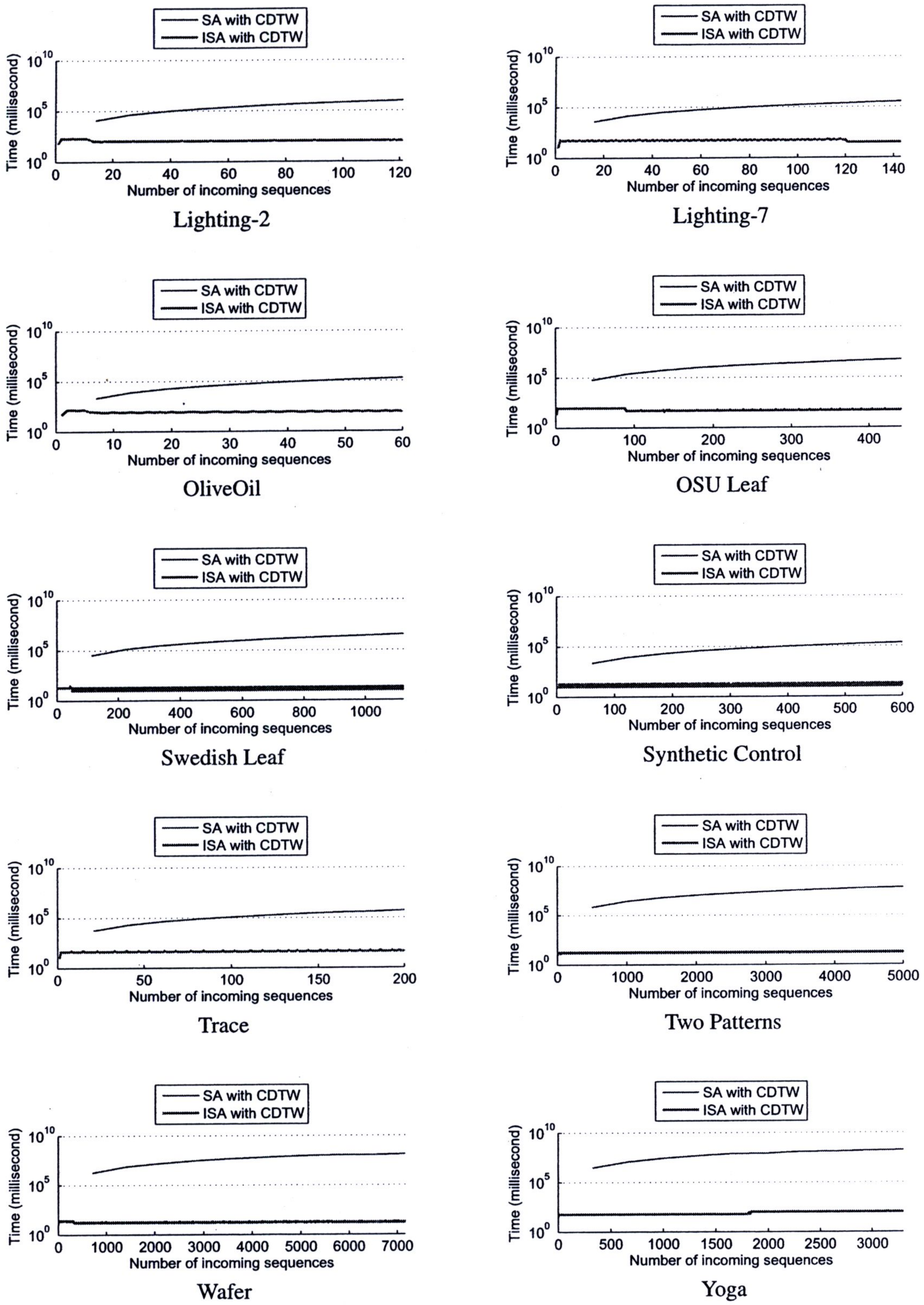


Figure E.2: Computational time of Incremental Shape-based Averaging and Shape-based Averaging with CDTW function when a new incoming sequence arrives. (cont.)

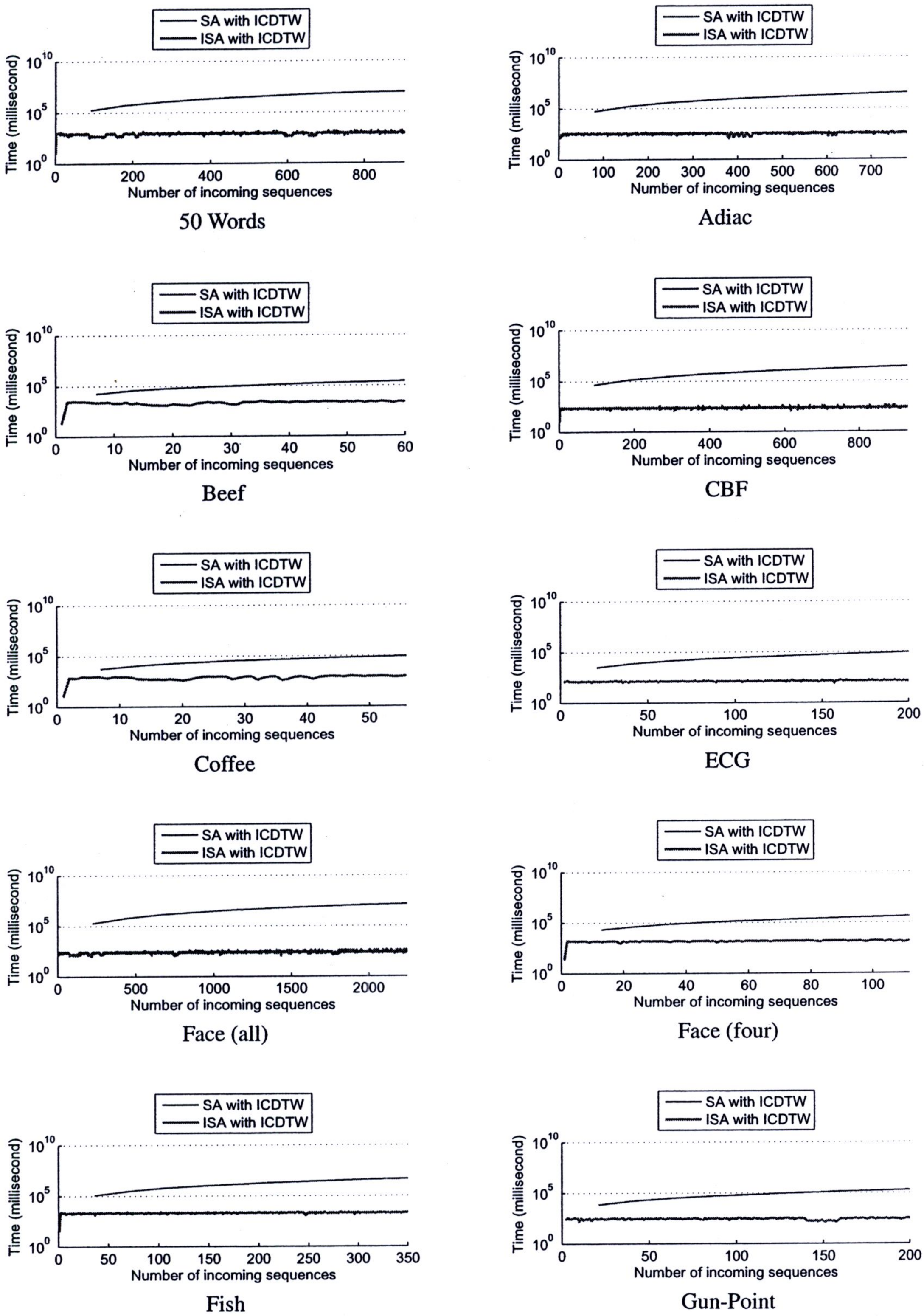
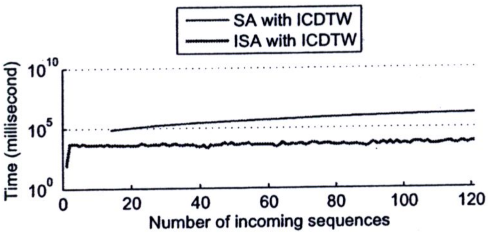
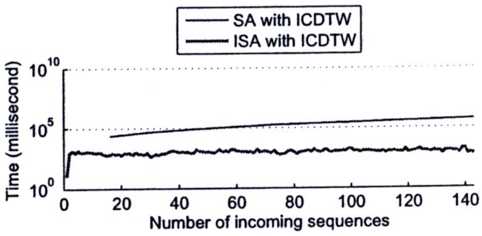


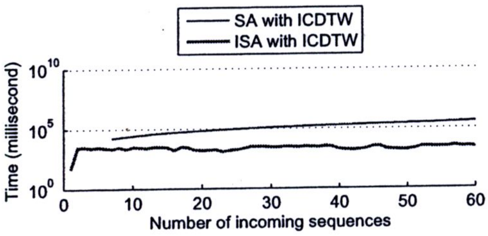
Figure E.3: Computational time of Incremental Shape-based Averaging and Shape-based Averaging with ICDTW function when a new incoming sequence arrives.



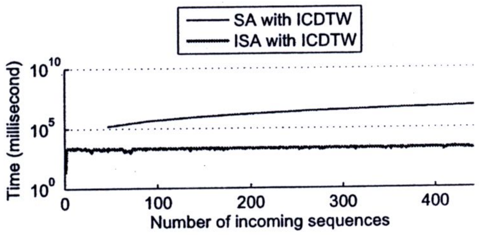
Lighting-2



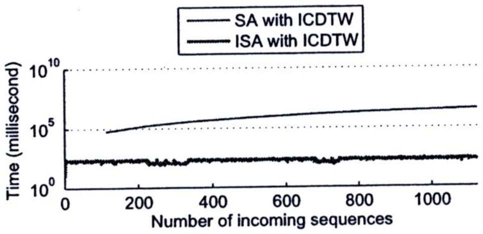
Lighting-7



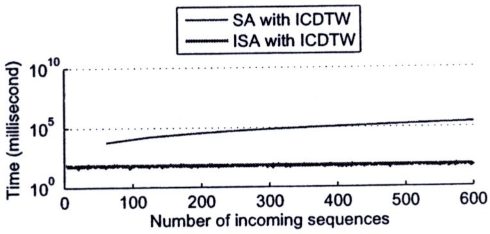
OliveOil



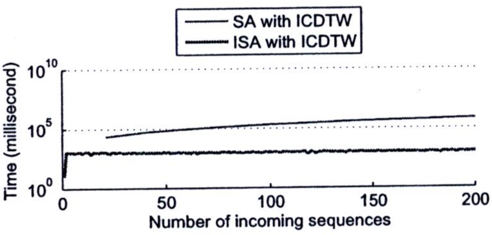
OSU Leaf



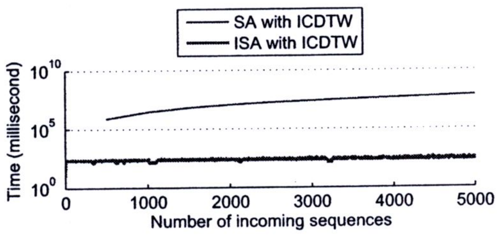
Swedish Leaf



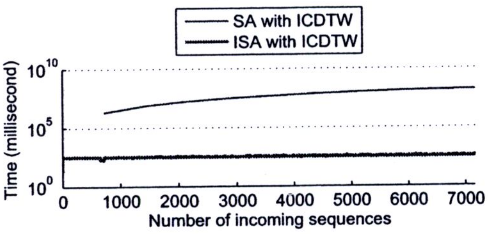
Synthetic Control



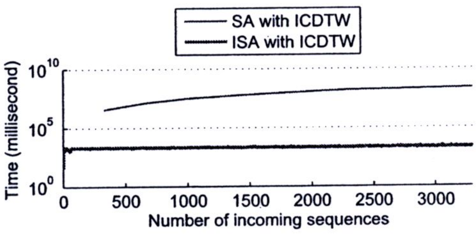
Trace



Two Patterns



Wafer



Yoga

Figure E.4: Computational time ofthan Shape-based Averaging around Incremental Shape-based Averaging and Shape-based Averaging with ICDTW function when a new incoming sequence arrives.

APPENDIX F

COMPLETE EXPERIMENTAL RESULTS OF THE SECOND EXPERIMENT ON CHAPTER V

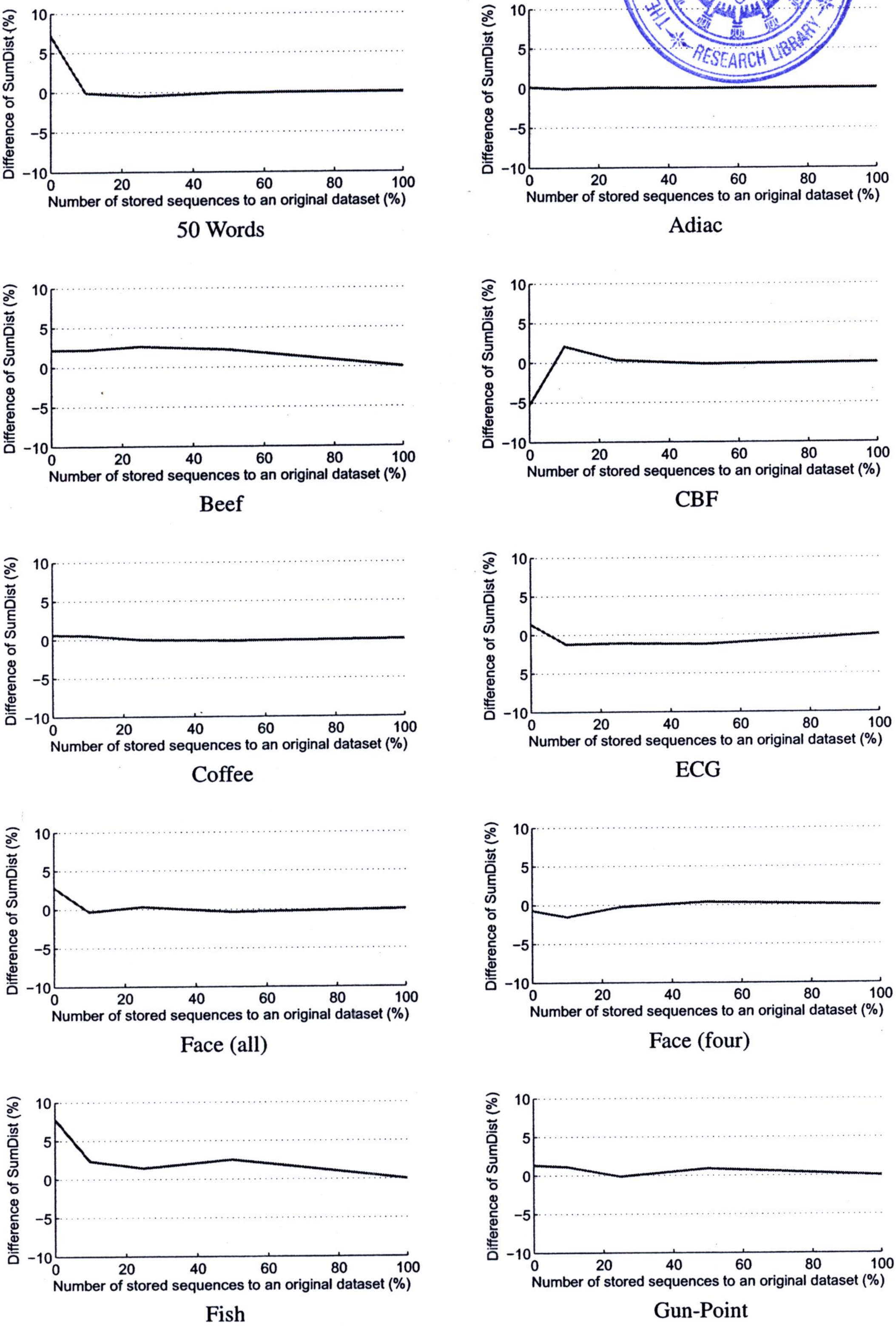


Figure F.1: Difference of SUMDIST of Incremental Shape-based Averaging with CDTW when the number of stored sequences to an original dataset is varied.

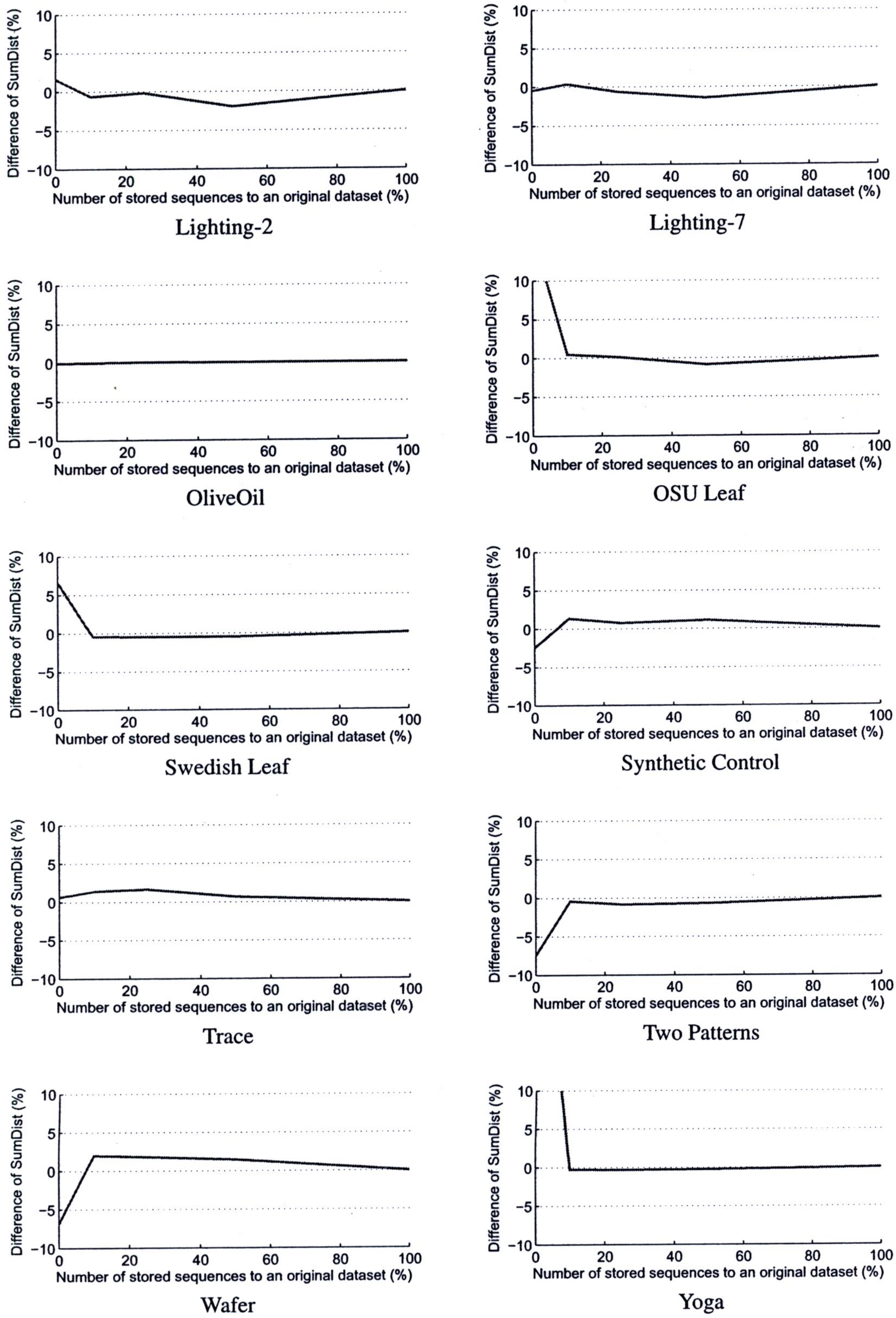


Figure F.2: Difference of SUMDIST of Incremental Shape-based Averaging with CDTW when the number of stored sequences to an original dataset is varied. (cont.)

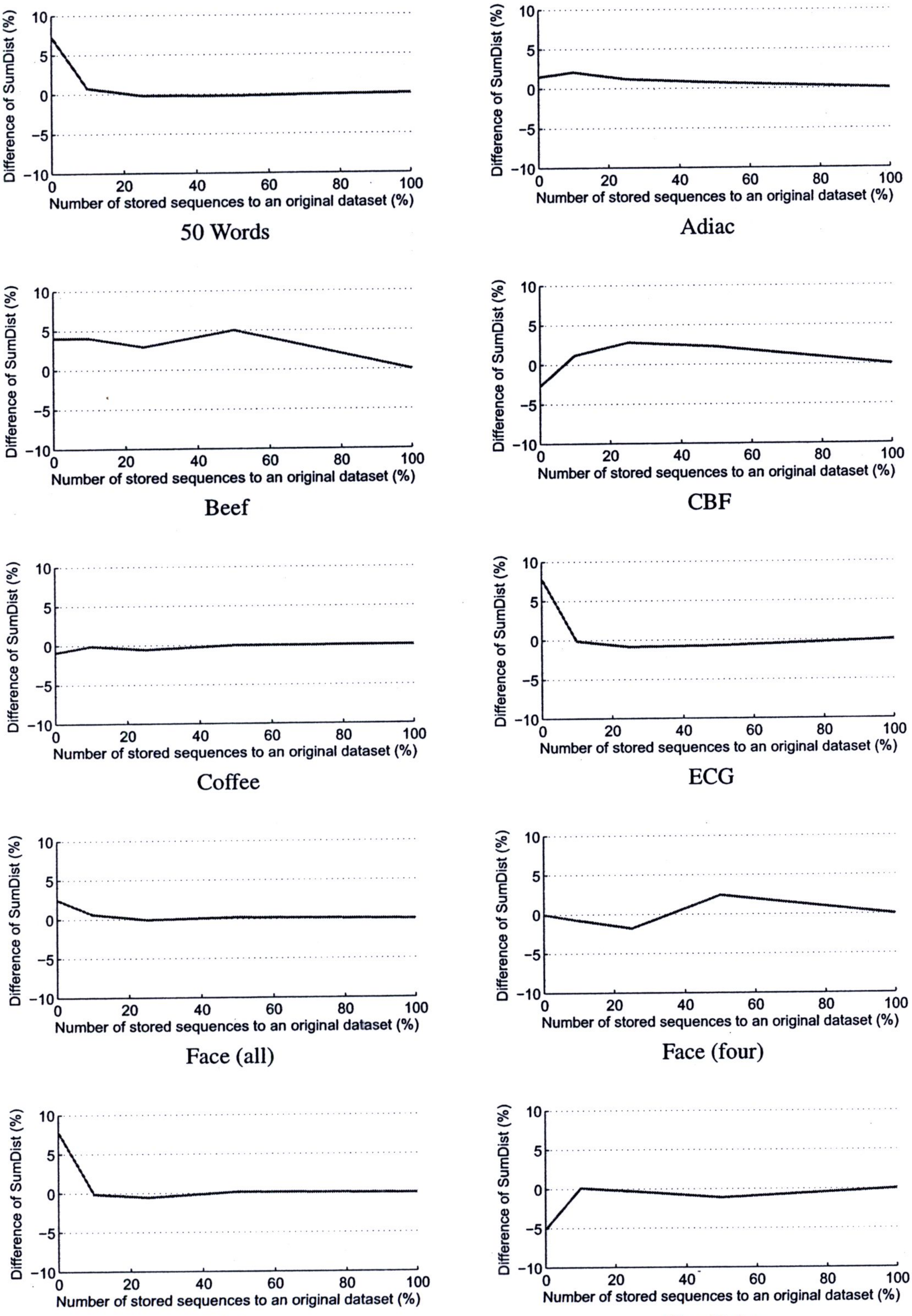
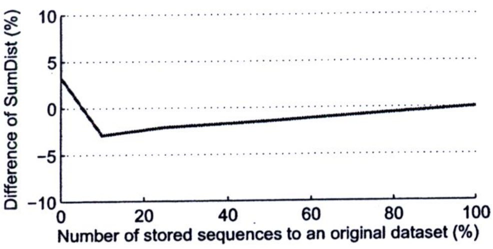
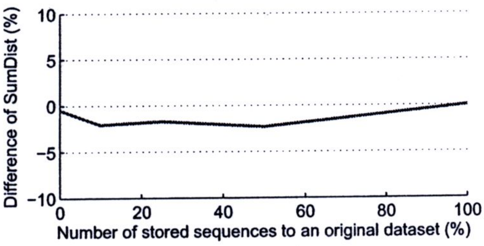


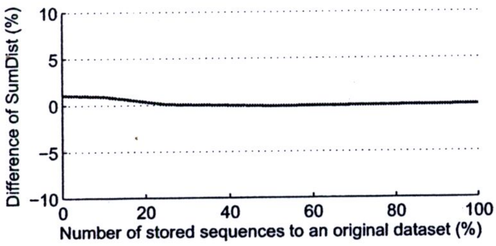
Figure F.3: Difference of SUMDIST of Incremental Shape-based Averaging with ICDTW when the number of stored sequences to an original dataset is varied.



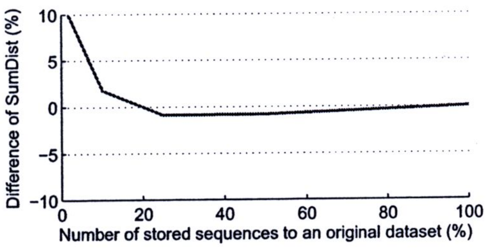
Lighting-2



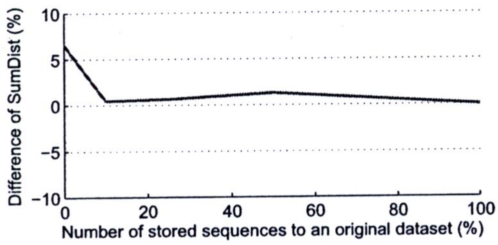
Lighting-7



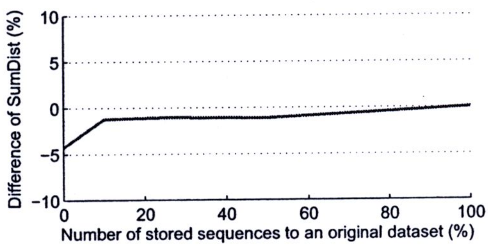
OliveOil



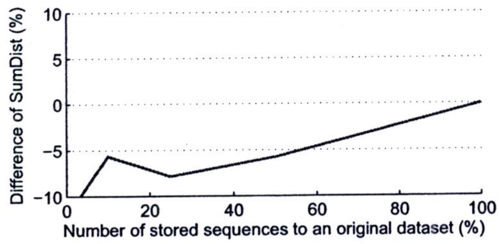
OSU Leaf



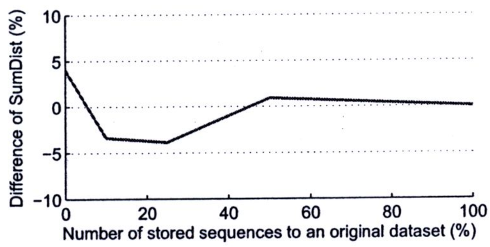
Swedish Leaf



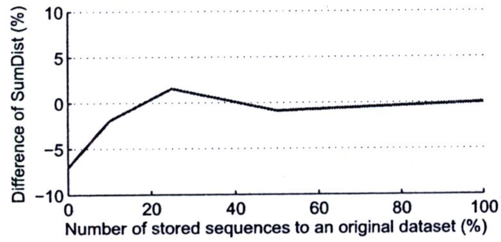
Synthetic Control



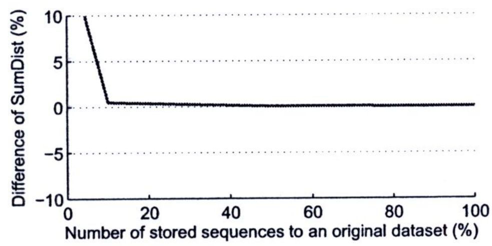
Trace



Two Patterns



Wafer



Yoga

Figure F.4: Difference of SUMDIST of Incremental Shape-based Averaging with ICDTW when the number of stored sequences to an original dataset is varied. (cont.)

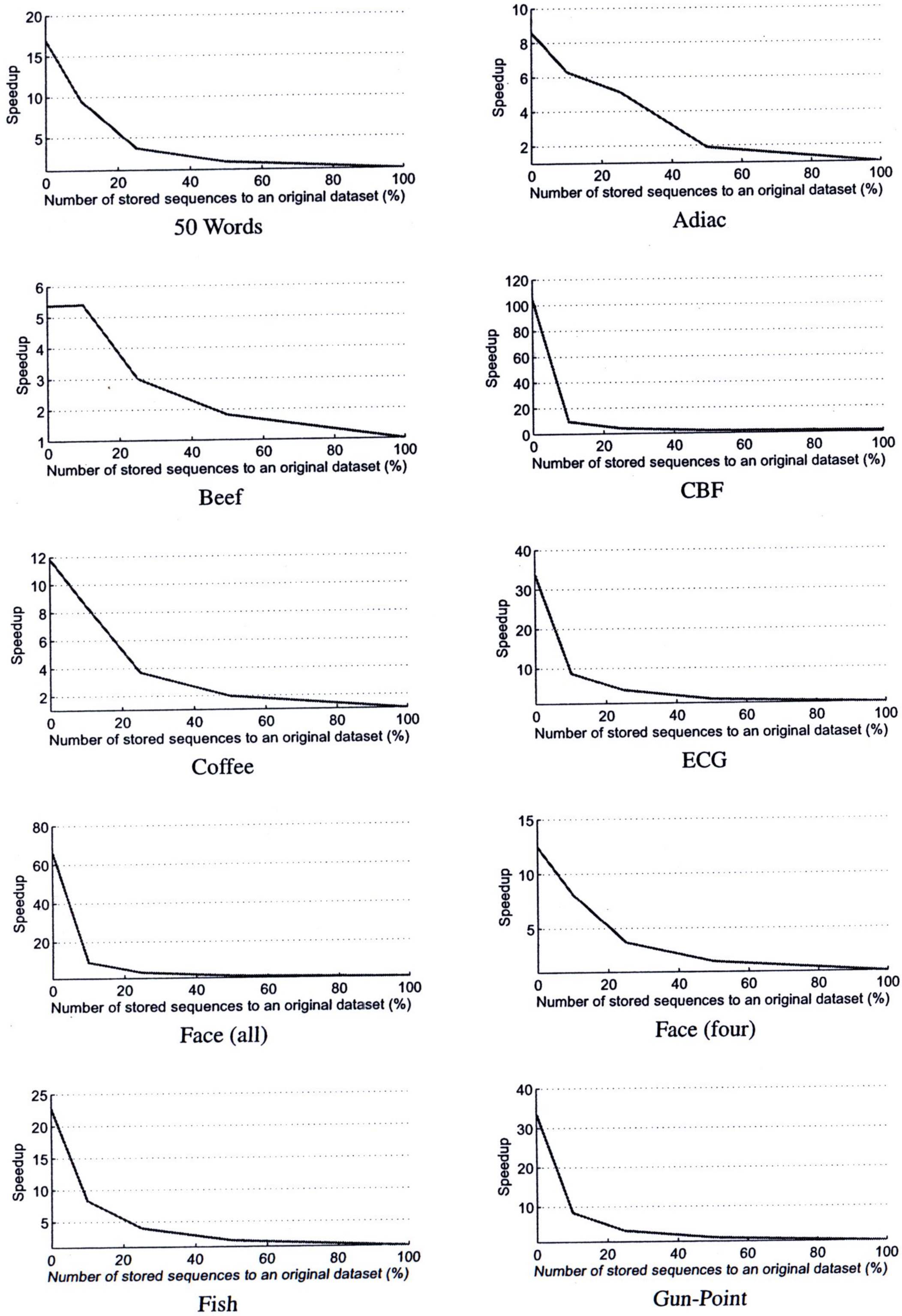


Figure F.5: Speedup of Incremental Shape-based Averaging with CDTW when the number of stored sequences to an original dataset is varied.

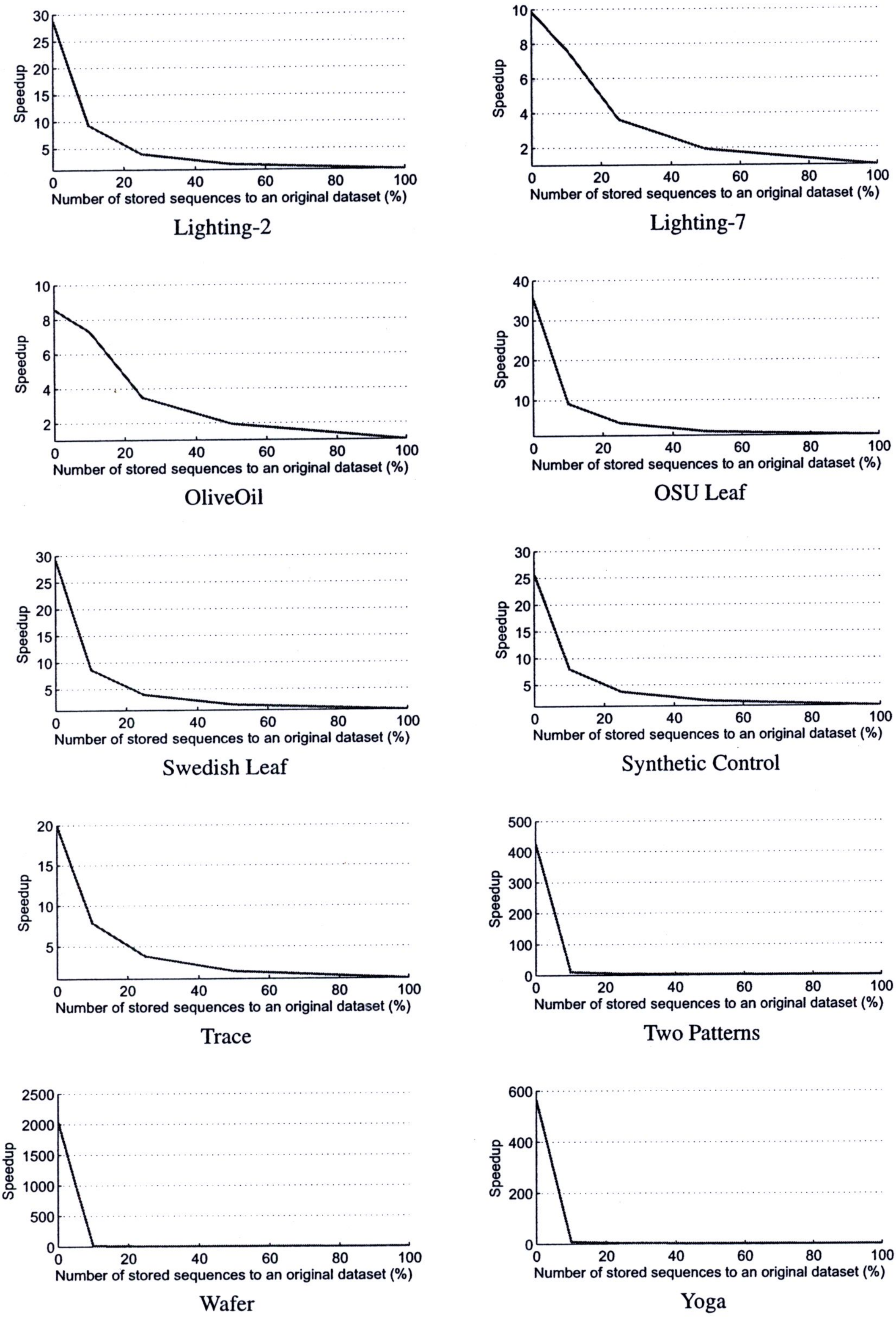


Figure F.6: Speedup of Incremental Shape-based Averaging with CDTW when the number of stored sequences to an original dataset is varied. (cont.)

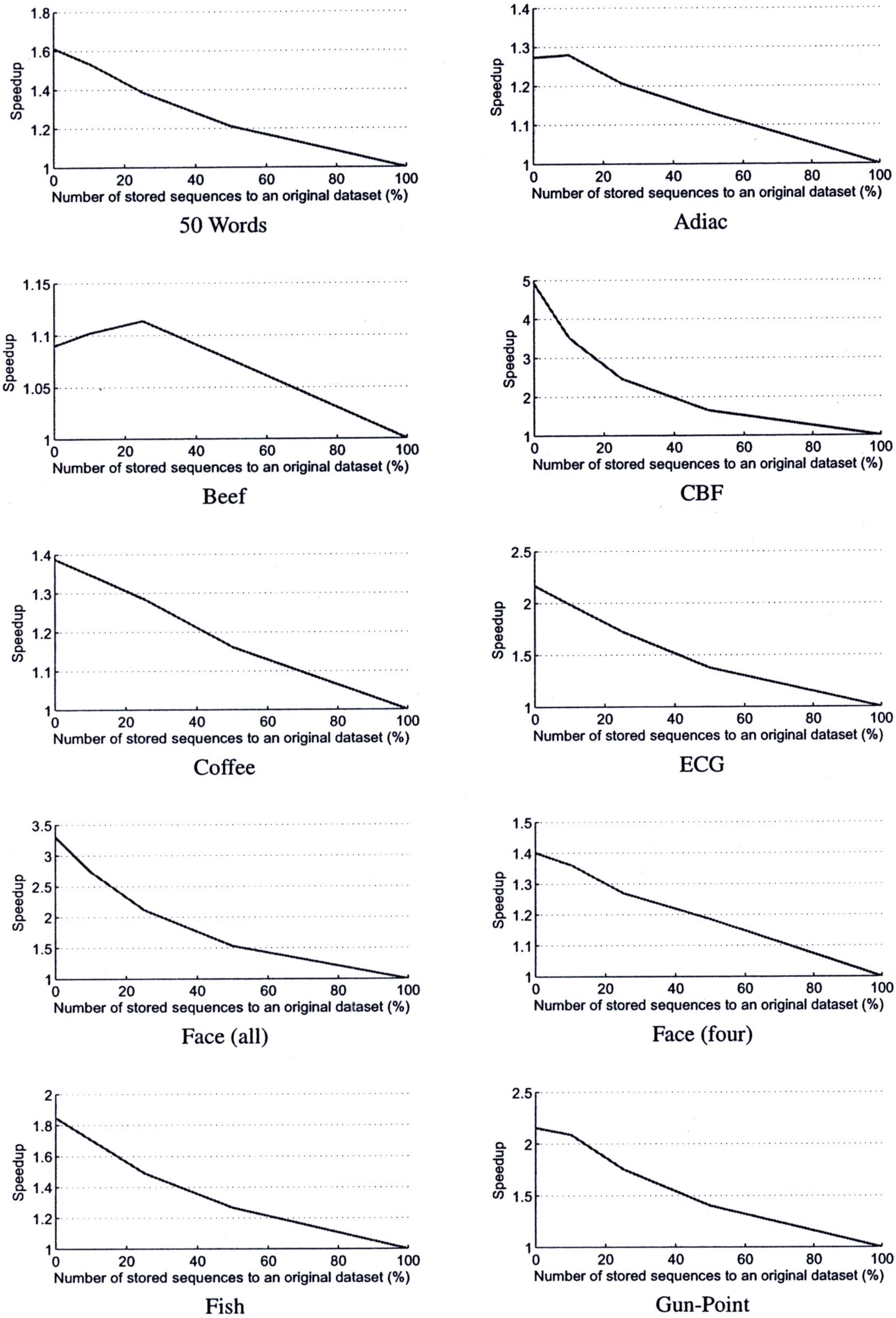
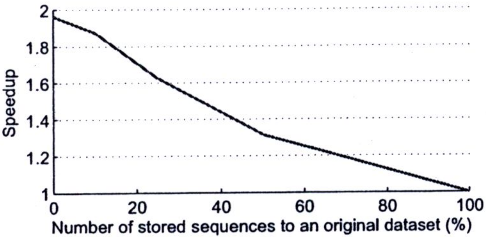
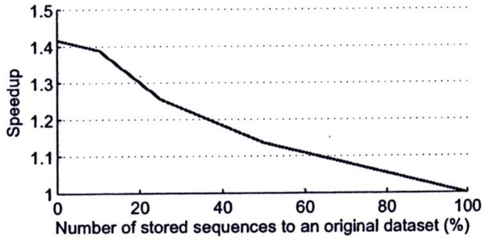


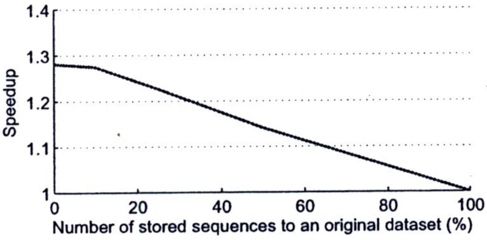
Figure F.7: Speedup of Incremental Shape-based Averaging with ICDTW when the number of stored sequences to an original dataset is varied.



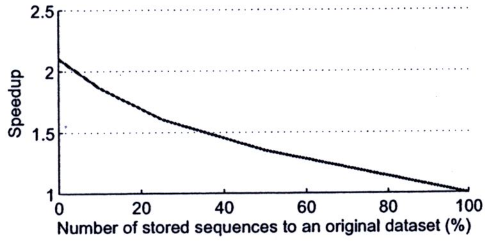
Lighting-2



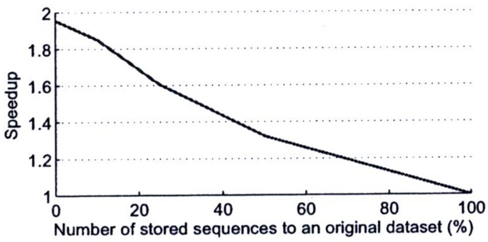
Lighting-7



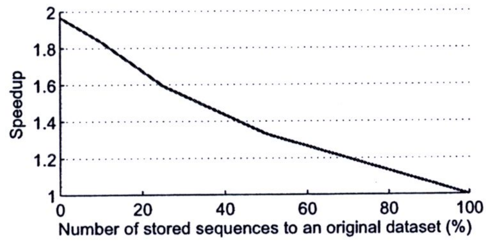
OliveOil



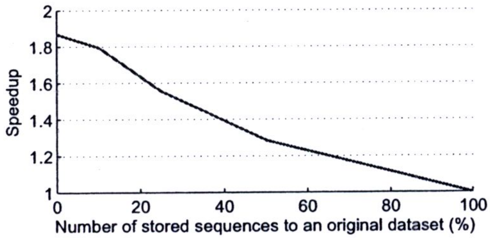
OSU Leaf



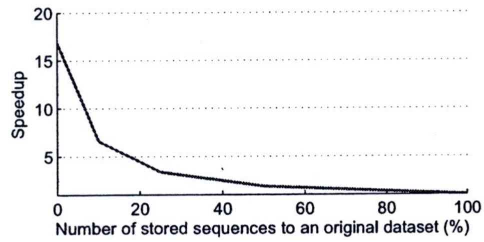
Swedish Leaf



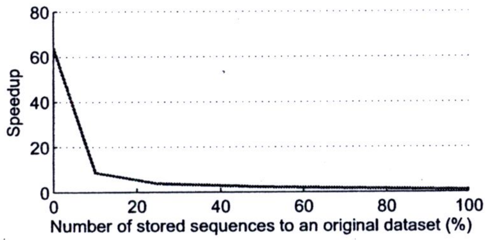
Synthetic Control



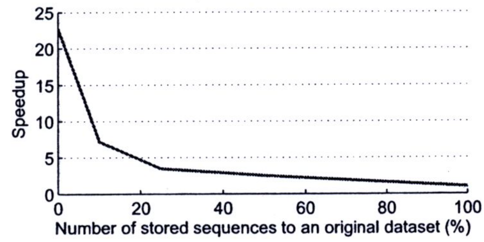
Trace



Two Patterns



Wafer



Yoga

Figure F.8: Speedup of Incremental Shape-based Averaging with CDTW when the number of stored sequences to an original dataset is varied. (cont.)

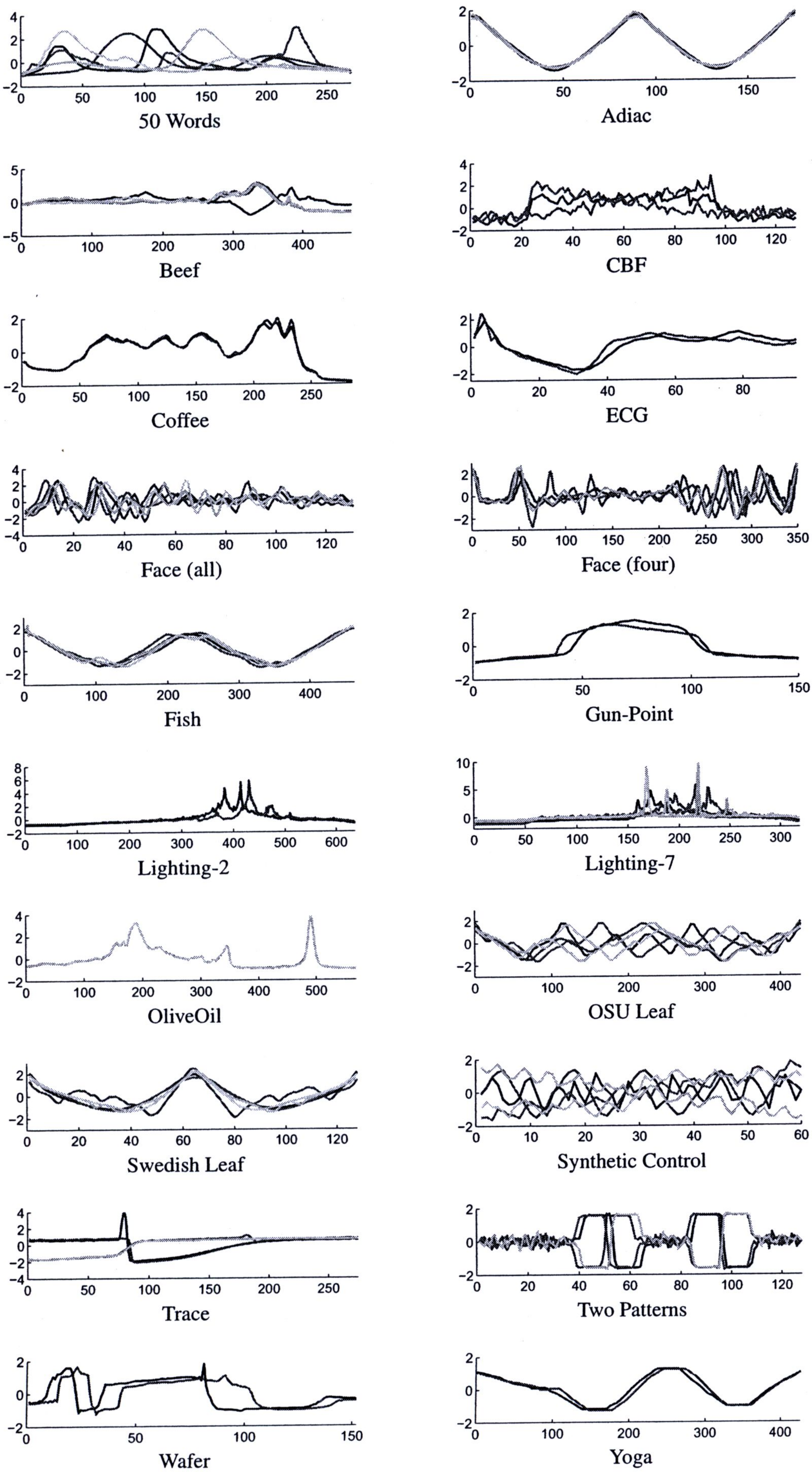


Figure F.9: Averaged results of some classes from Incremental Shape-based Averaging with CDTW when $\alpha = 1$.

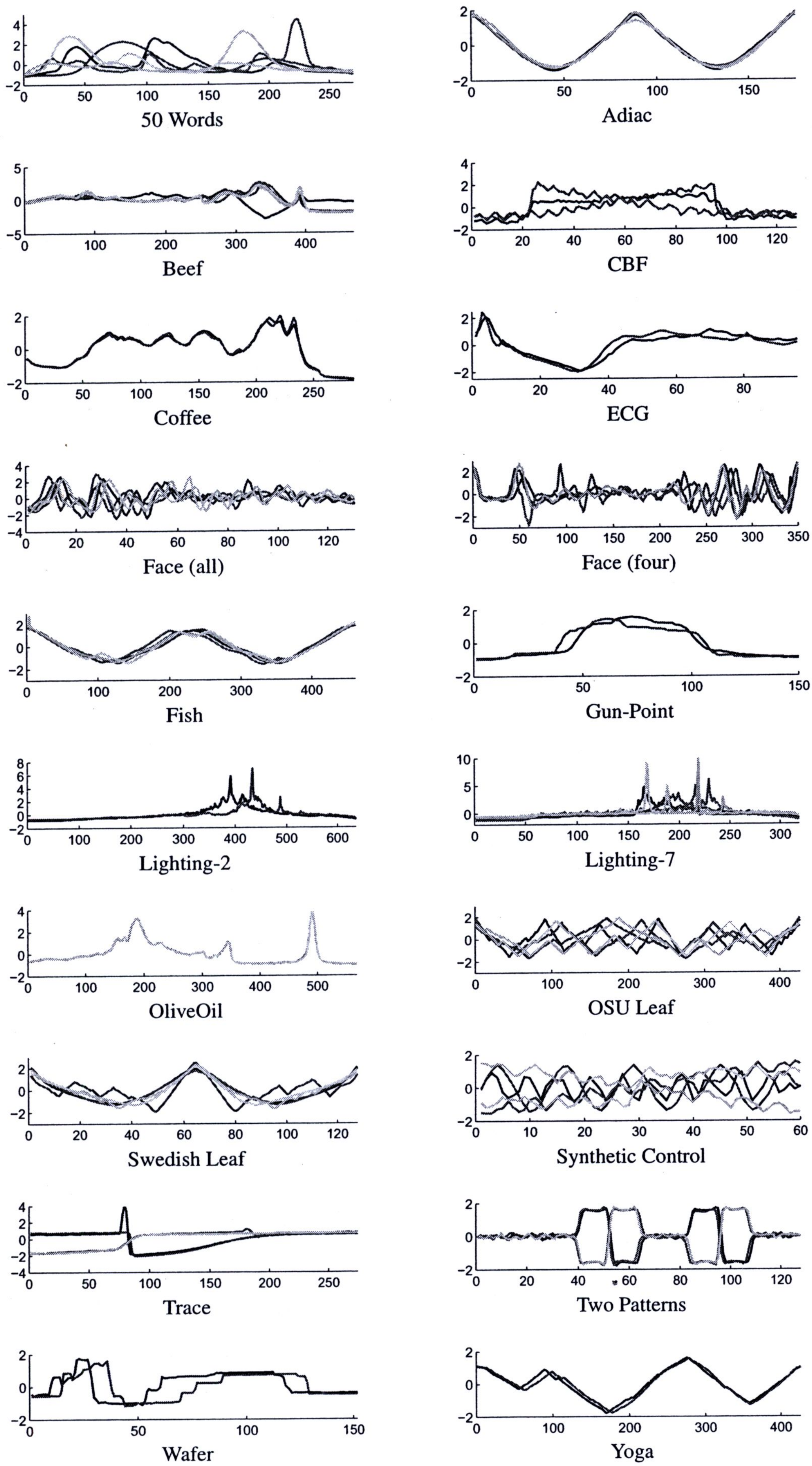


Figure F.10: Averaged results of some classes from Incremental Shape-based Averaging with CDTW when α is 25% of total number of each class.

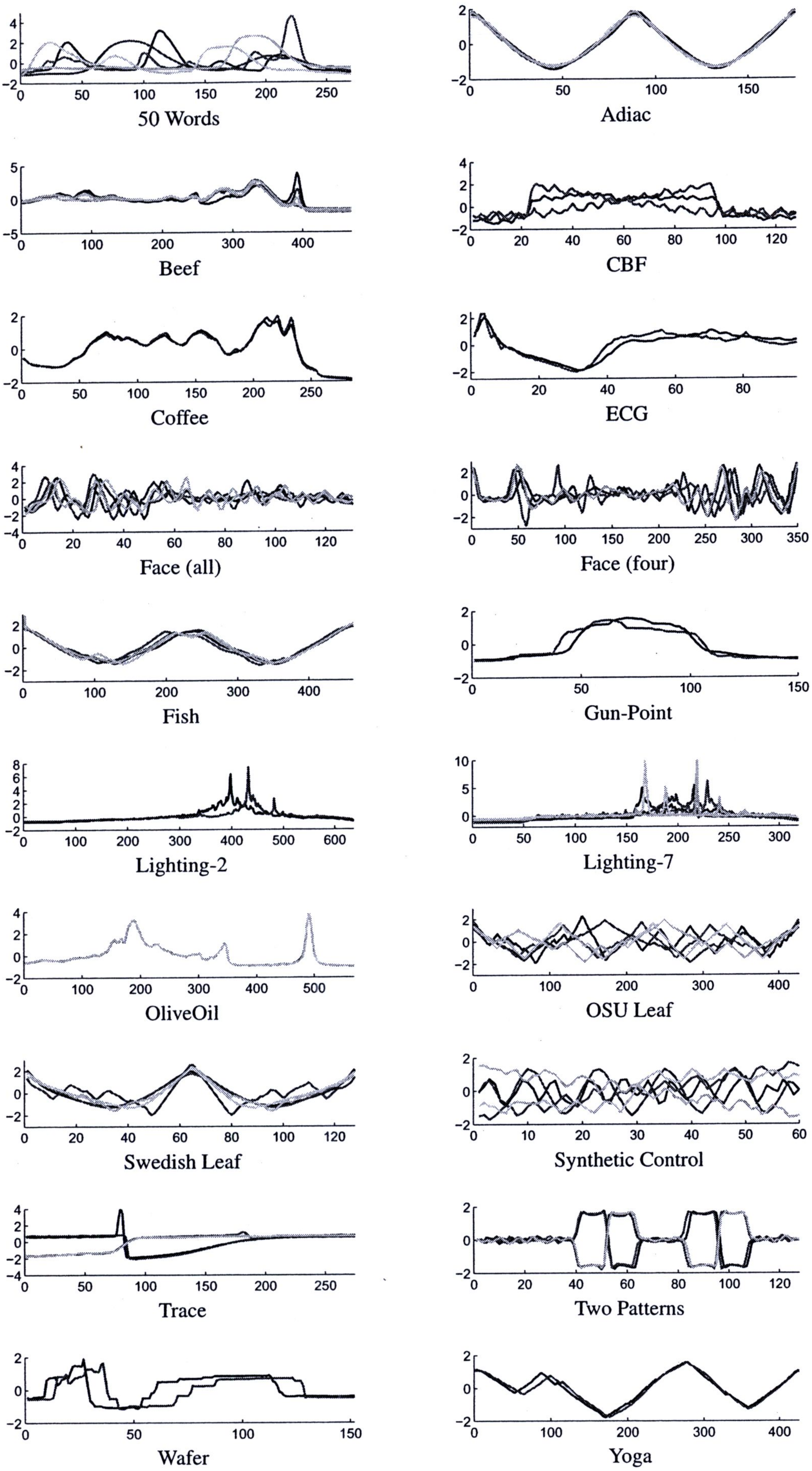


Figure F.11: Averaged results of some classes from Incremental Shape-based Averaging with CDTW when α is 50% of total number of each class.

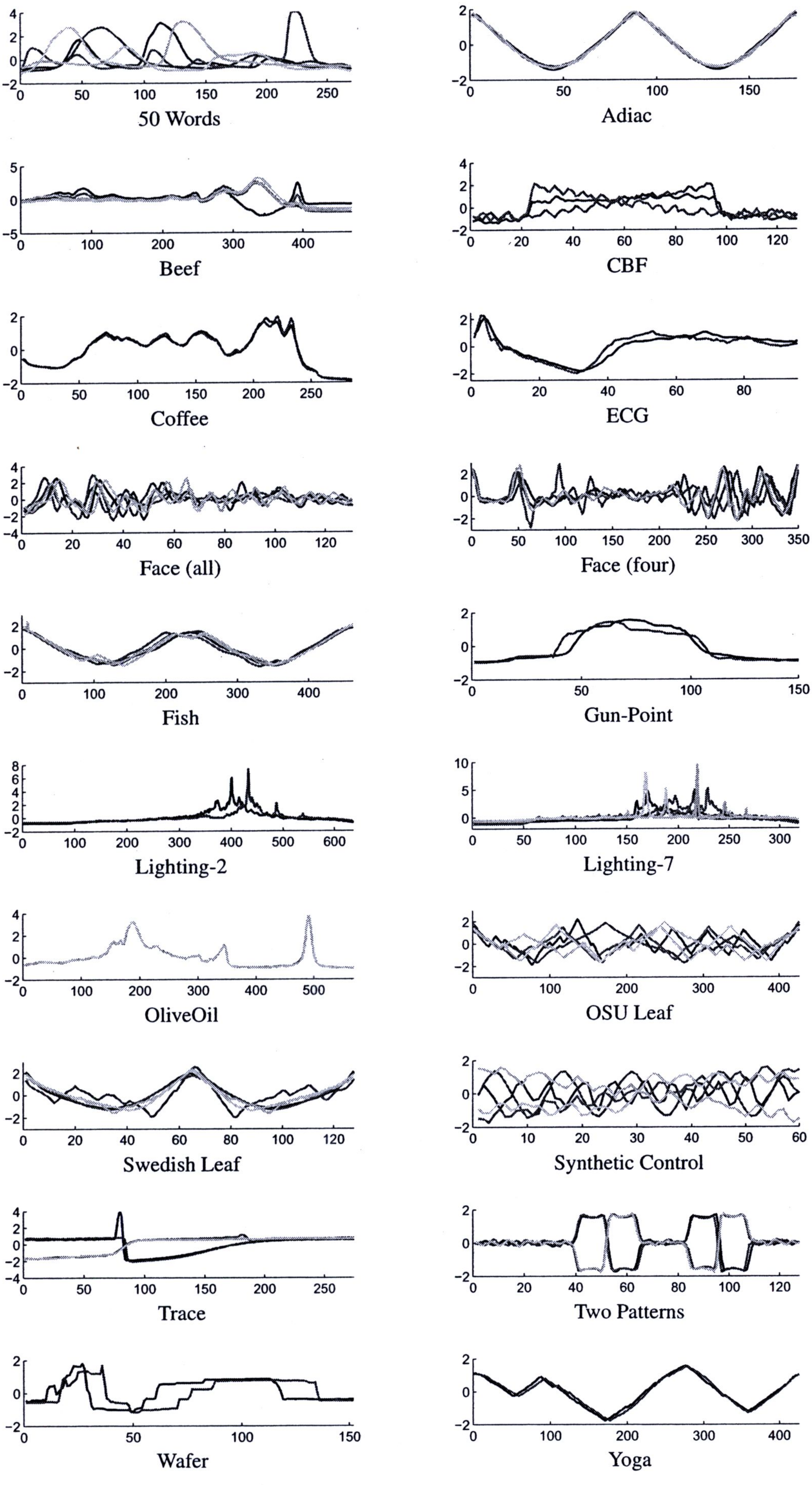


Figure F.12: Averaged results of some classes from Incremental Shape-based Averaging with CDTW when α is 100% of total number of each class.

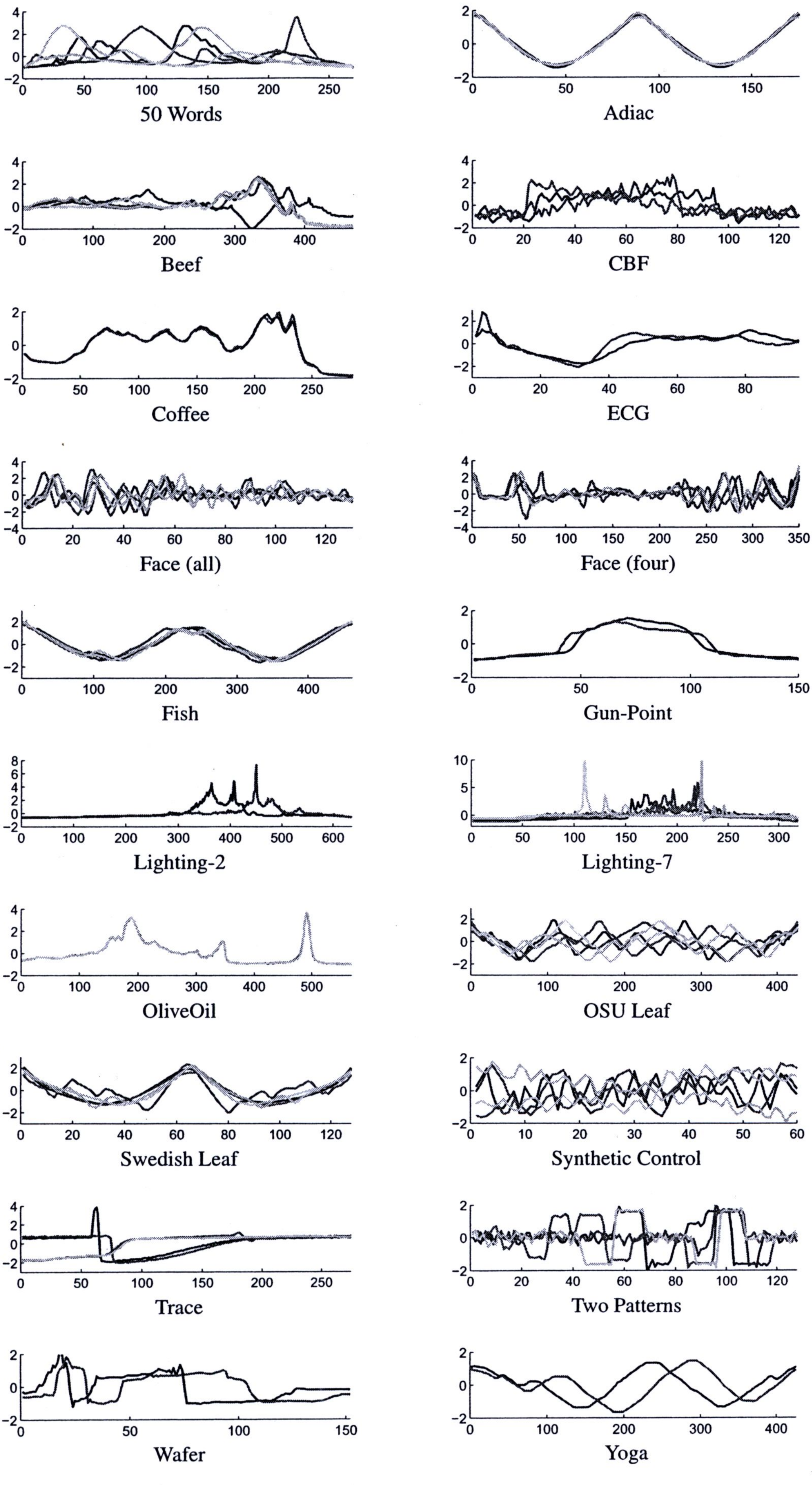


Figure F.13: Averaged results of some classes from Incremental Shape-based Averaging with ICDTW when $\alpha = 1$.

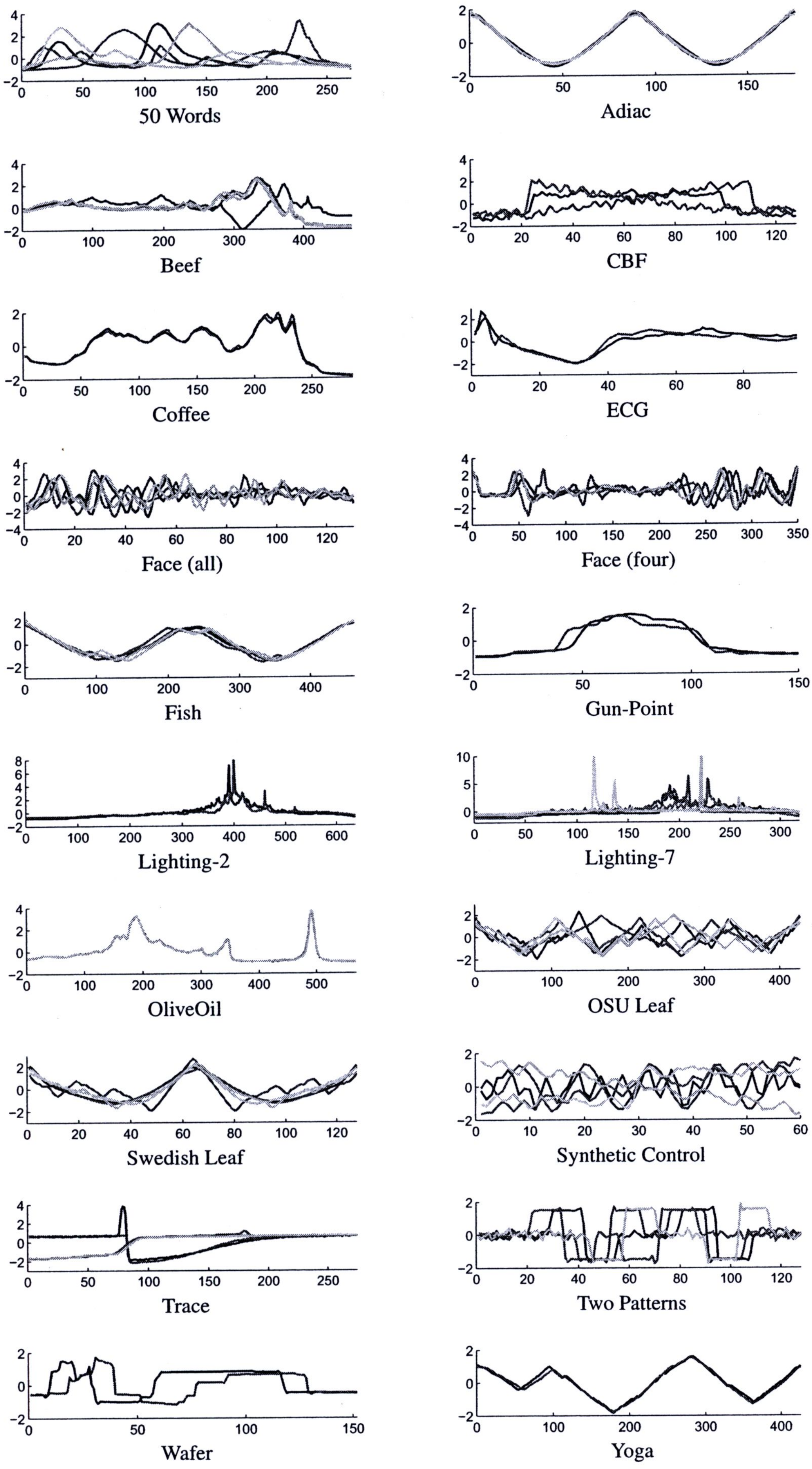


Figure F.14: Averaged results of some classes from Incremental Shape-based Averaging with ICDTW when α is 25% of total number of each class.

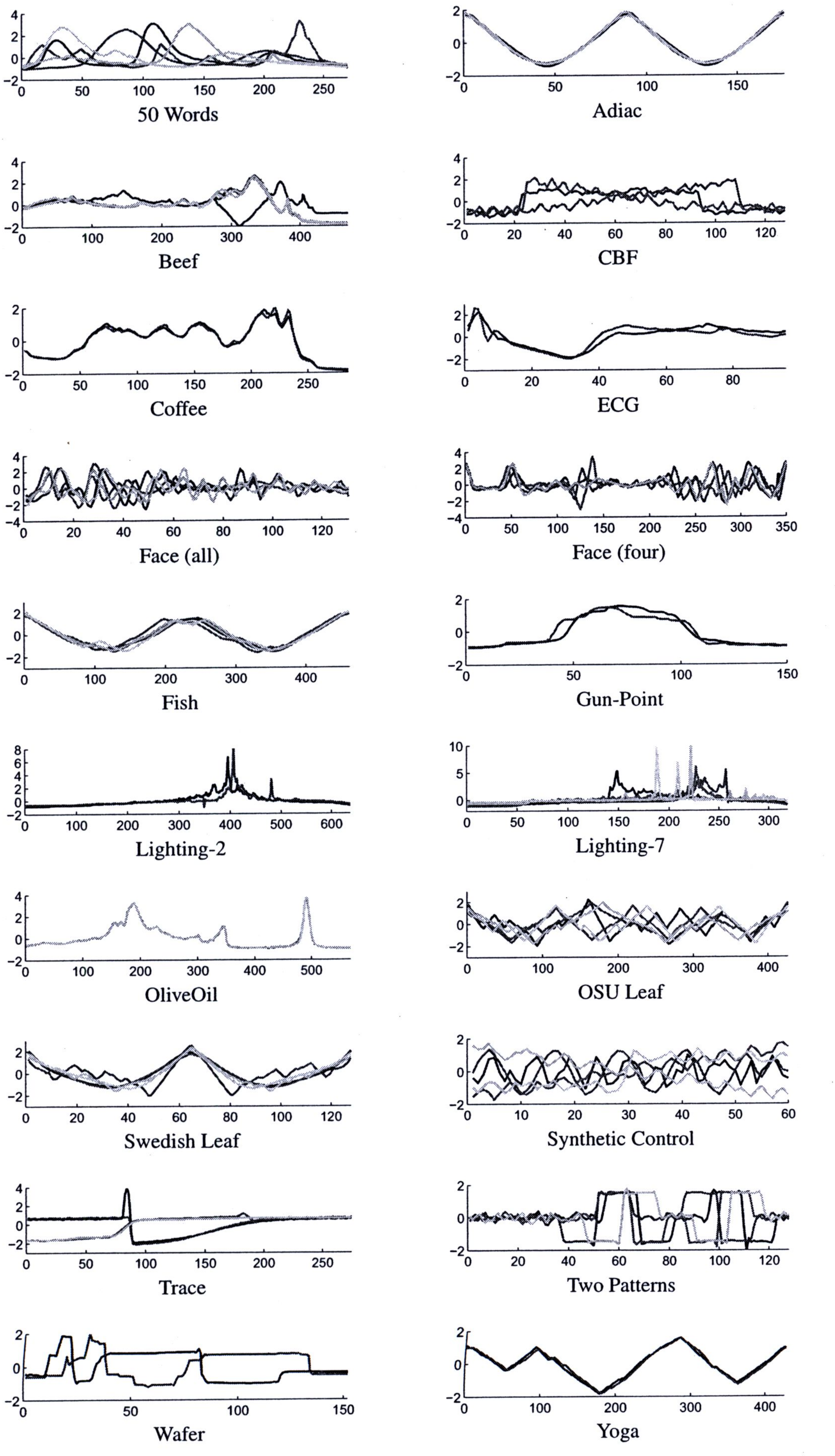


Figure F.15: Averaged results of some classes from Incremental Shape-based Averaging with ICDTW when α is 50% of total number of each class.

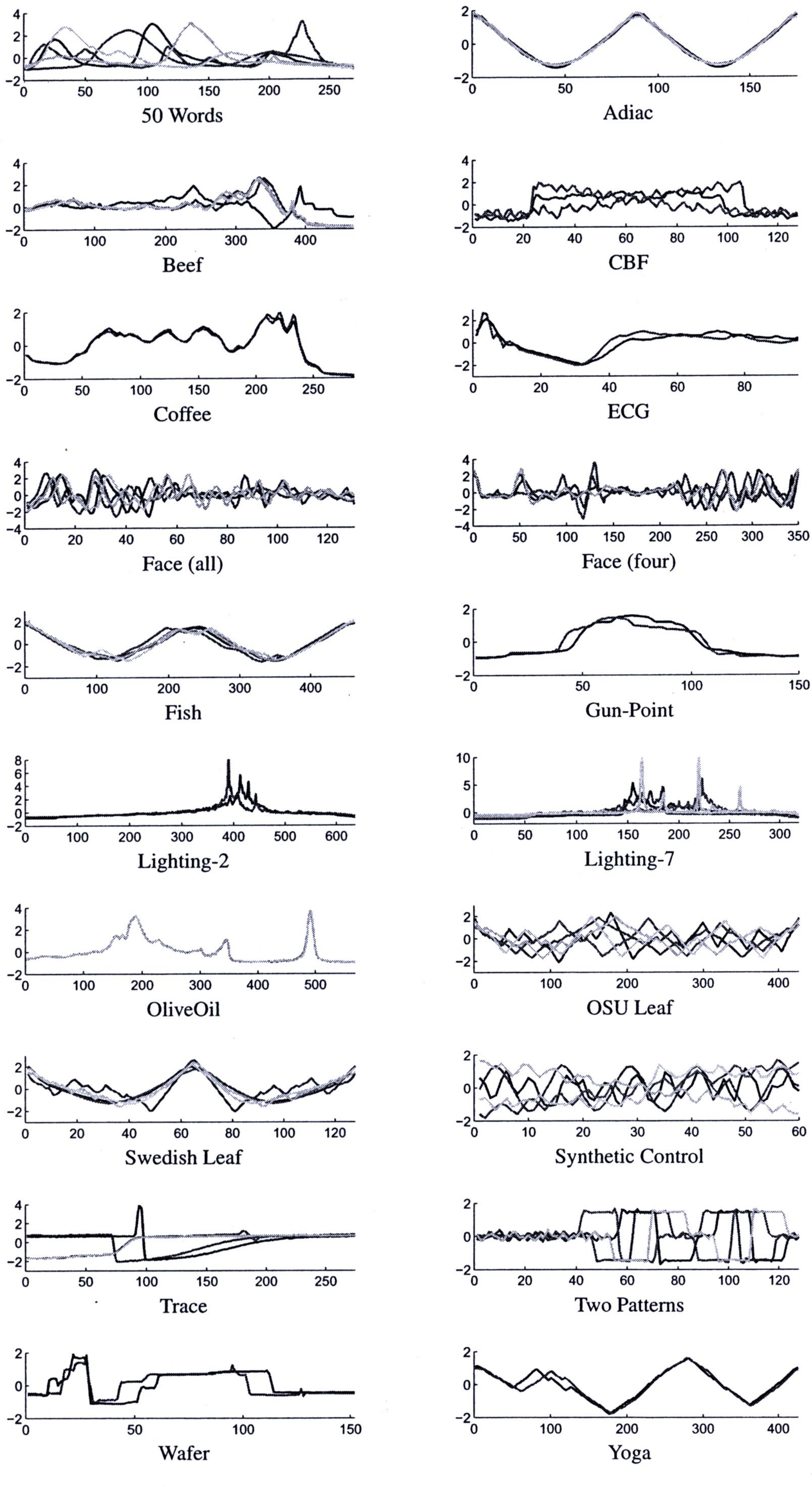


Figure F.16: Averaged results of some classes from Incremental Shape-based Averaging with ICDTW when α is 100% of total number of each class.

APPENDIX G

COMPLETE EXPERIMENTAL RESULTS OF THE FIRST EXPERIMENT IN CHAPTER VI

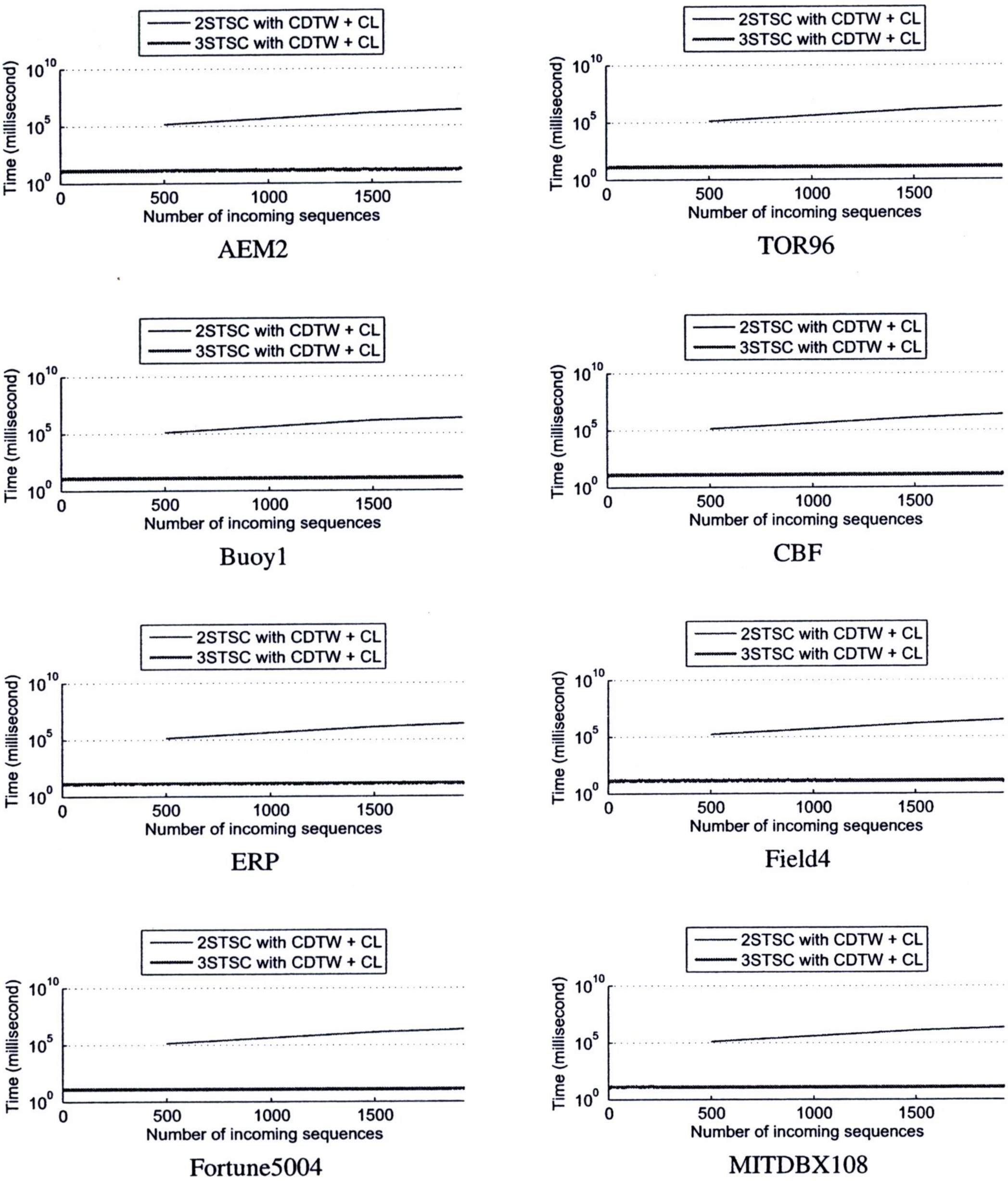


Figure G.1: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 3$ and $w = 64$.

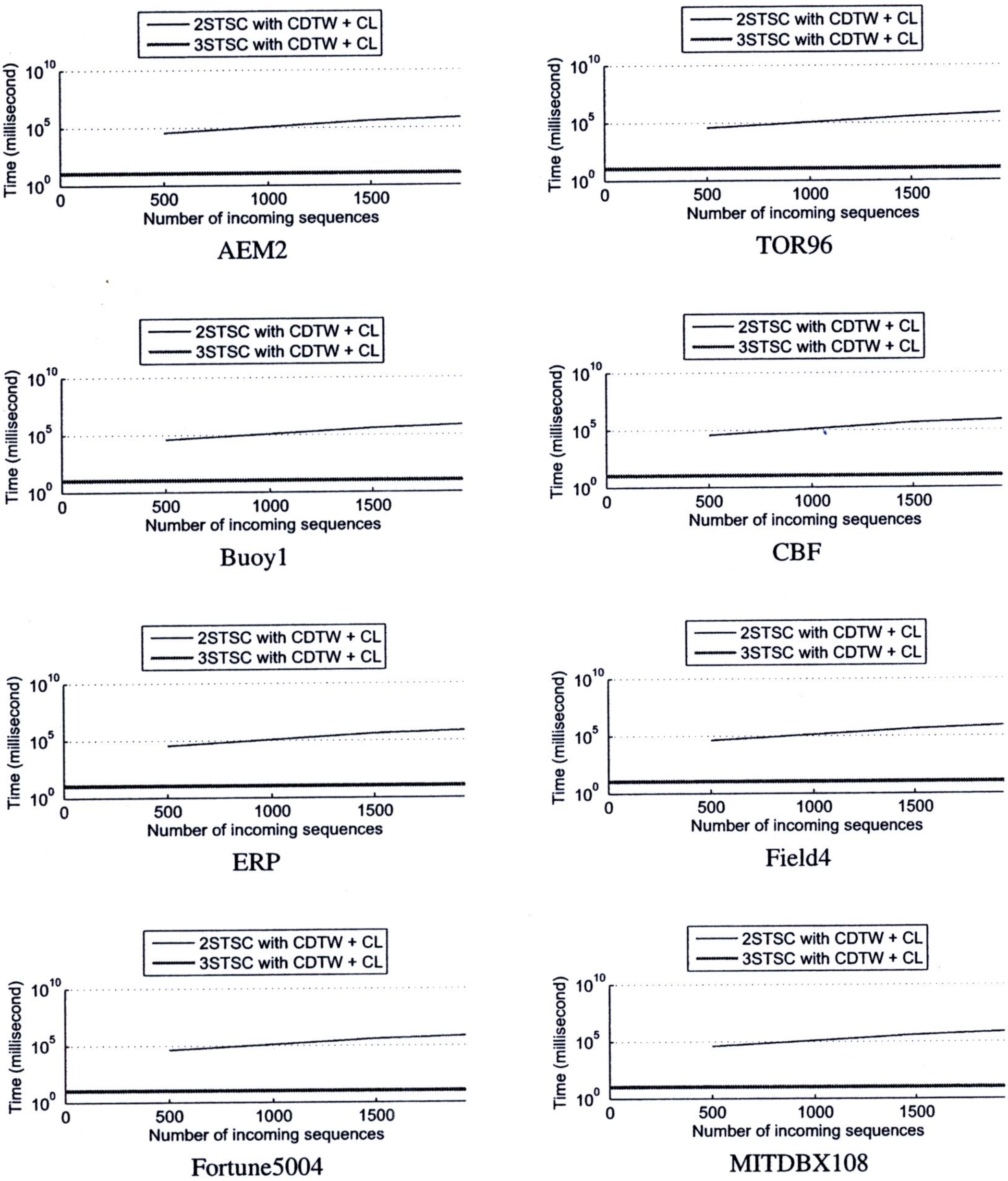


Figure G.2: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 3$ and $w = 32$

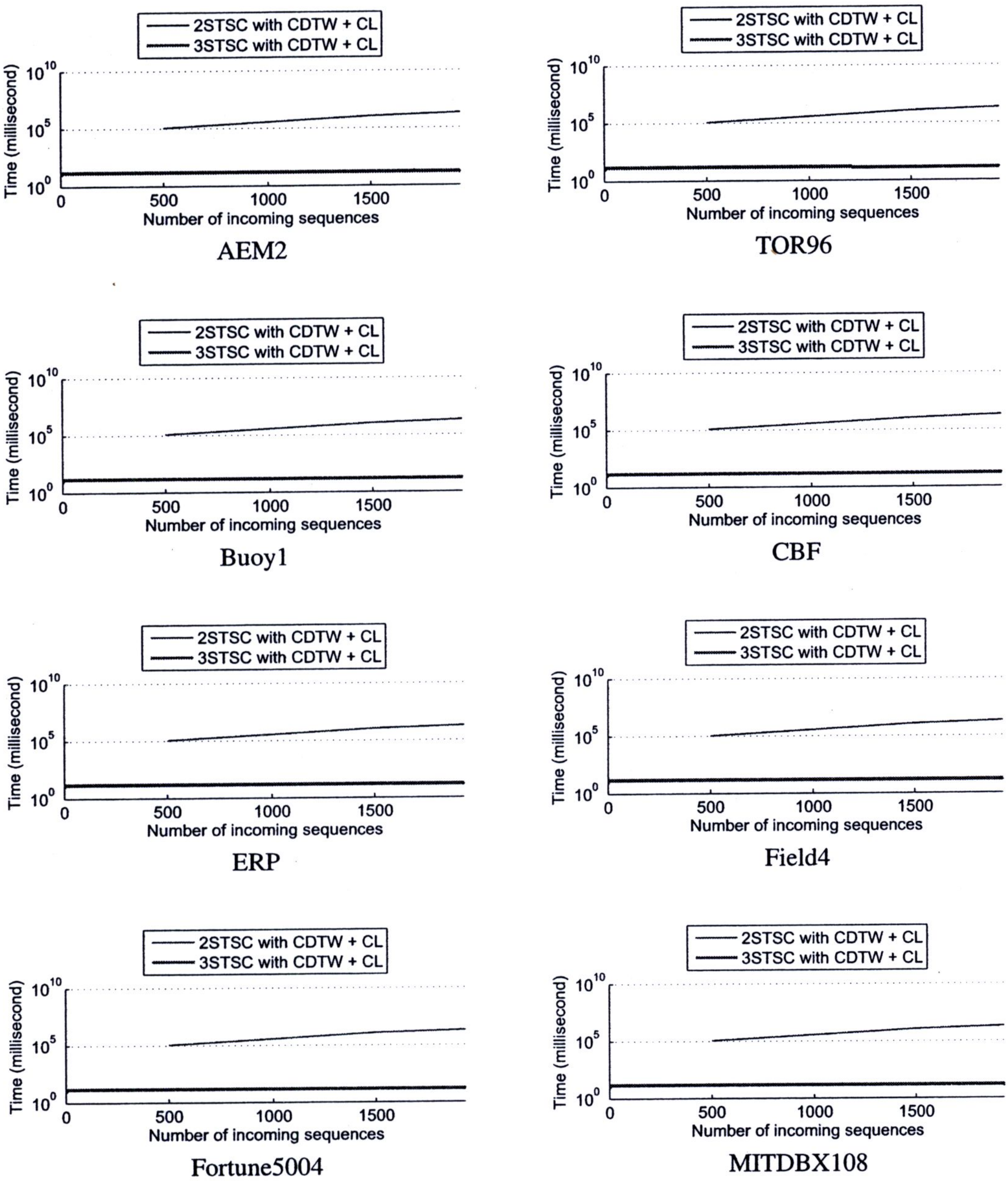


Figure G.3: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 5$ and $w = 64$.

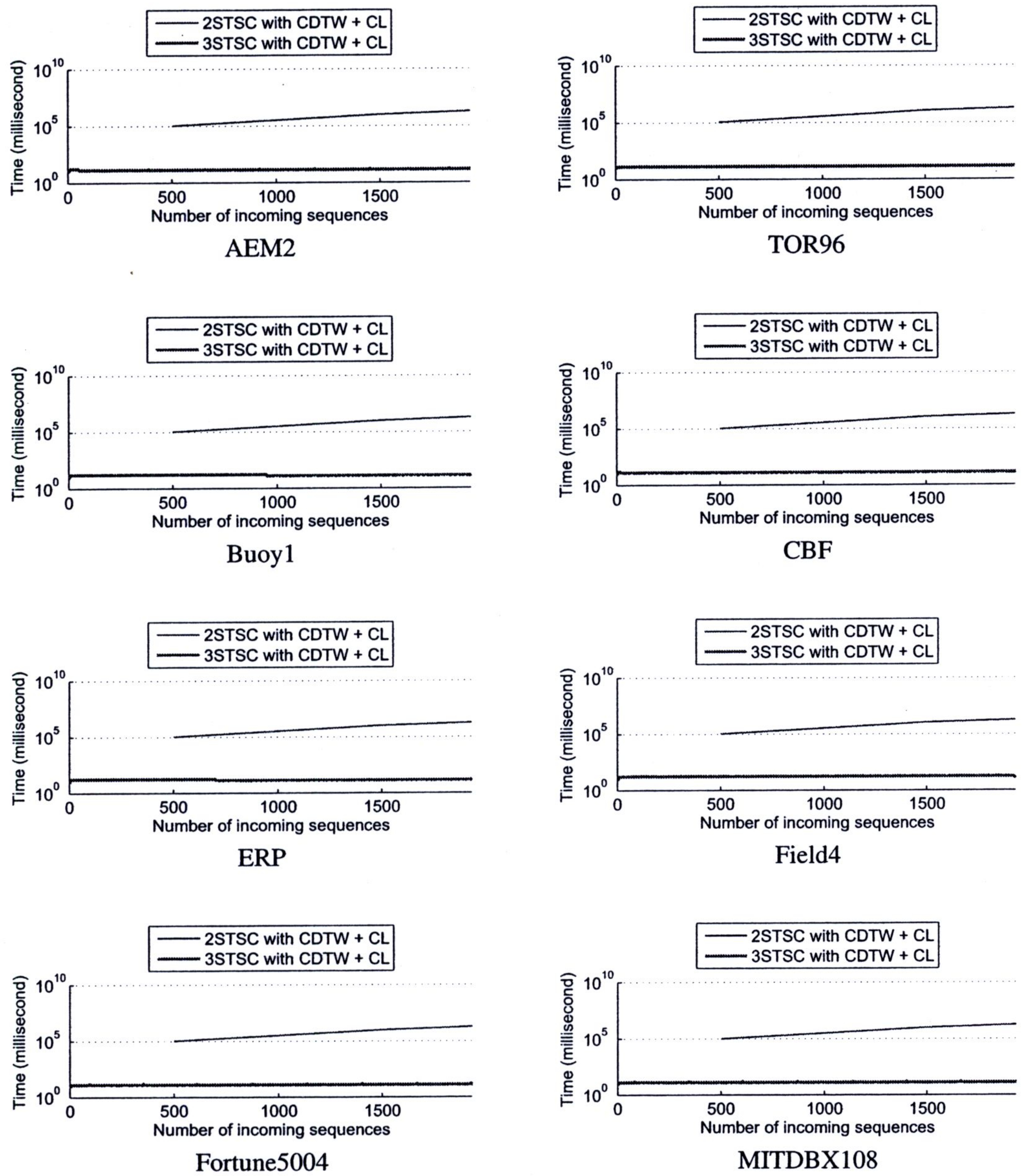


Figure G.4: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 7$ and $w = 64$.

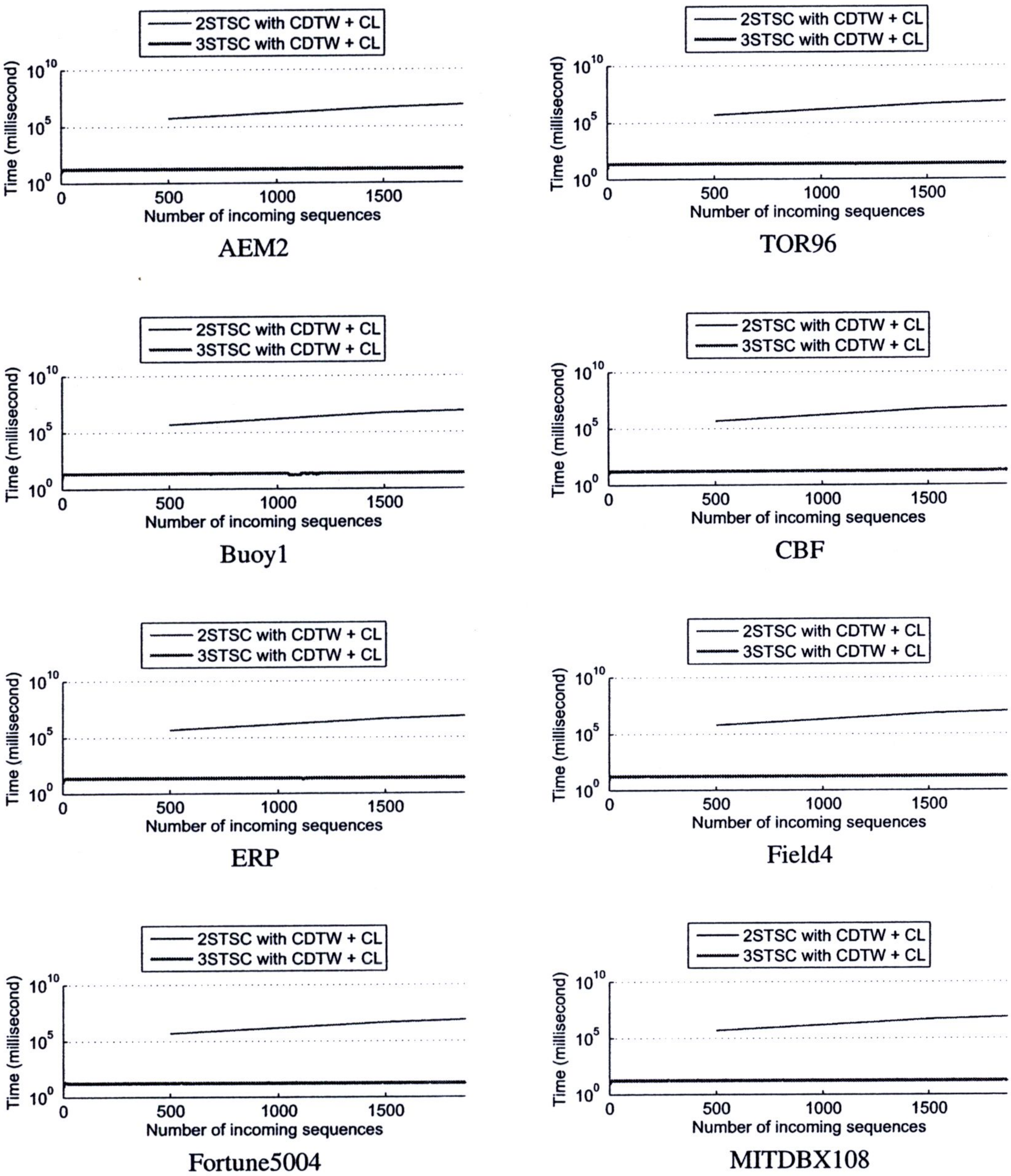


Figure G.5: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 3$ and $w = 128$.

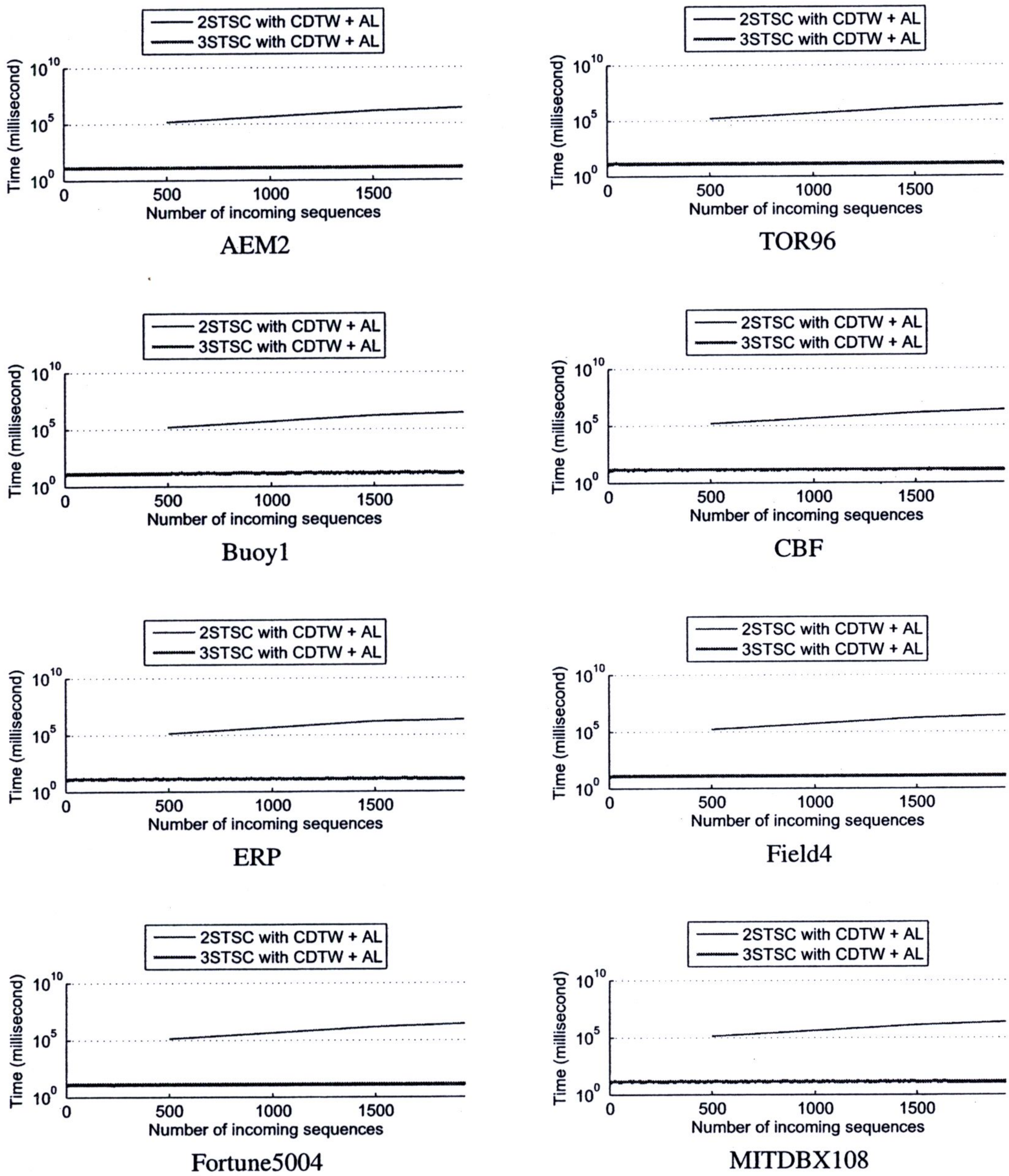


Figure G.6: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 3$ and $w = 64$.

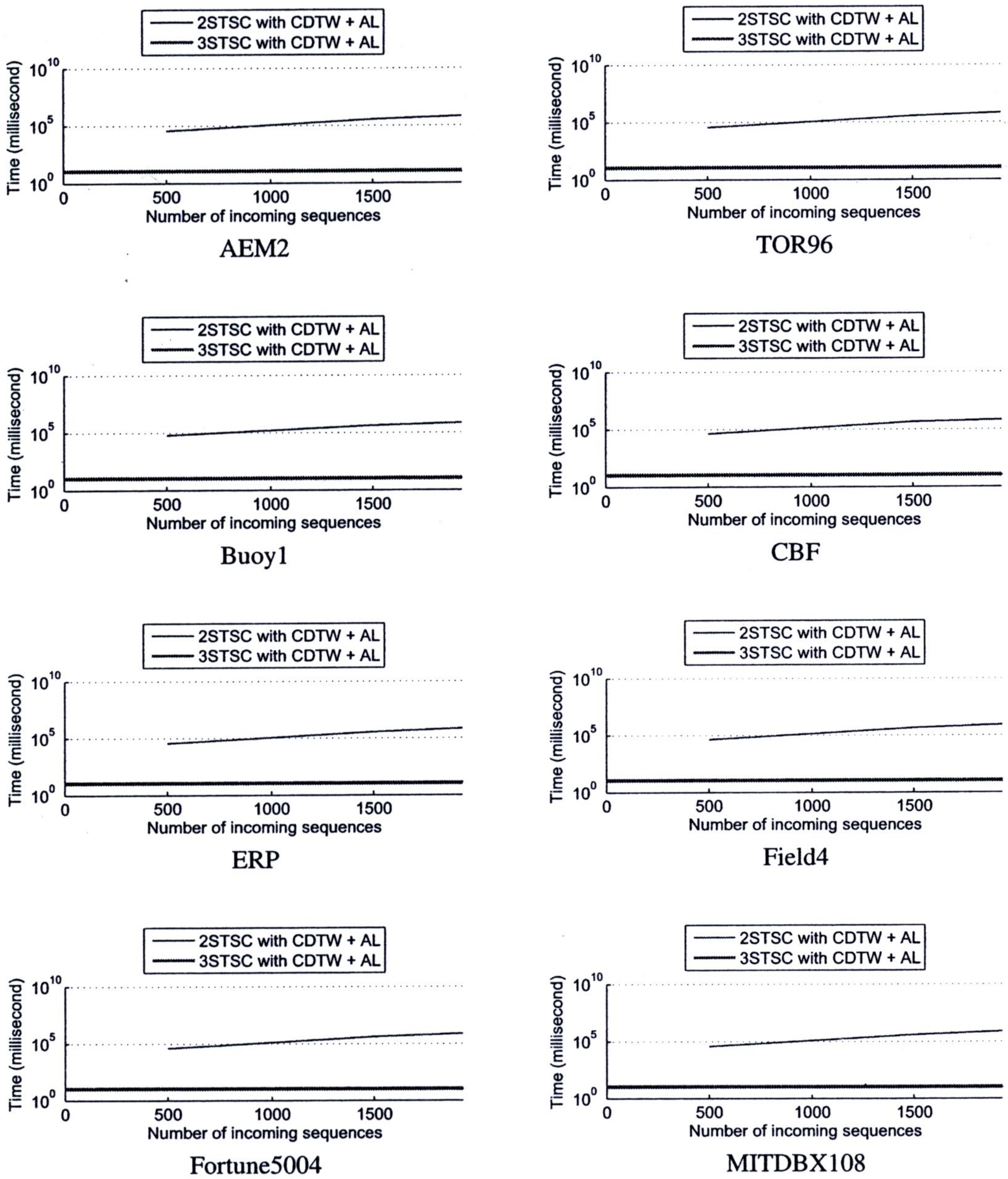


Figure G.7: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 3$ and $w = 32$.

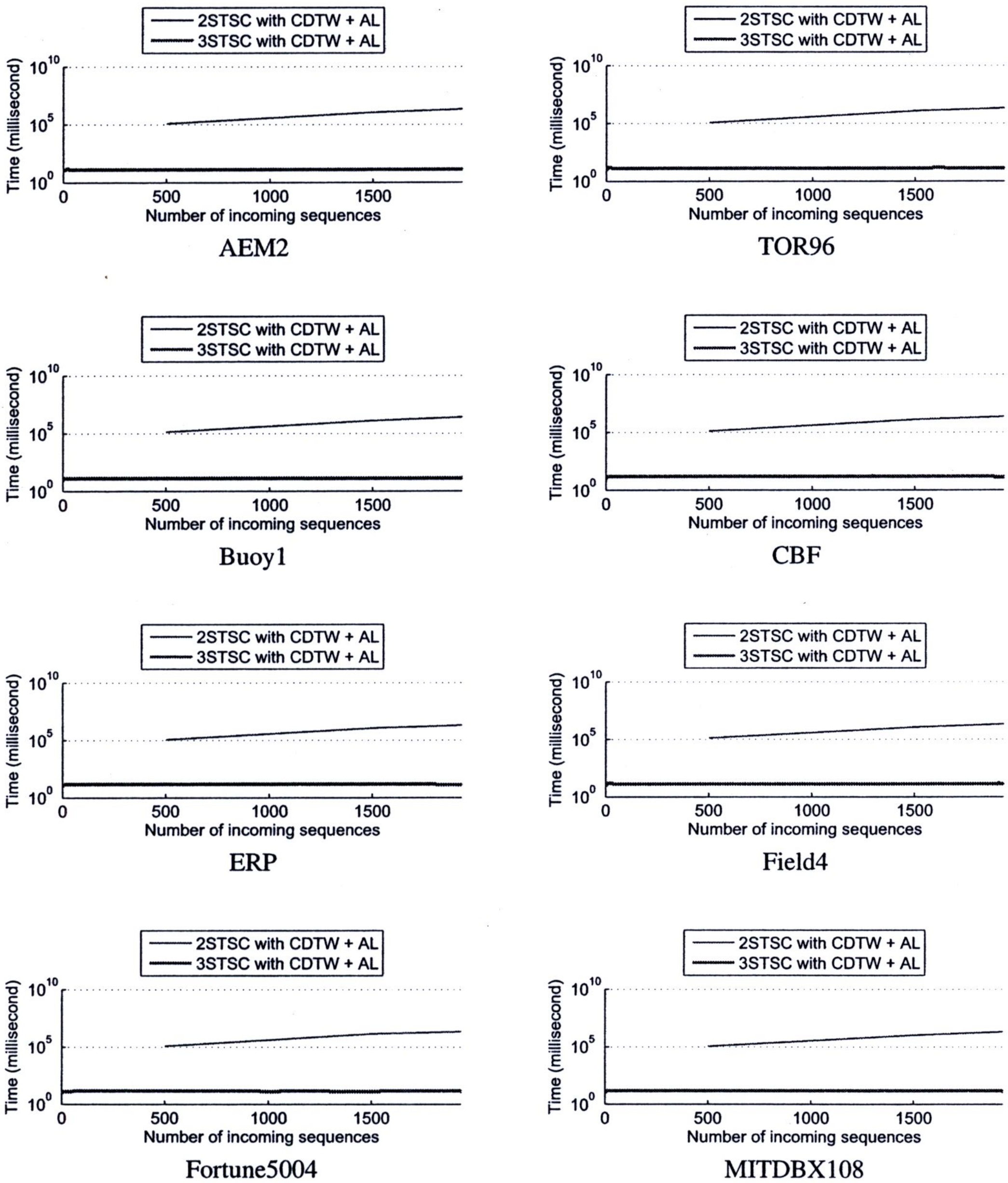


Figure G.8: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 5$ and $w = 64$.

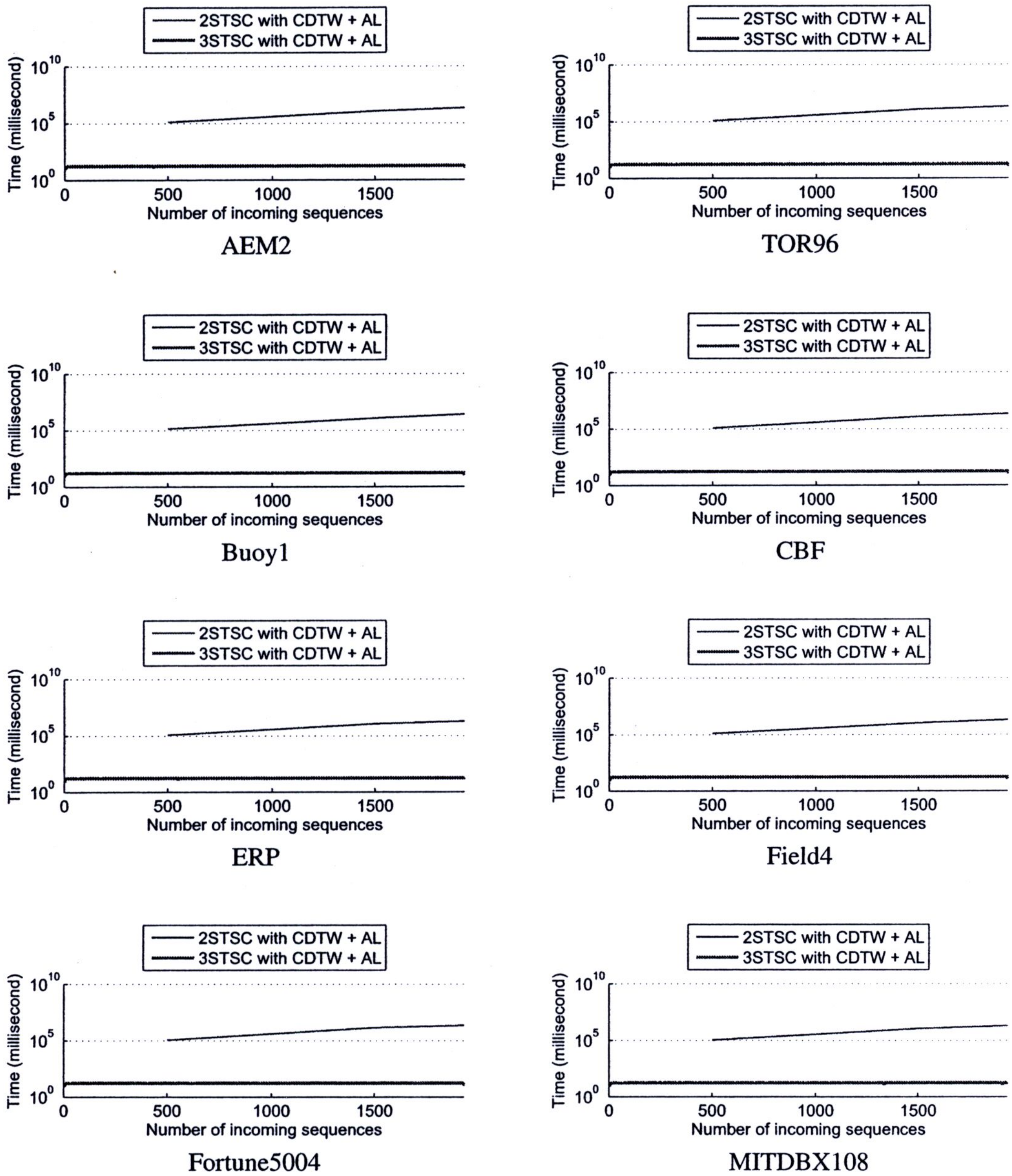


Figure G.9: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 7$ and $w = 64$.

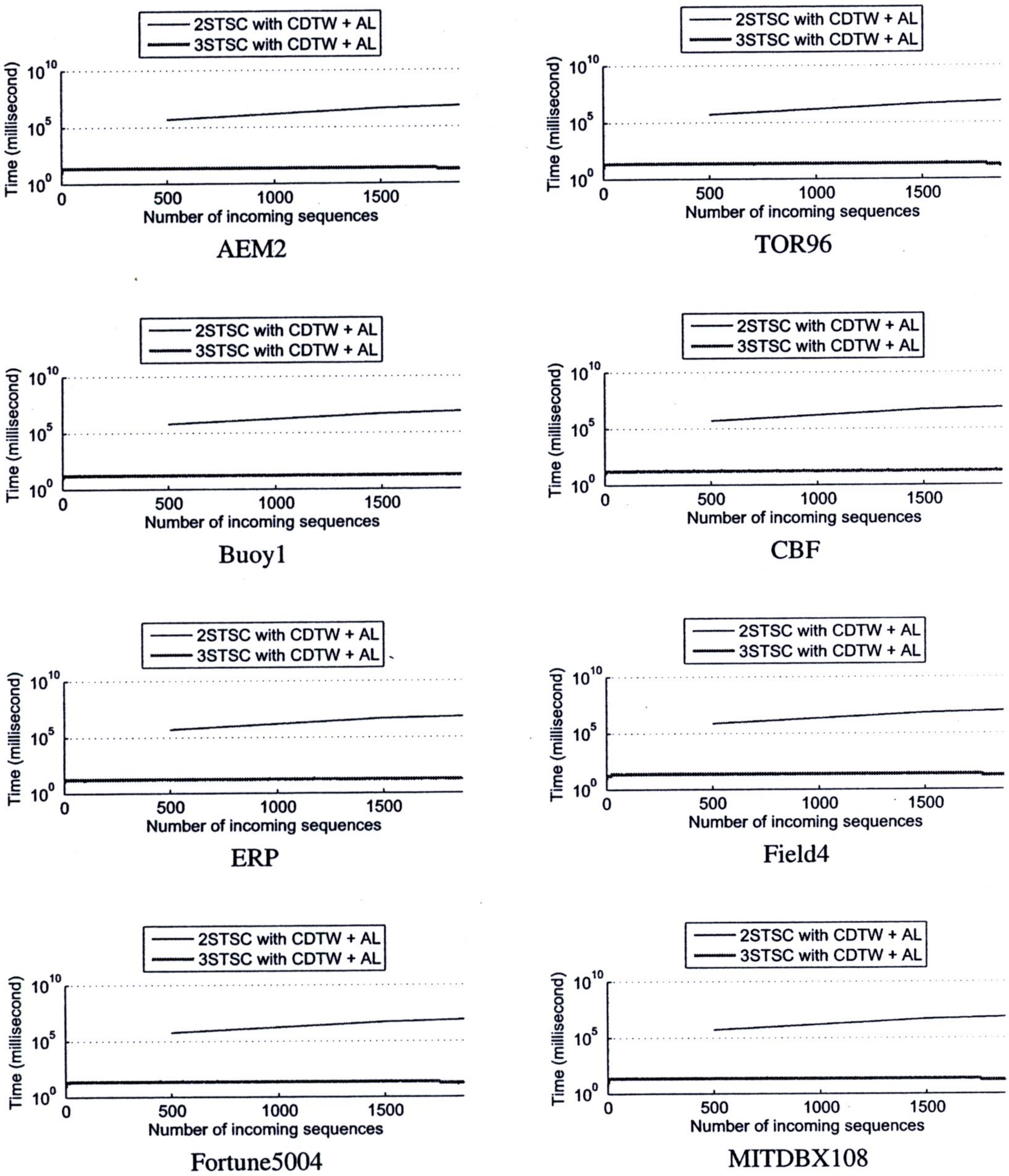
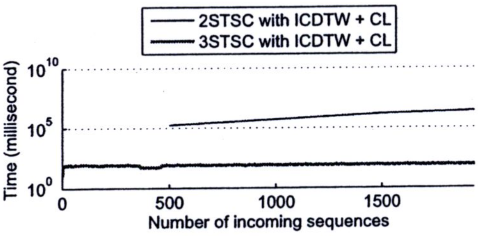
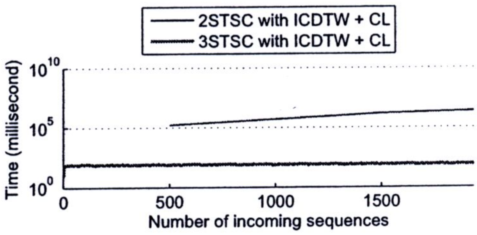


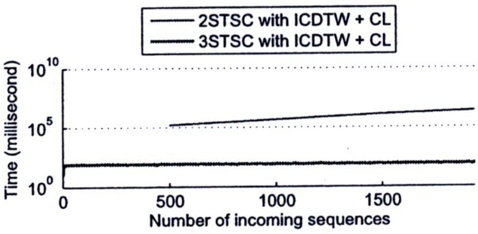
Figure G.10: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 3$ and $w = 128$.



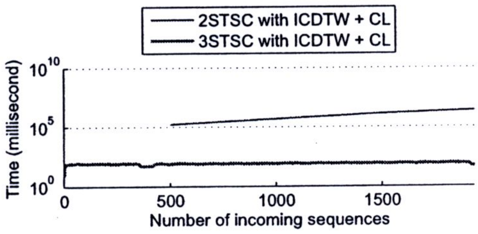
AEM2



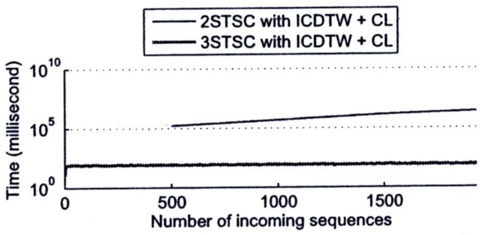
TOR96



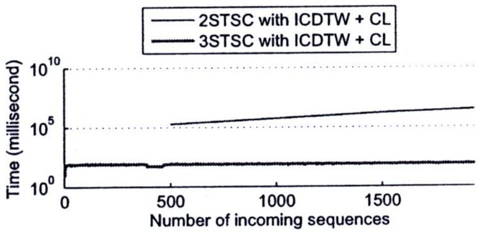
Buoy1



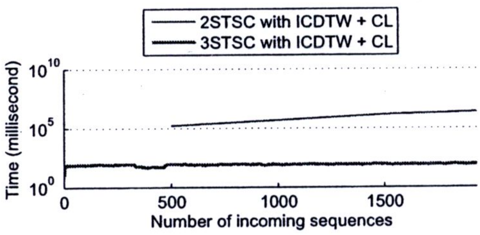
CBF



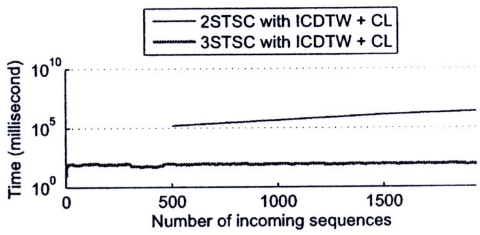
ERP



Field4



Fortune5004



MITDBX108

Figure G.11: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 3$ and $w = 64$.

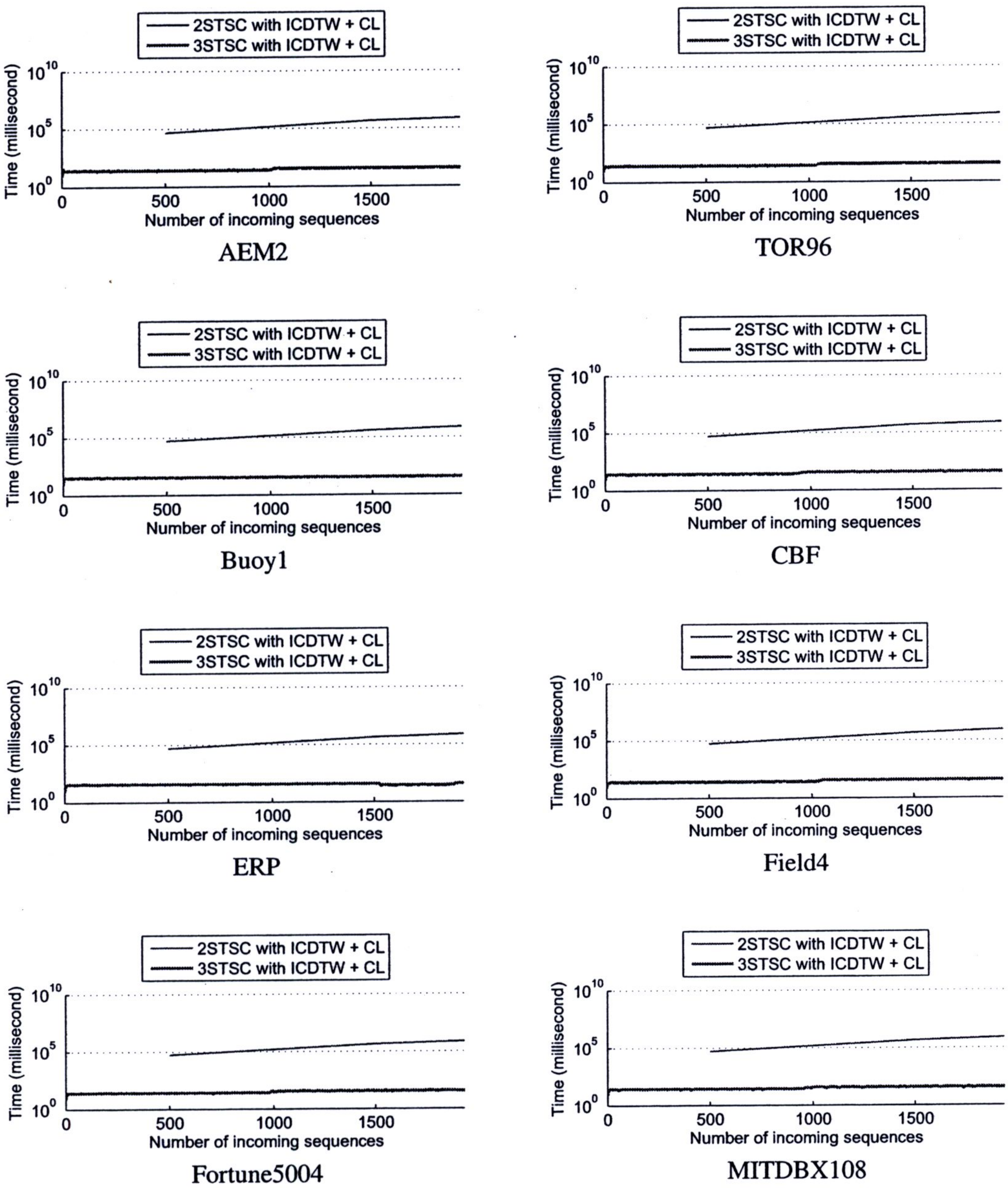
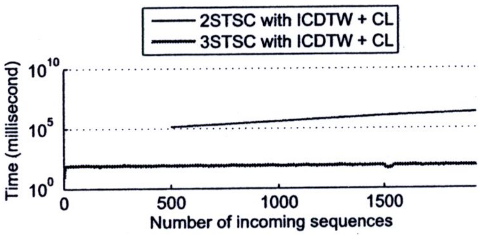
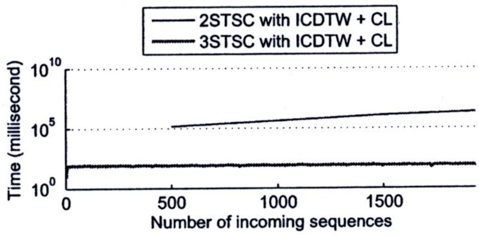


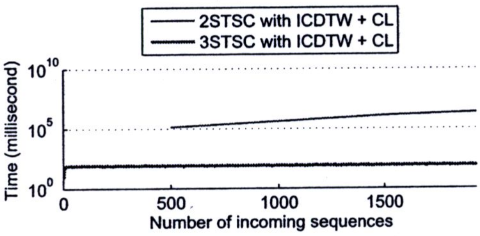
Figure G.12: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 3$ and $w = 32$.



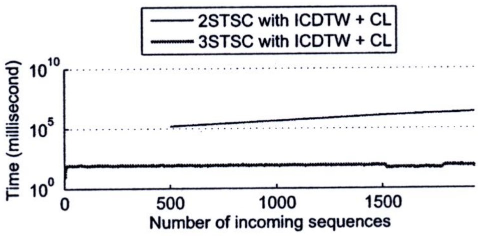
AEM2



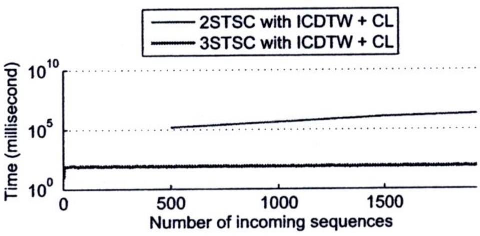
TOR96



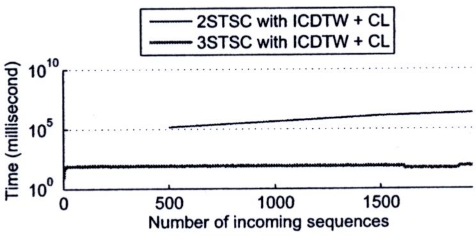
Buoy1



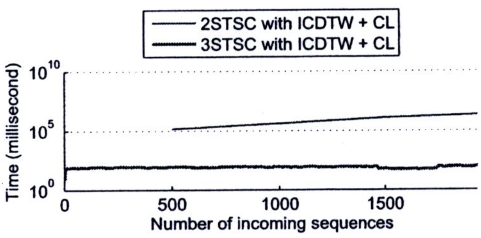
CBF



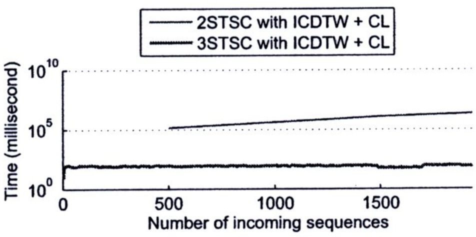
ERP



Field4



Fortune5004



MITDBX108

Figure G.13: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 5$ and $w = 64$.

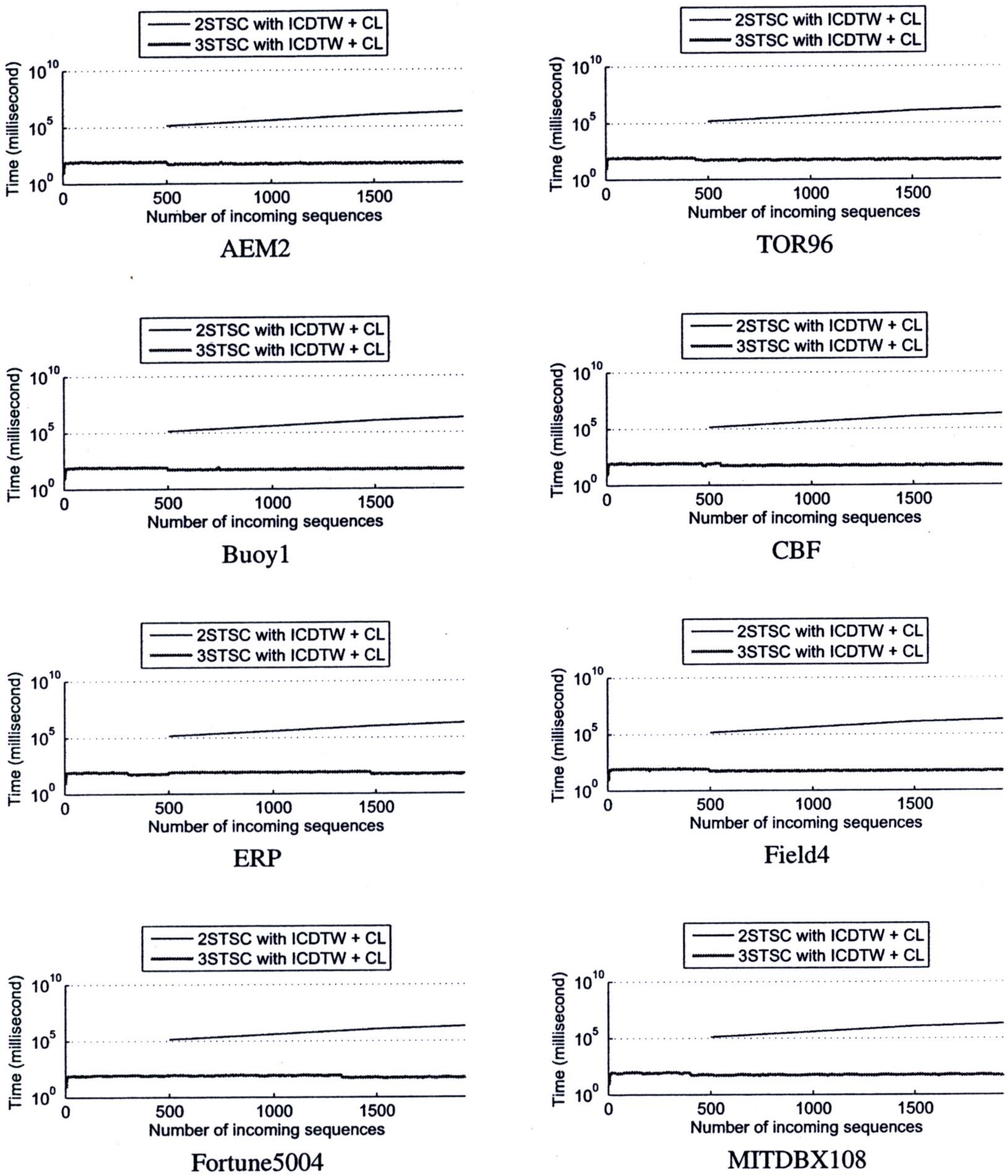
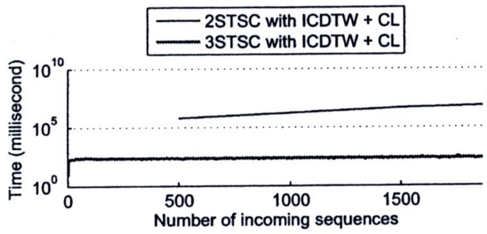
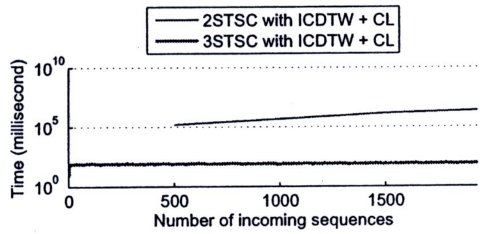


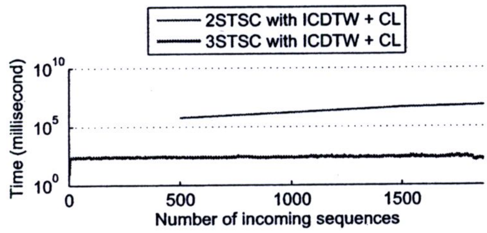
Figure G.14: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 7$ and $w = 64$.



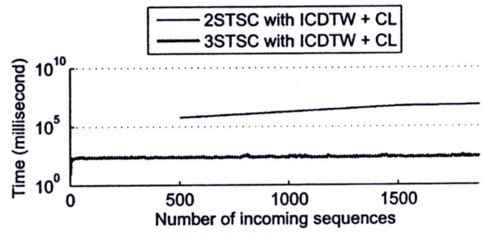
AEM2



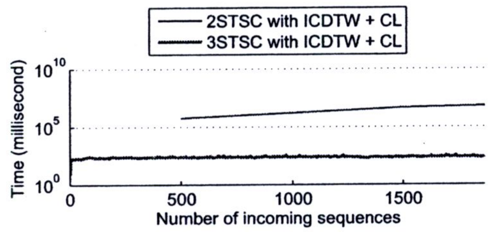
TOR96



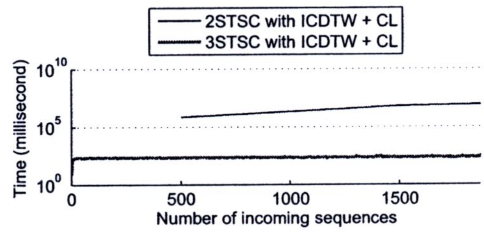
Buoy1



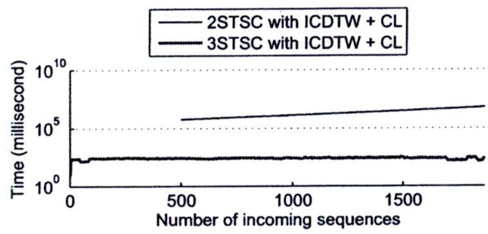
CBF



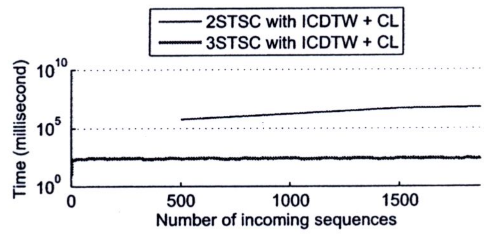
ERP



Field4



Fortune5004



MITDBX108

Figure G.15: Computational time of 3STSC and 2STSC with CDTW function and complete linkage when a new incoming sequence arrives, where $k = 3$ and $w = 128$.

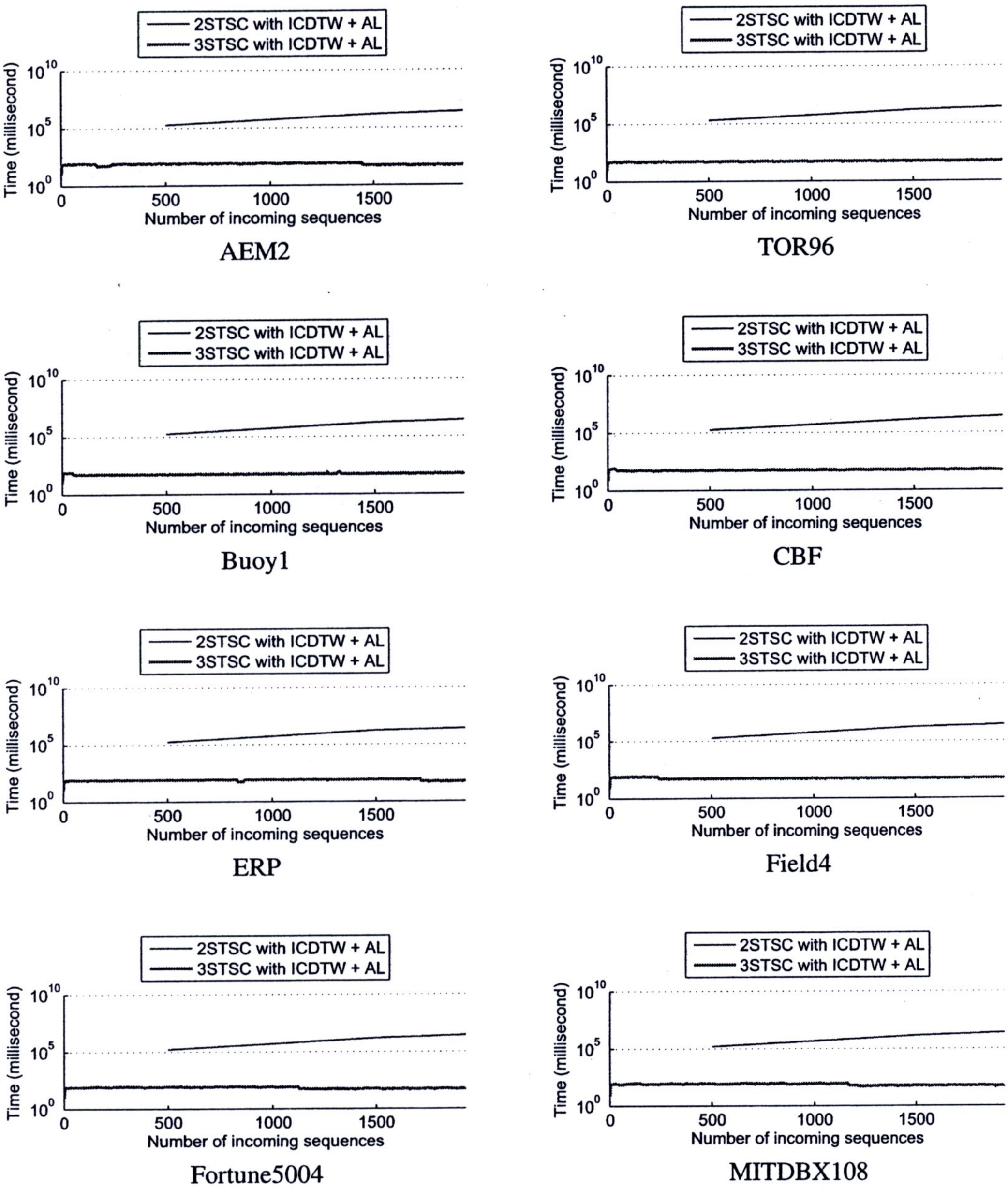
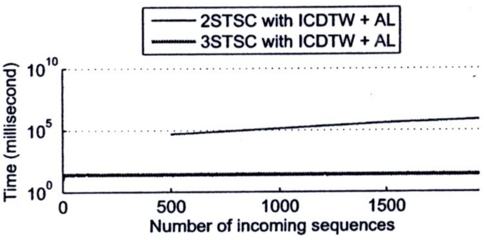
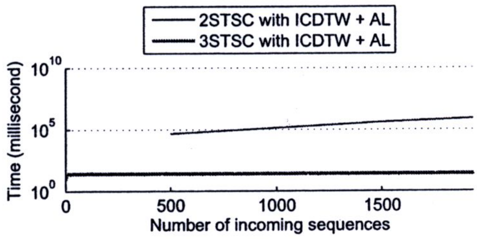


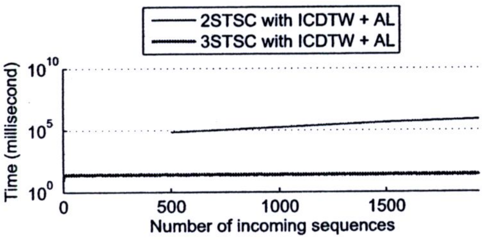
Figure G.16: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 3$ and $w = 64$



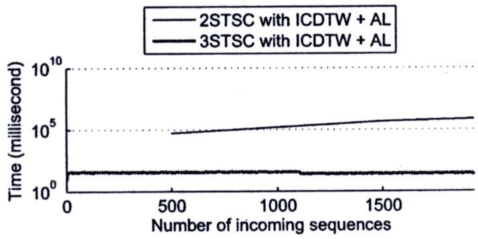
AEM2



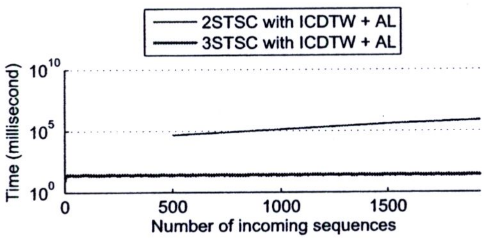
TOR96



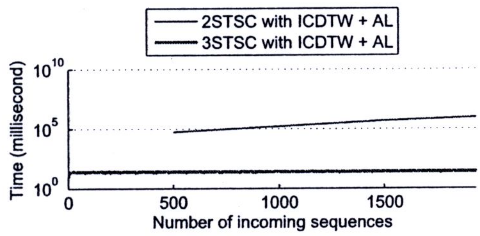
Buoy1



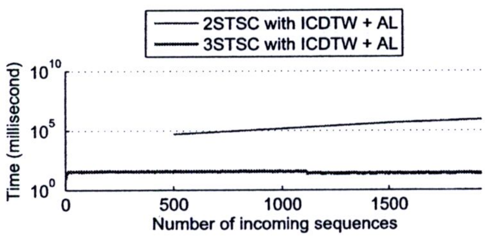
CBF



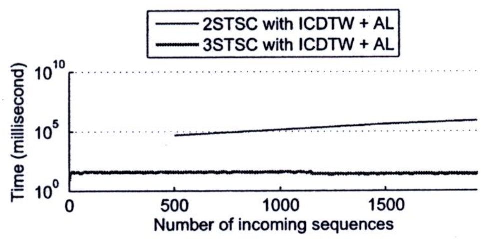
ERP



Field4



Fortune5004



MITDBX108

Figure G.17: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 3$ and $w = 32$.

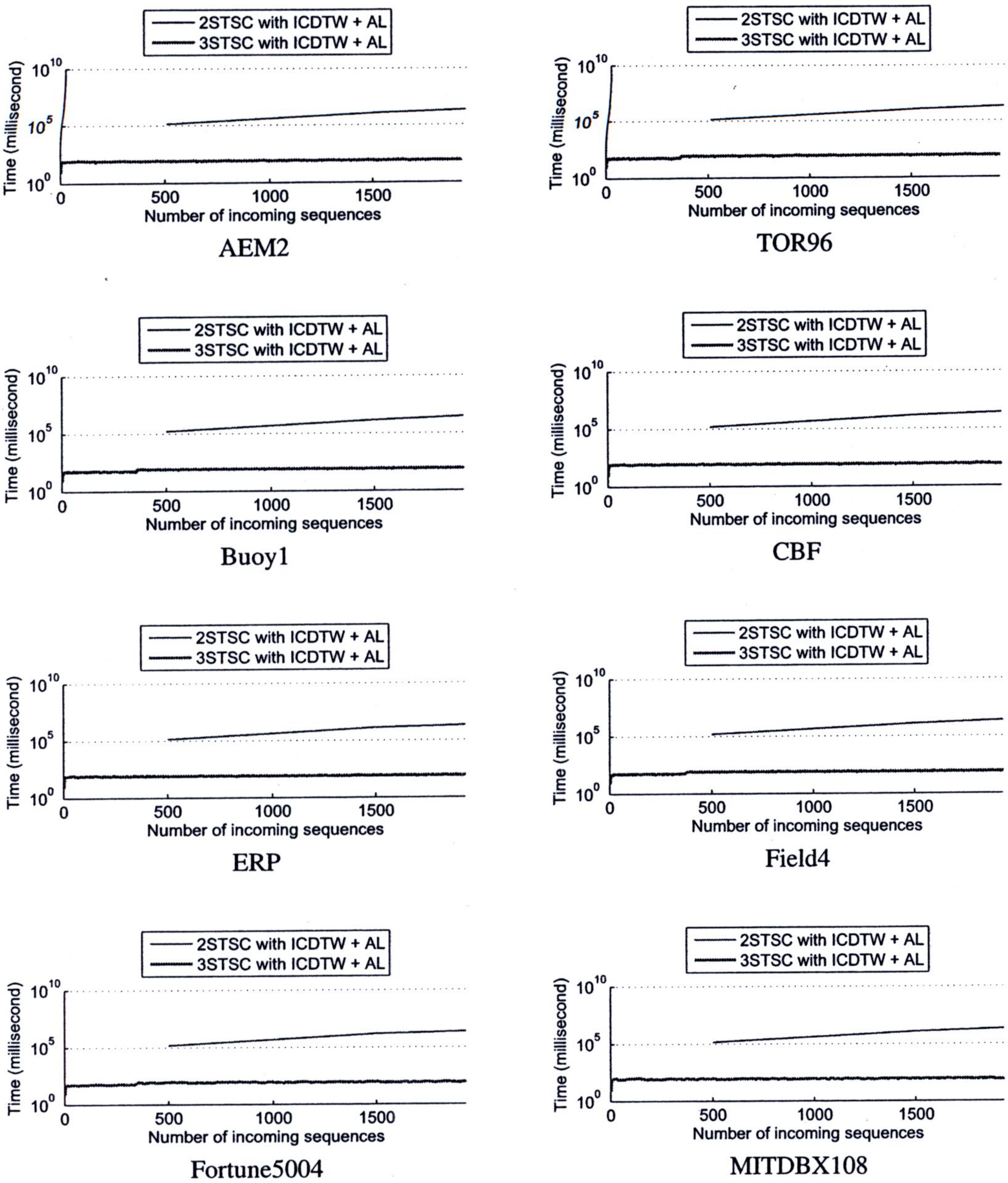


Figure G.18: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 5$ and $w = 64$.

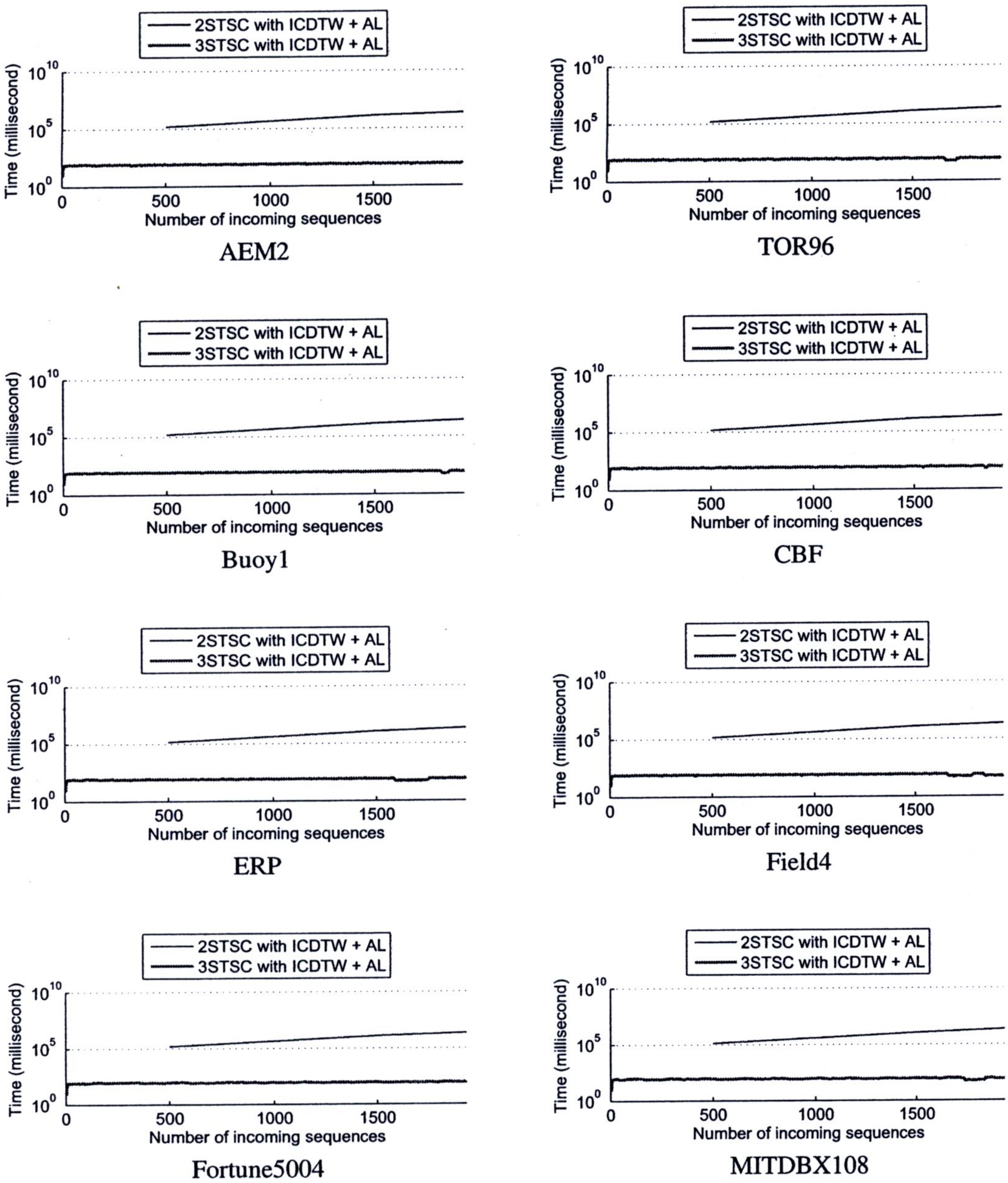
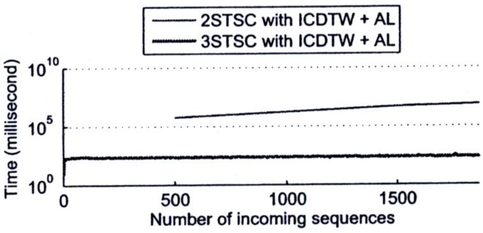
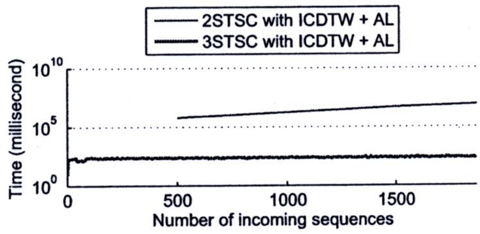


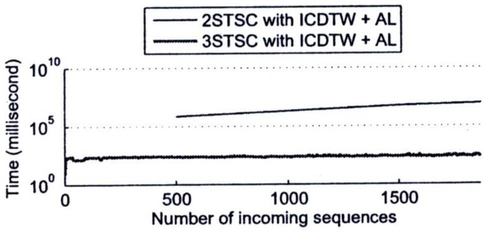
Figure G.19: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 7$ and $w = 64$.



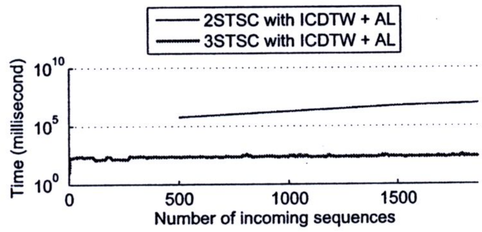
AEM2



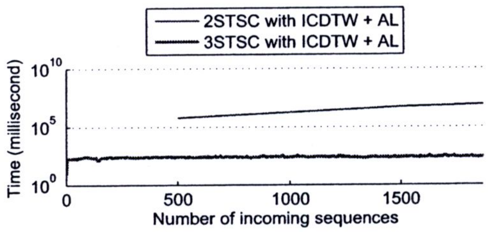
TOR96



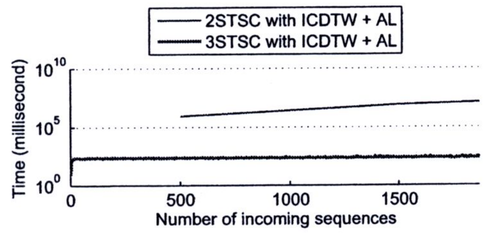
Buoy1



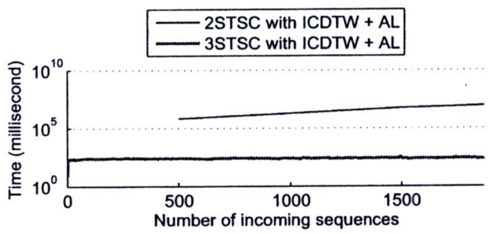
CBF



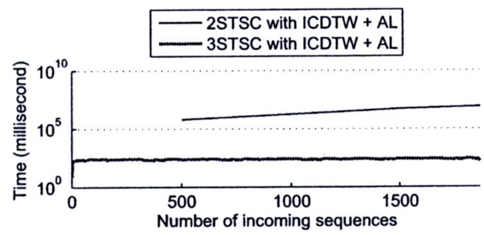
ERP



Field4



Fortune5004



MITDBX108

Figure G.20: Computational time of 3STSC and 2STSC with CDTW function and average linkage when a new incoming sequence arrives, where $k = 3$ and $w = 128$.

APPENDIX H

COMPLETE EXPERIMENTAL RESULTS OF THE SECOND EXPERIMENT IN CHAPTER VI

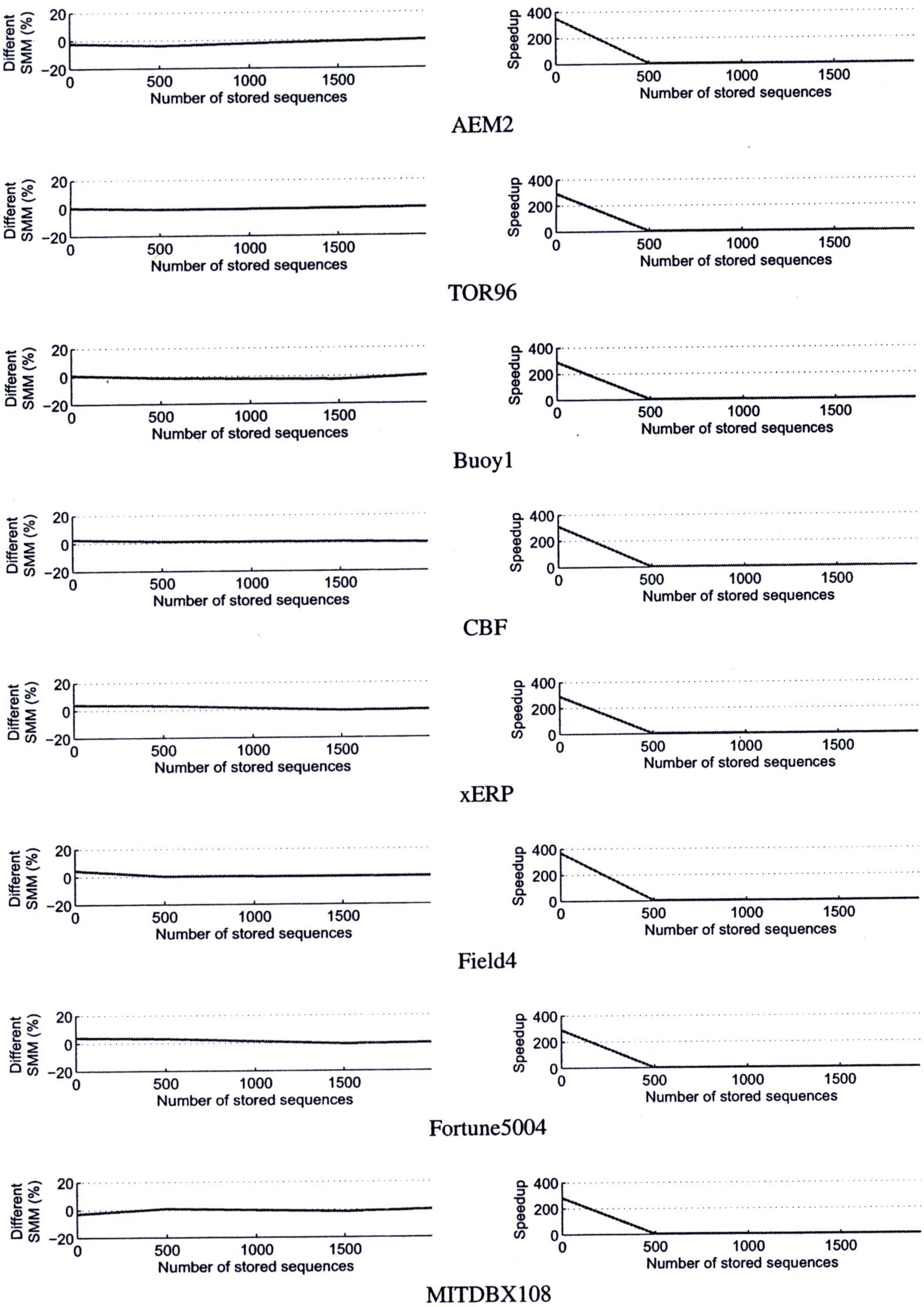


Figure H.1: Percentage difference of SMM and speedup of 3STSC with CDTW function and complete linkage when $k = 3$, $w = 64$, and number of stored sequences are varied.

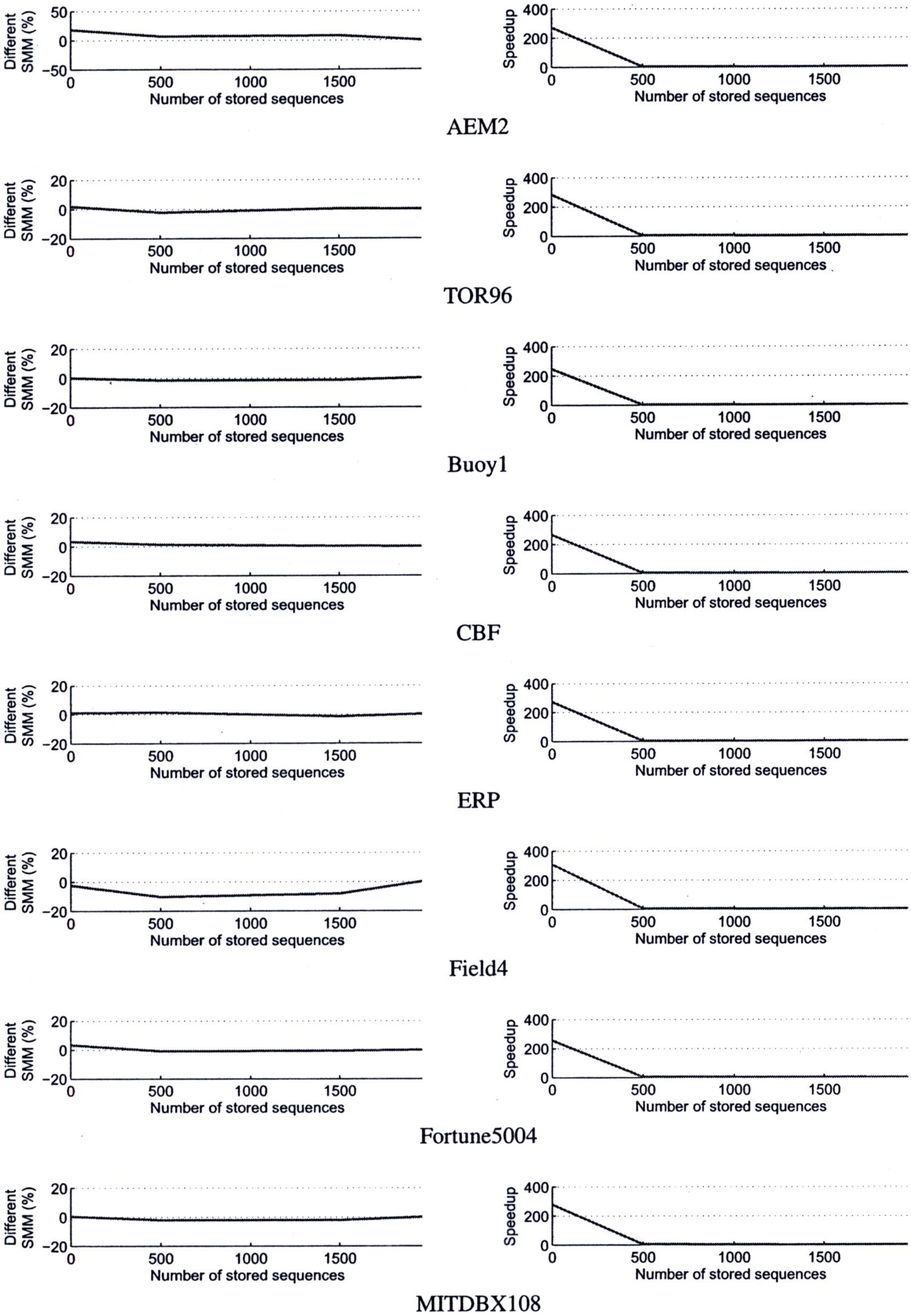


Figure H.2: Percentage difference of SMM and speedup of 3STSC with CDTW function and complete linkage when $k = 3$, $w = 32$, and number of stored sequences are varied.

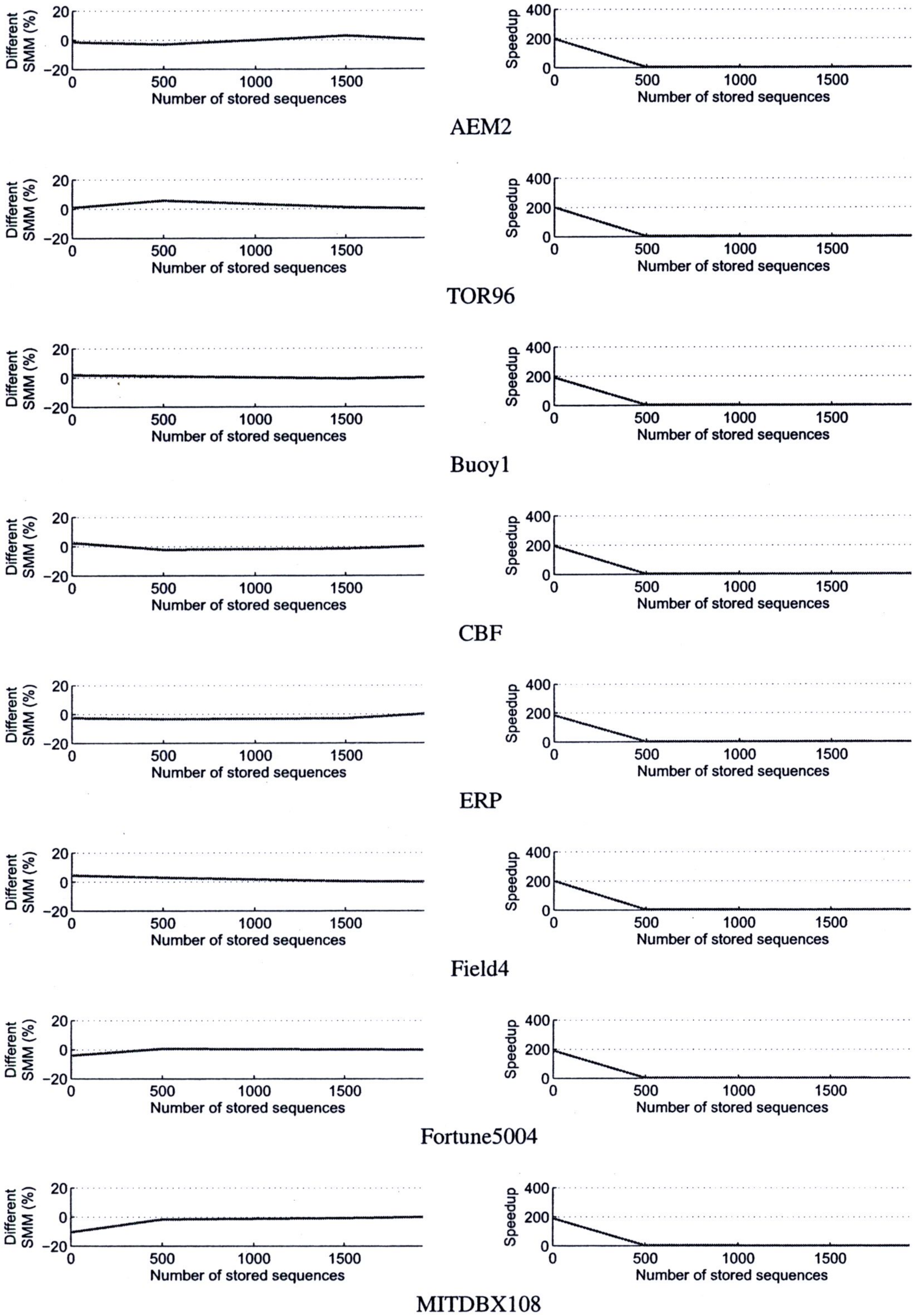


Figure H.3: Percentage difference of SMM and speedup of 3STSC with CDTW function and complete linkage when $k = 5$, $w = 64$, and number of stored sequences are varied.

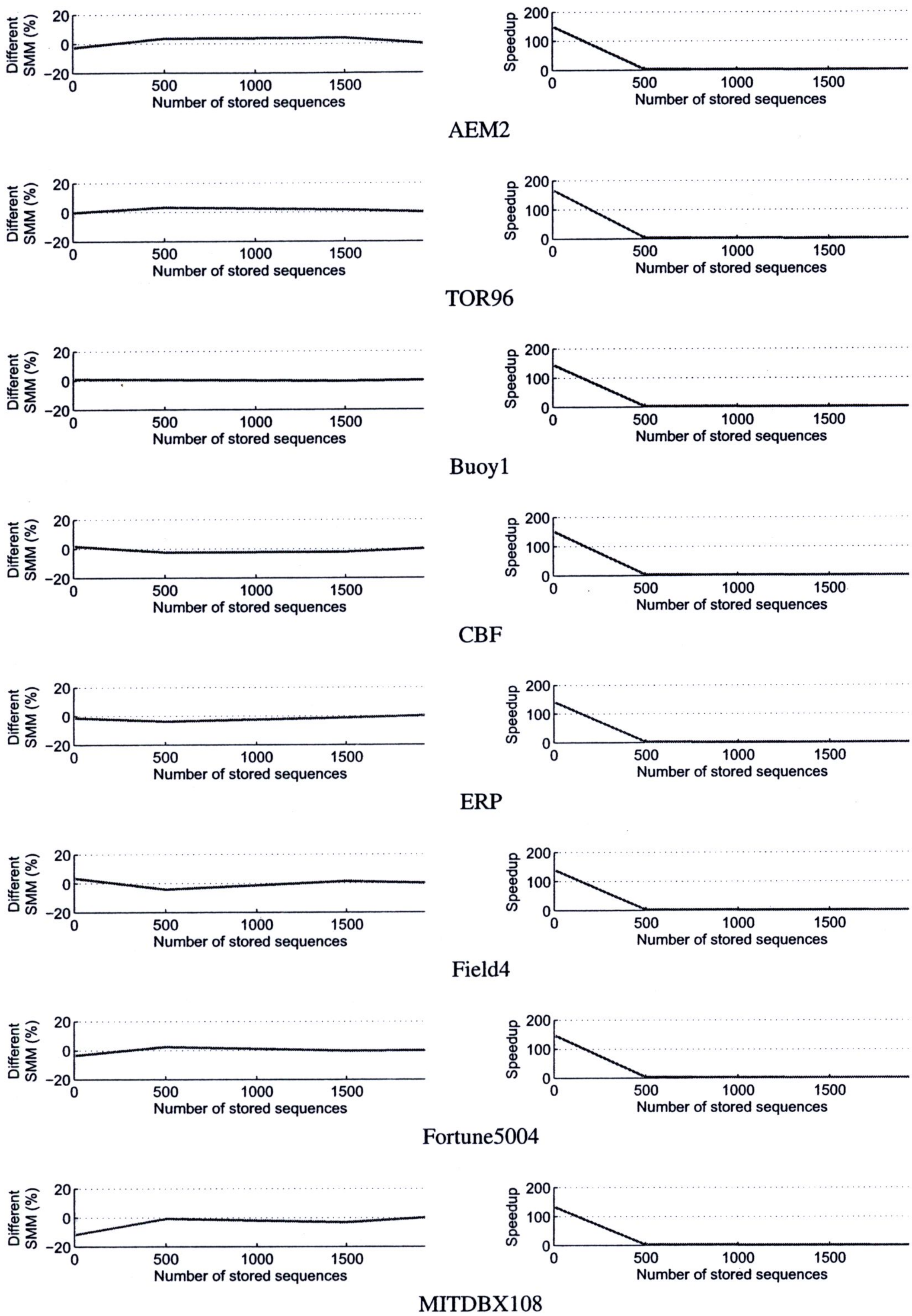
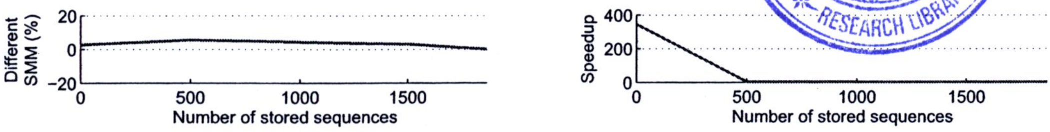
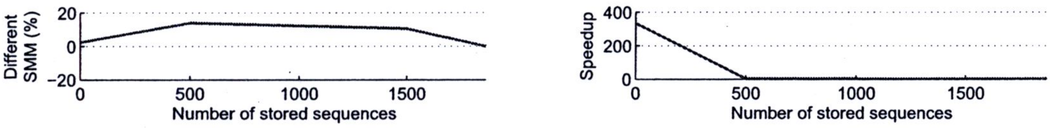


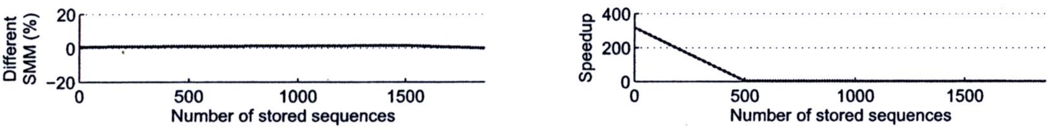
Figure H.4: Percentage difference of SMM and speedup of 3STSC with CDTW function and complete linkage when $k = 7$, $w = 64$, and number of stored sequences are varied.



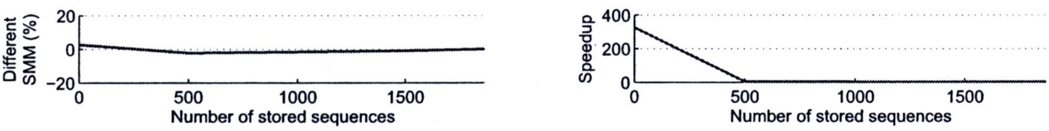
AEM2



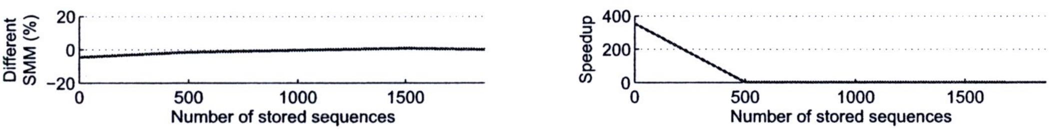
TOR96



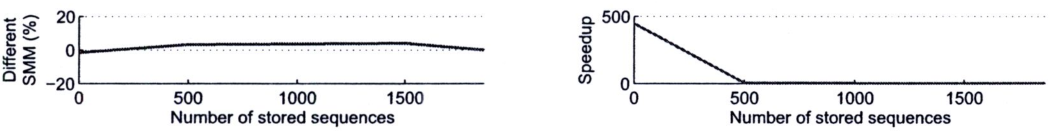
Buoy1



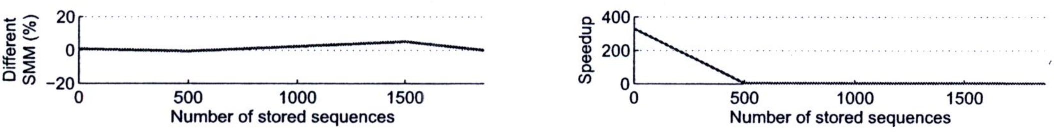
CBF



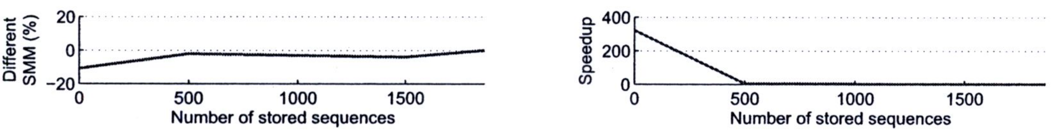
ERP



Field4

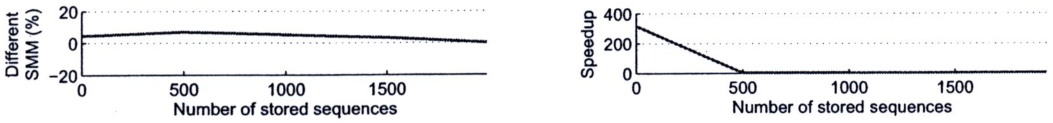


Fortune5004

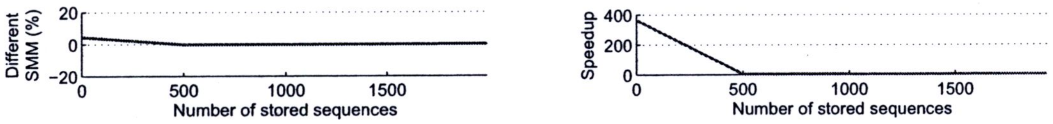


MITDBX108

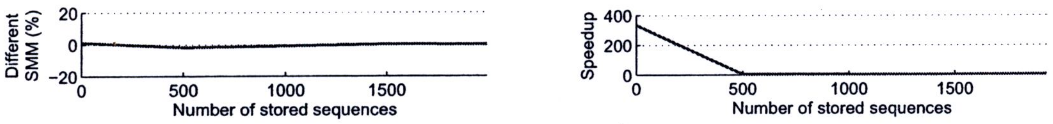
Figure H.5: Percentage difference of SMM and speedup of 3STSC with CDTW function and complete linkage when $k = 3$, $w = 128$, and number of stored sequences are varied.



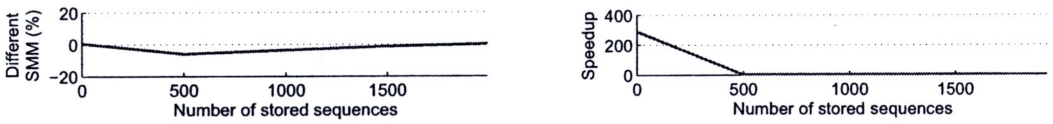
AEM2



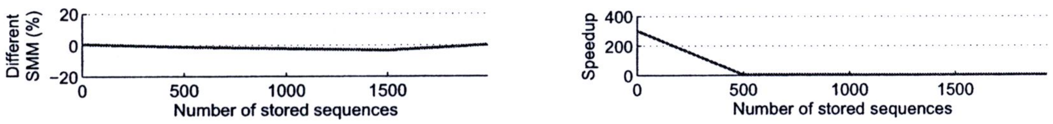
TOR96



Buoy1

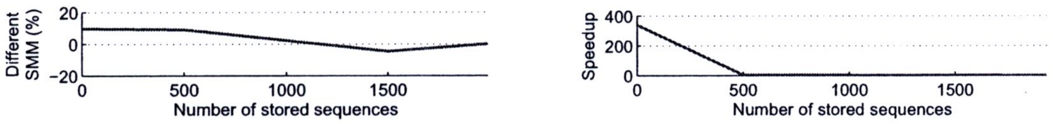


CBF

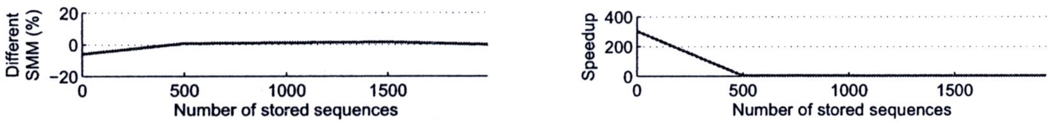


ERP

Field4



Fortune5004



MITDBX108

Figure H.6: Percentage difference of SMM and speedup of 3STSC with CDTW function and average linkage when $k = 3$, $w = 64$, and number of stored sequences are varied.

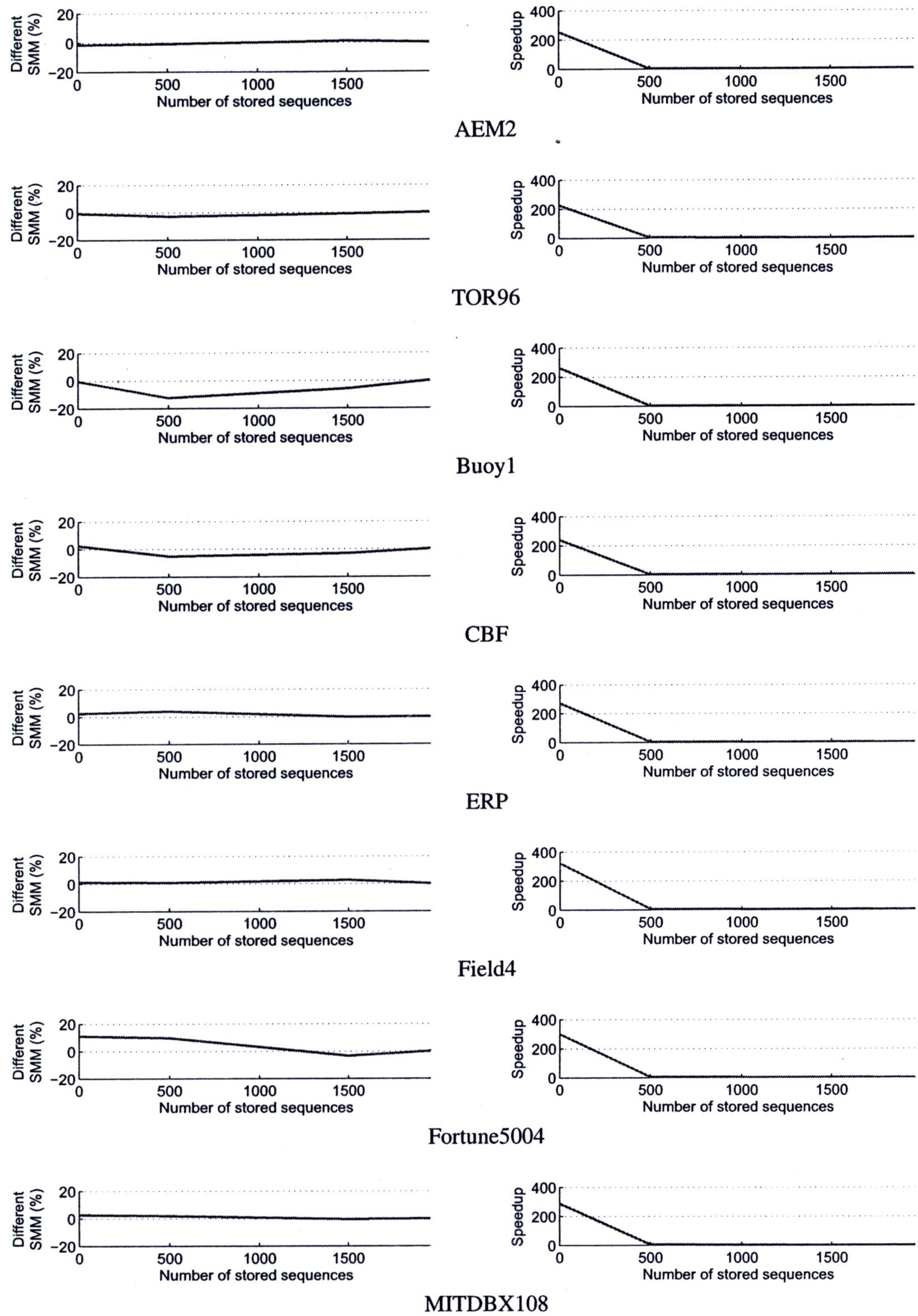


Figure H.7: Percentage difference of SMM and speedup of 3STSC with CDTW function and average linkage when $k = 3$, $w = 64$, and number of stored sequences are varied.

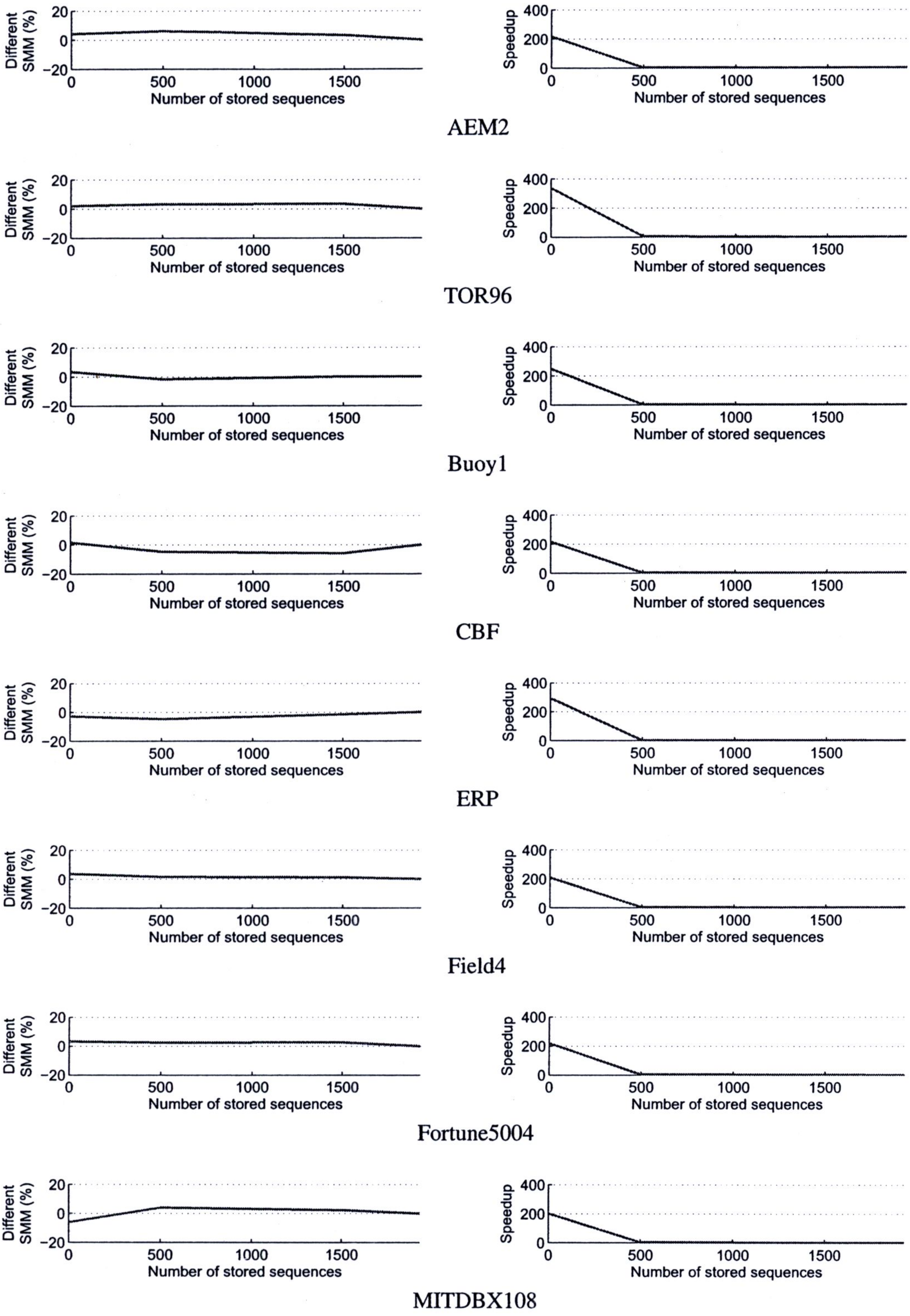


Figure H.8: Percentage difference of SMM and speedup of 3STSC with CDTW function and average linkage when $k = 5$, $w = 64$, and number of stored sequences are varied.

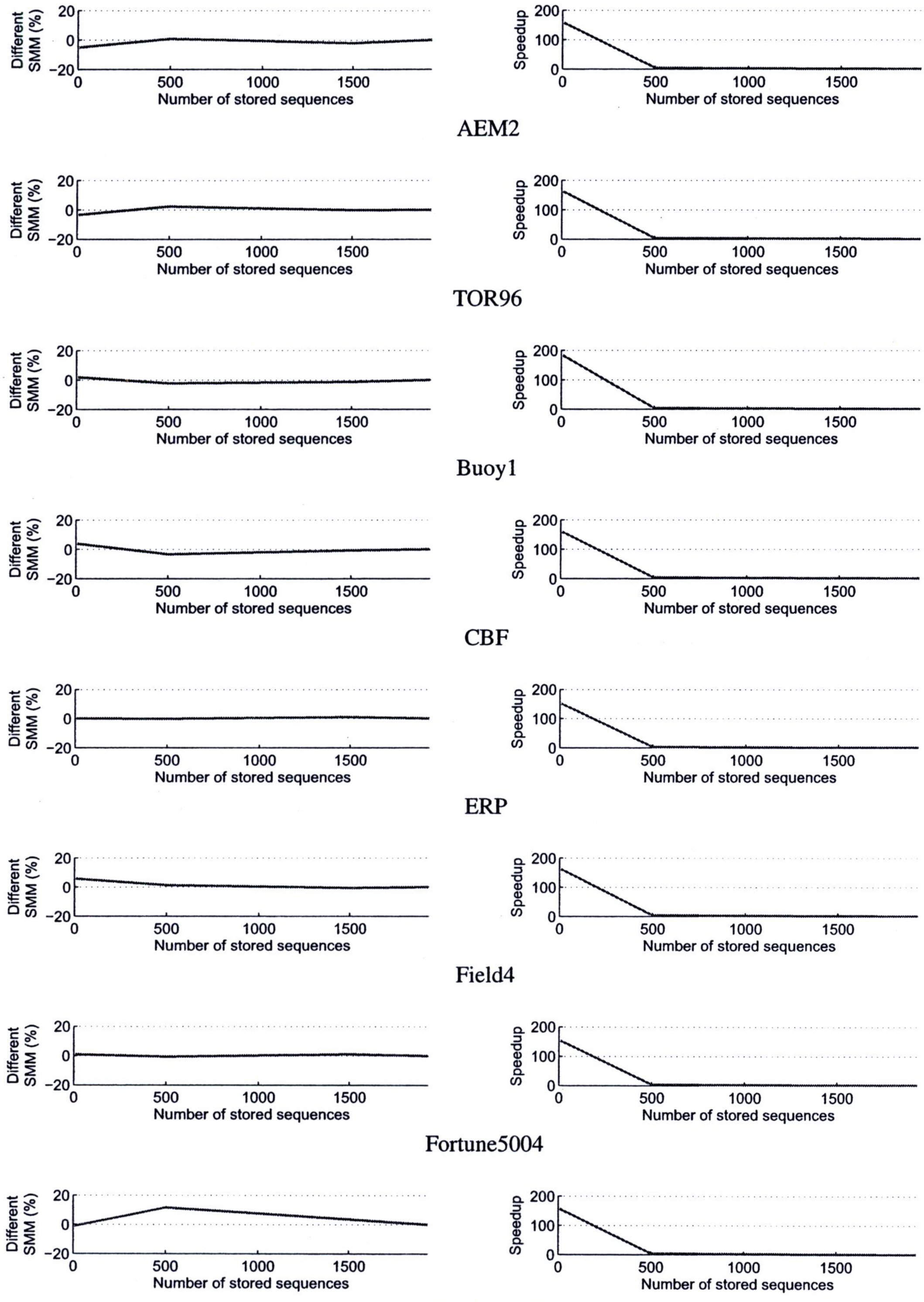


Figure H.9: Percentage difference of SMM and speedup of 3STSC with CDTW function and average linkage when $k = 7$, $w = 64$, and number of stored sequences are varied.

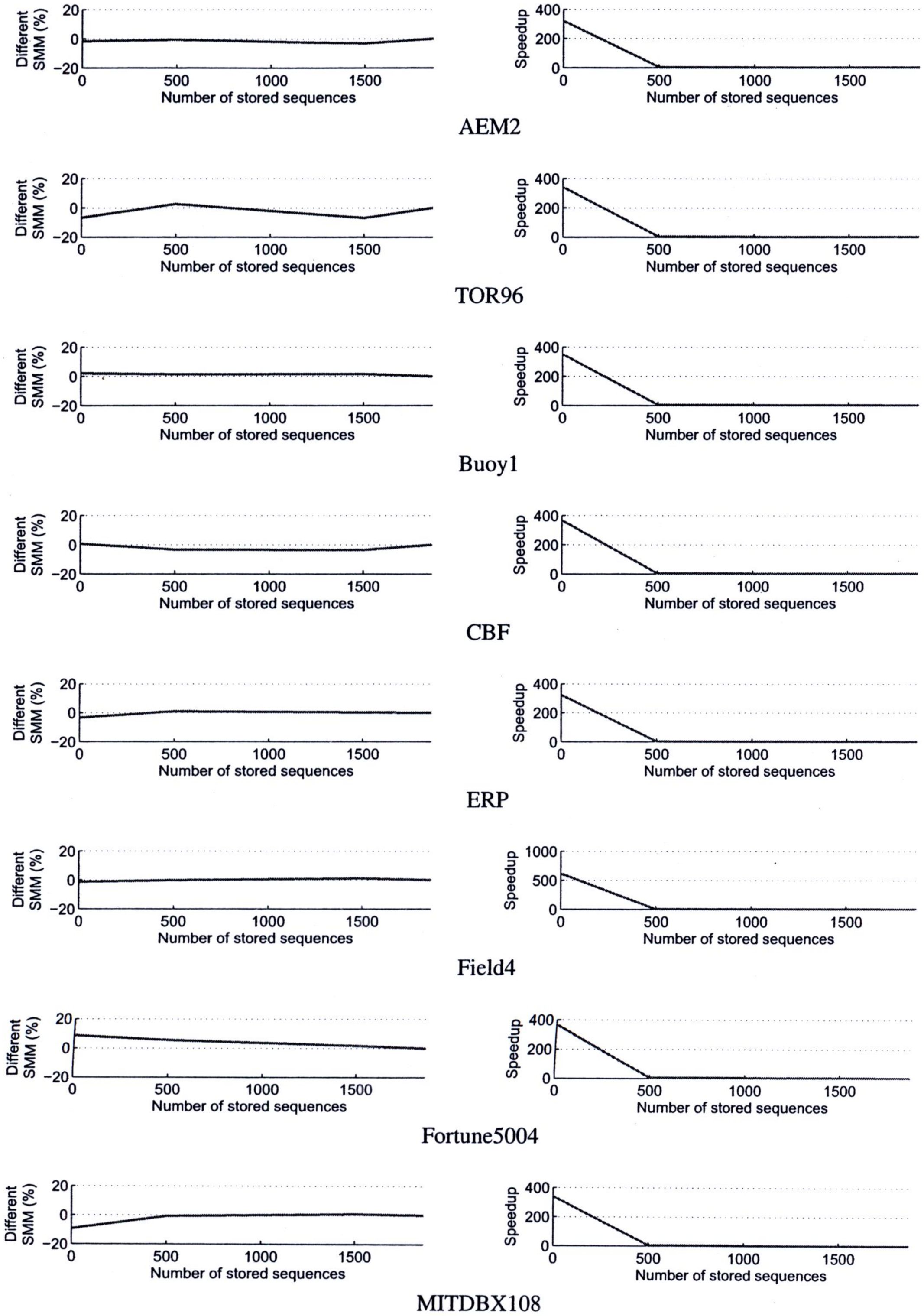
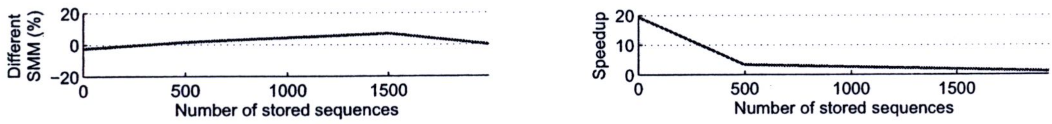
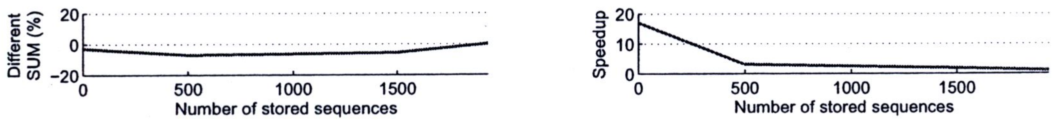


Figure H.10: Percentage difference of SMM and speedup of 3STSC with CDTW function and average linkage when $k = 3$, $w = 128$, and number of stored sequences are varied.



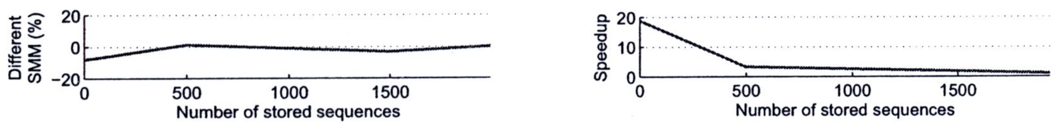
AEM2



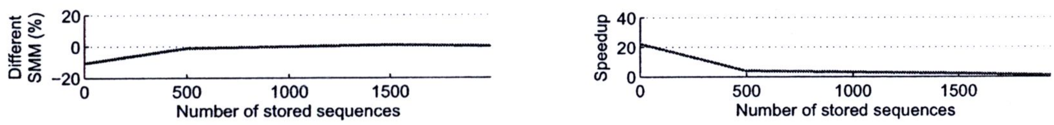
TOR96

Buoy1

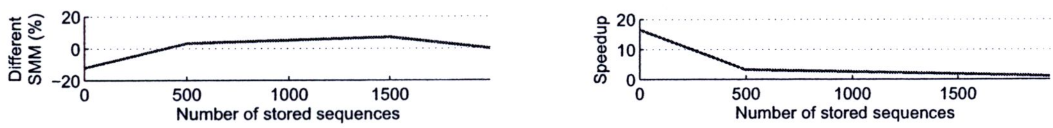
CBF



ERP



Field4



Fortune5004

MITDBX108

Figure H.11: Percentage difference of SMM and speedup of 3STSC with ICDTW function and complete linkage when $k = 3$, $w = 64$, and number of stored sequences are varied.

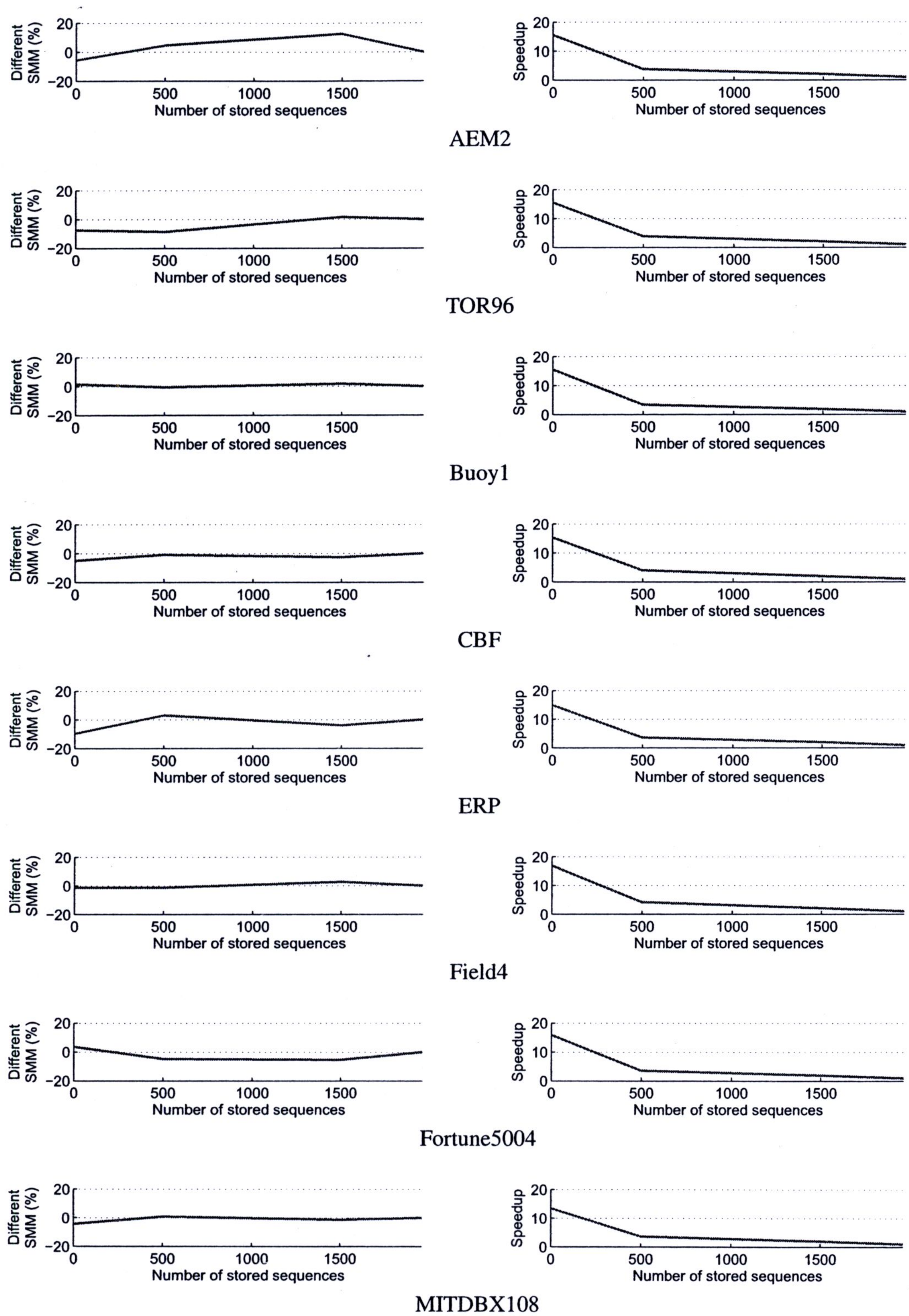


Figure H.12: Percentage difference of SMM and speedup of 3STSC with ICDTW function and complete linkage when $k = 3$, $w = 32$, and number of stored sequences are varied.

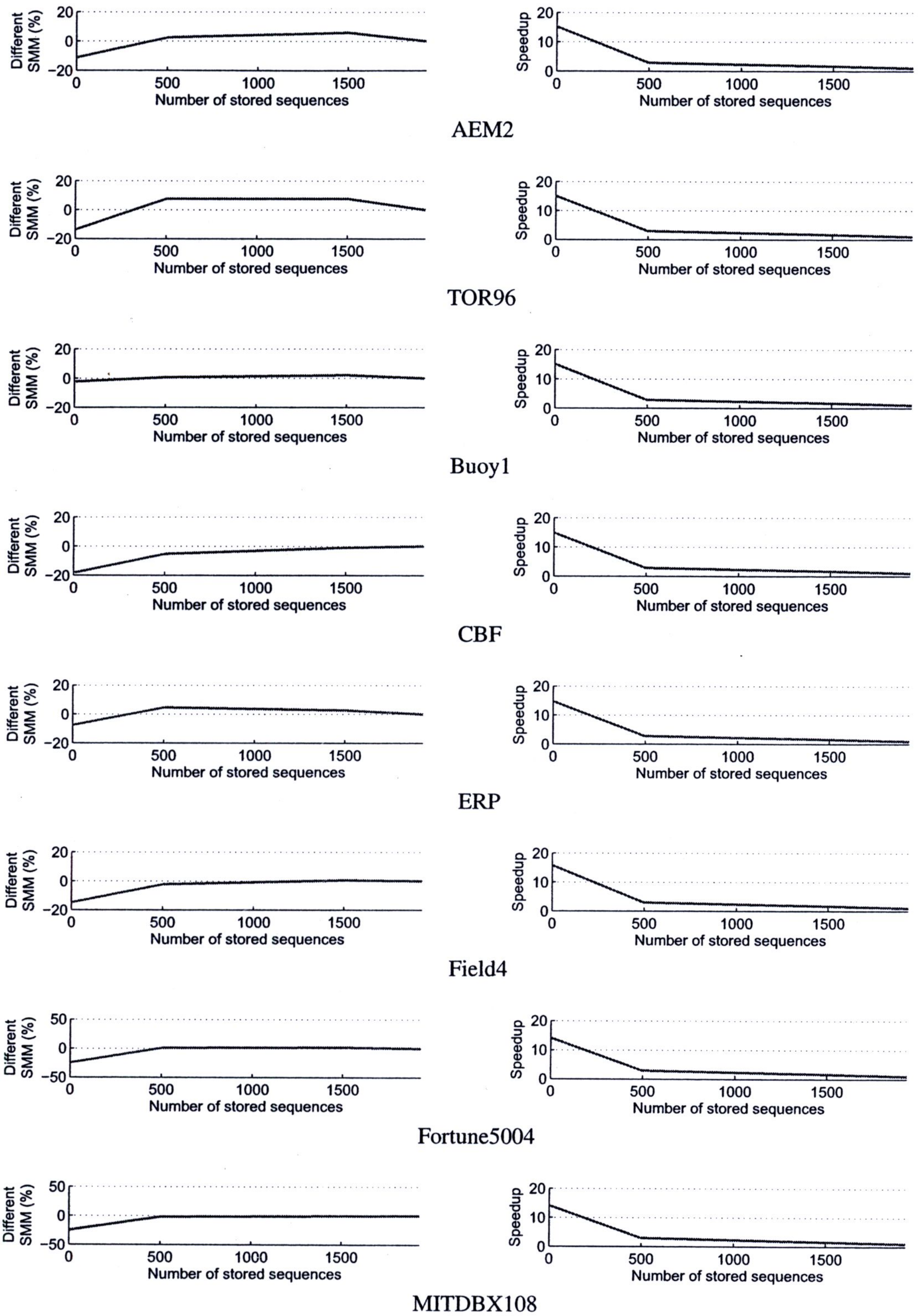
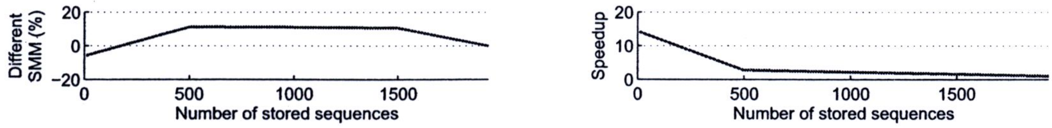


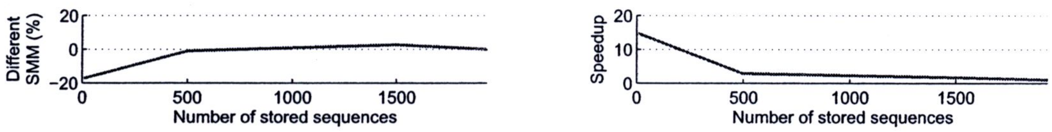
Figure H.13: Percentage difference of SMM and speedup of 3STSC with ICDTW function and complete linkage when $k = 5$, $w = 64$, and number of stored sequences are varied.



AEM2

TOR96

Buoy1

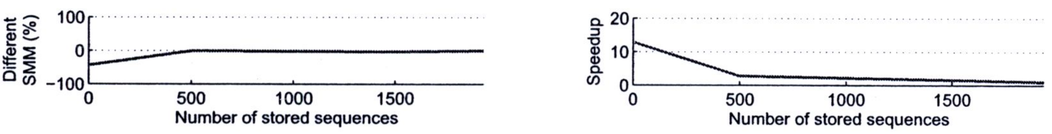


CBF

ERP

Field4

Fortune5004



MITDBX108

Figure H.14: Percentage difference of SMM and speedup of 3STSC with ICDTW function and complete linkage when $k = 7$, $w = 64$, and number of stored sequences are varied.

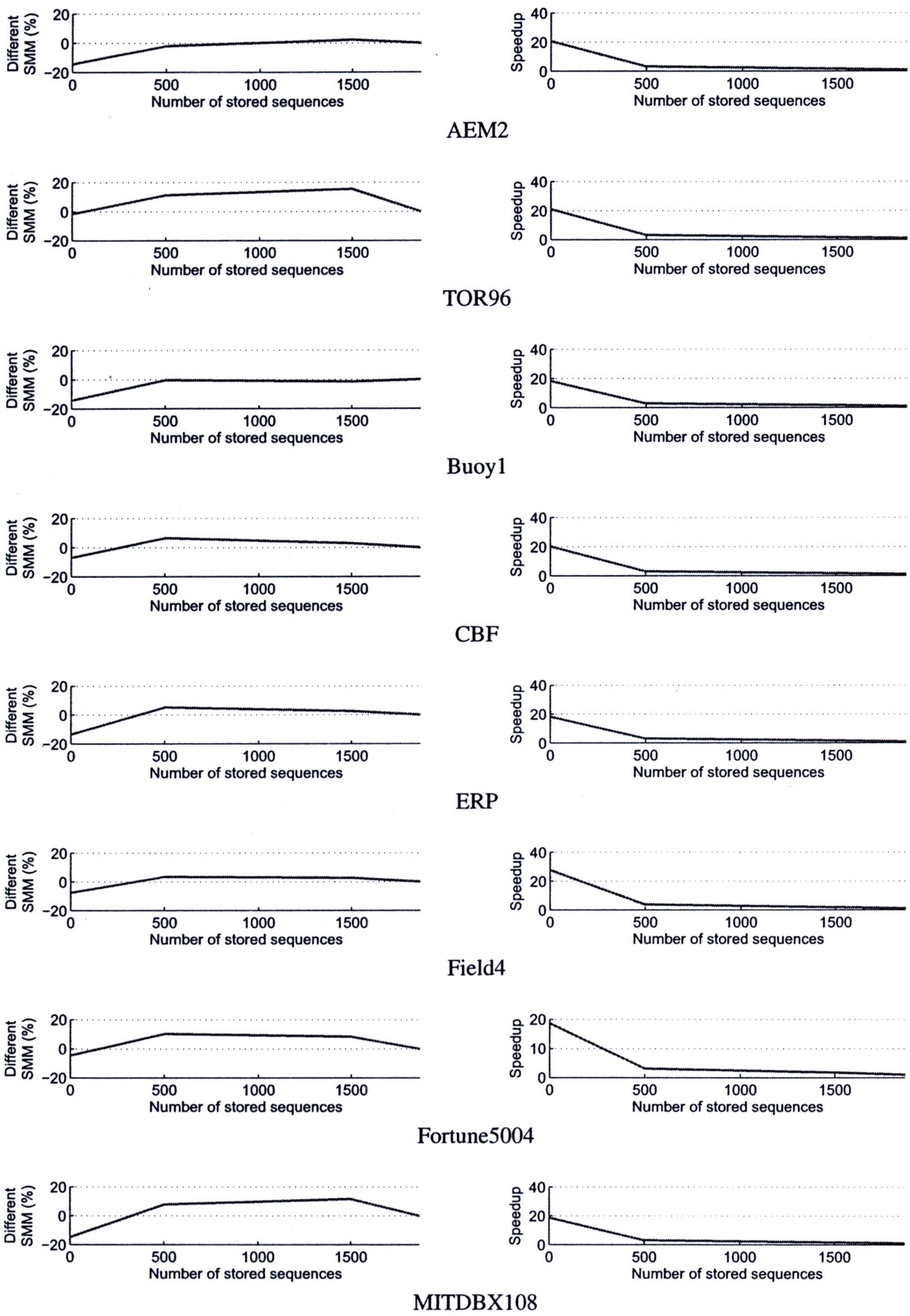


Figure H.15: Percentage difference of SMM and speedup of 3STSC with ICDTW function and complete linkage when $k = 3$, $w = 128$, and number of stored sequences are varied.

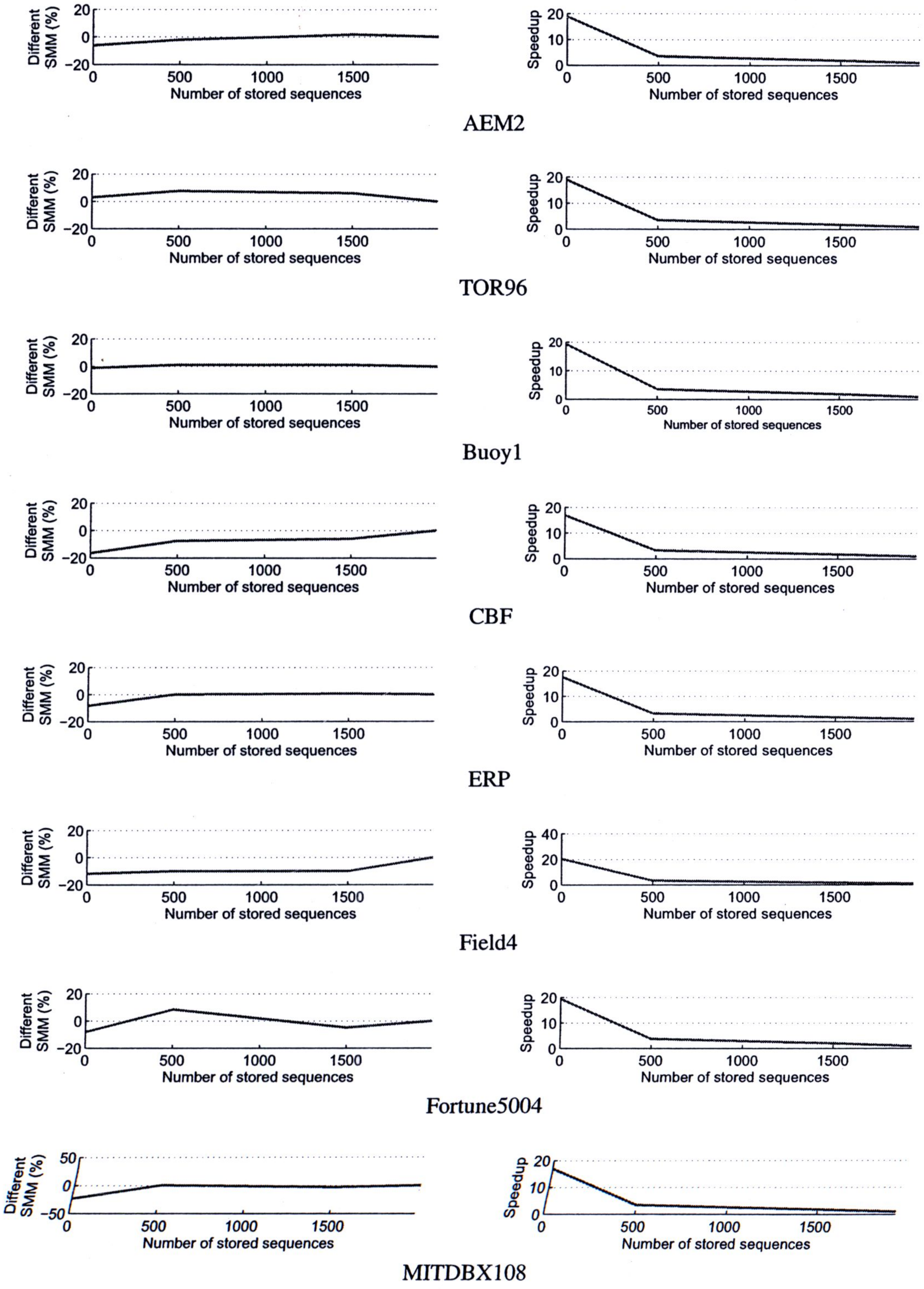


Figure H.16: Percentage difference of SMM and speedup of 3STSC with ICDTW function and average linkage when $k = 3$, $w = 64$, and number of stored sequences are varied.

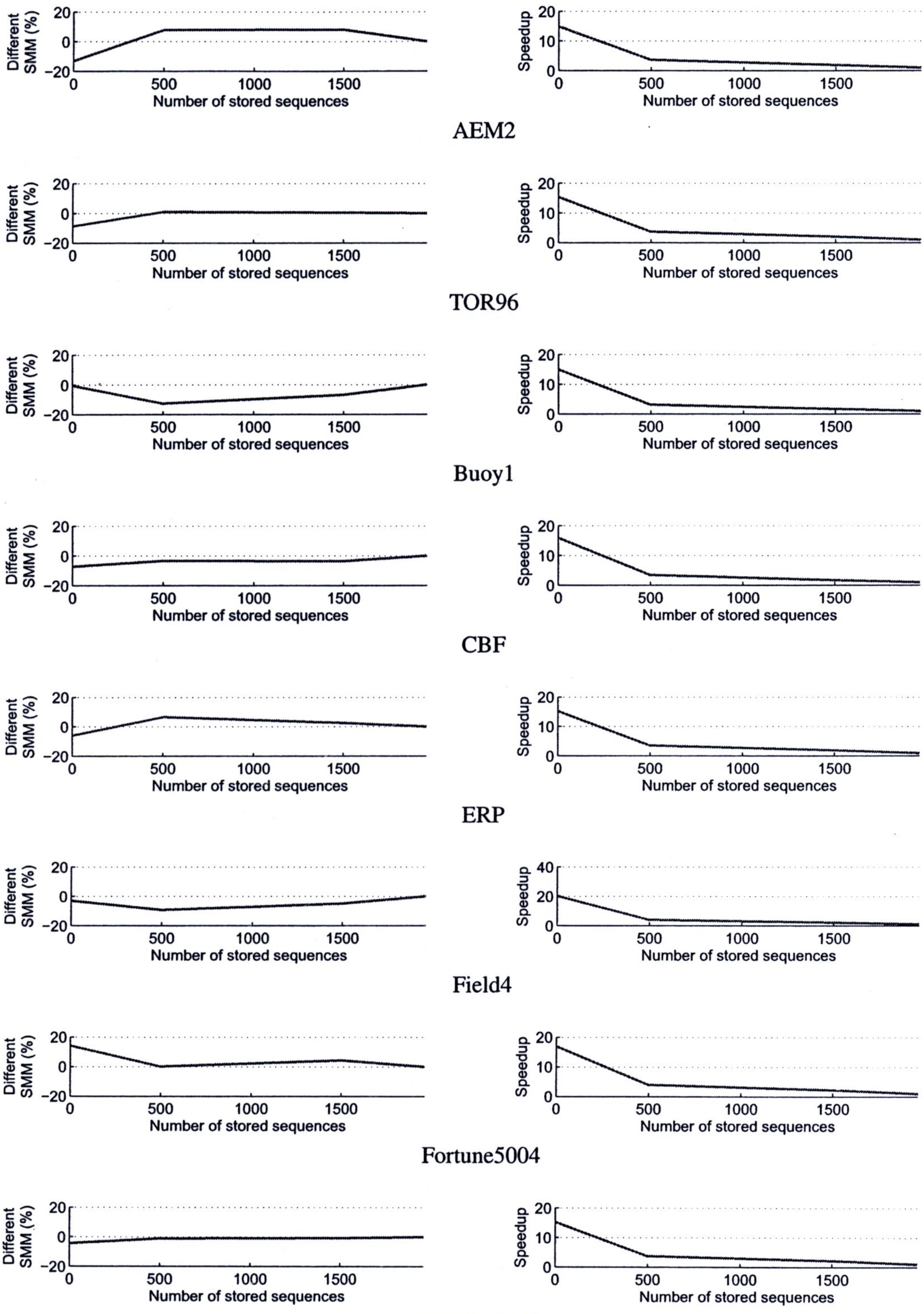
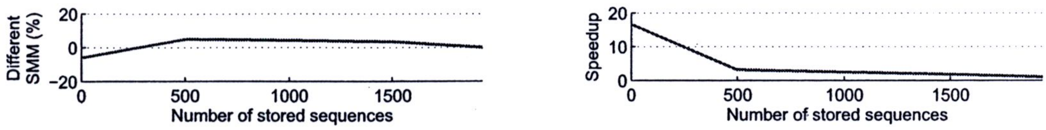
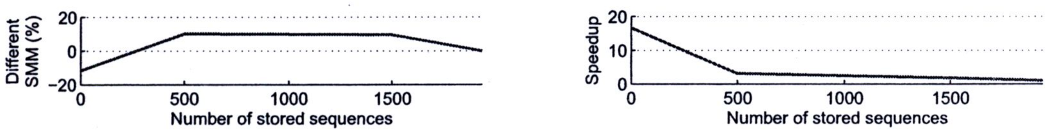


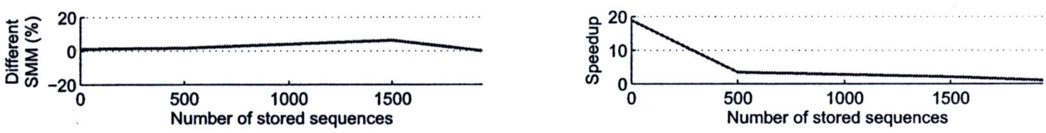
Figure H.17: Percentage difference of SMM and speedup of 3STSC with ICDTW function and average linkage when $k = 3$, $w = 32$, and number of stored sequences are varied.



AEM2

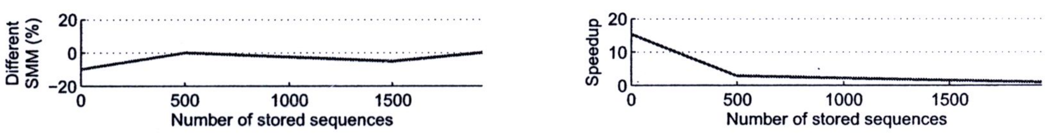


TOR96



Buoy1

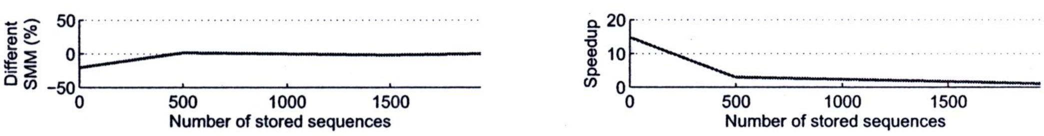
CBF



ERP

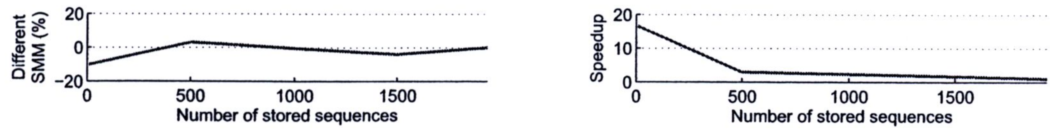
Field4

Fortune5004

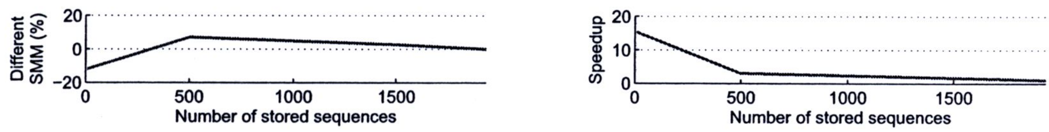


MITDBX108

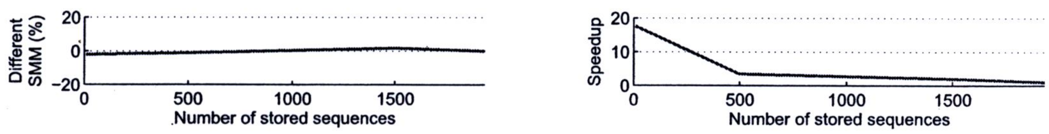
Figure H.18: Percentage difference of SMM and speedup of 3STSC with ICDTW function and average linkage when $k = 3$, $w = 64$, and number of stored sequences are varied.



AEM2

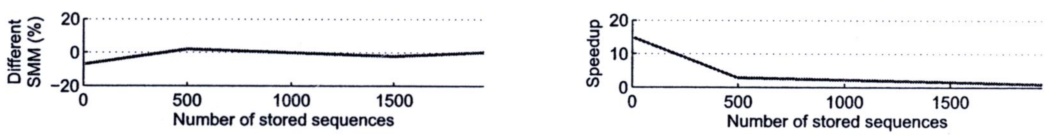


TOR96

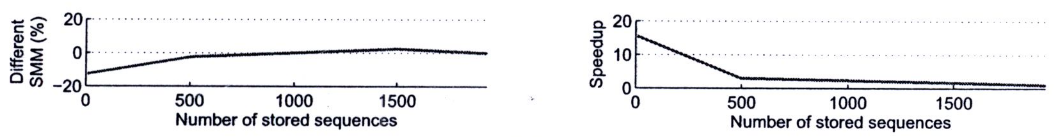


Buoy1

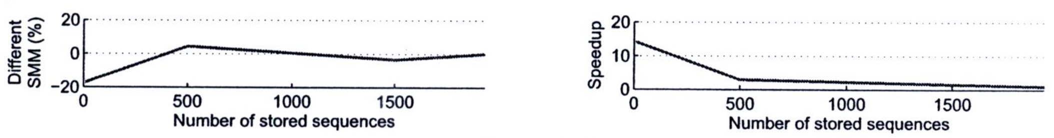
CBF



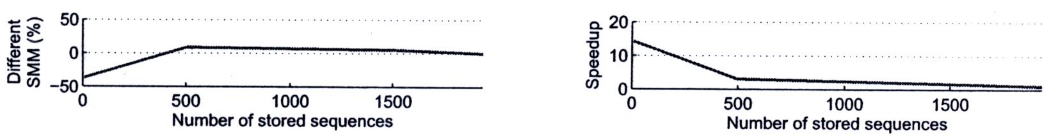
ERP



Field4



Fortune5004



MITDBX108

Figure H.19: Percentage difference of SMM and speedup of 3STSC with ICDTW function and average linkage when $k = 7$, $w = 64$, and number of stored sequences are varied.

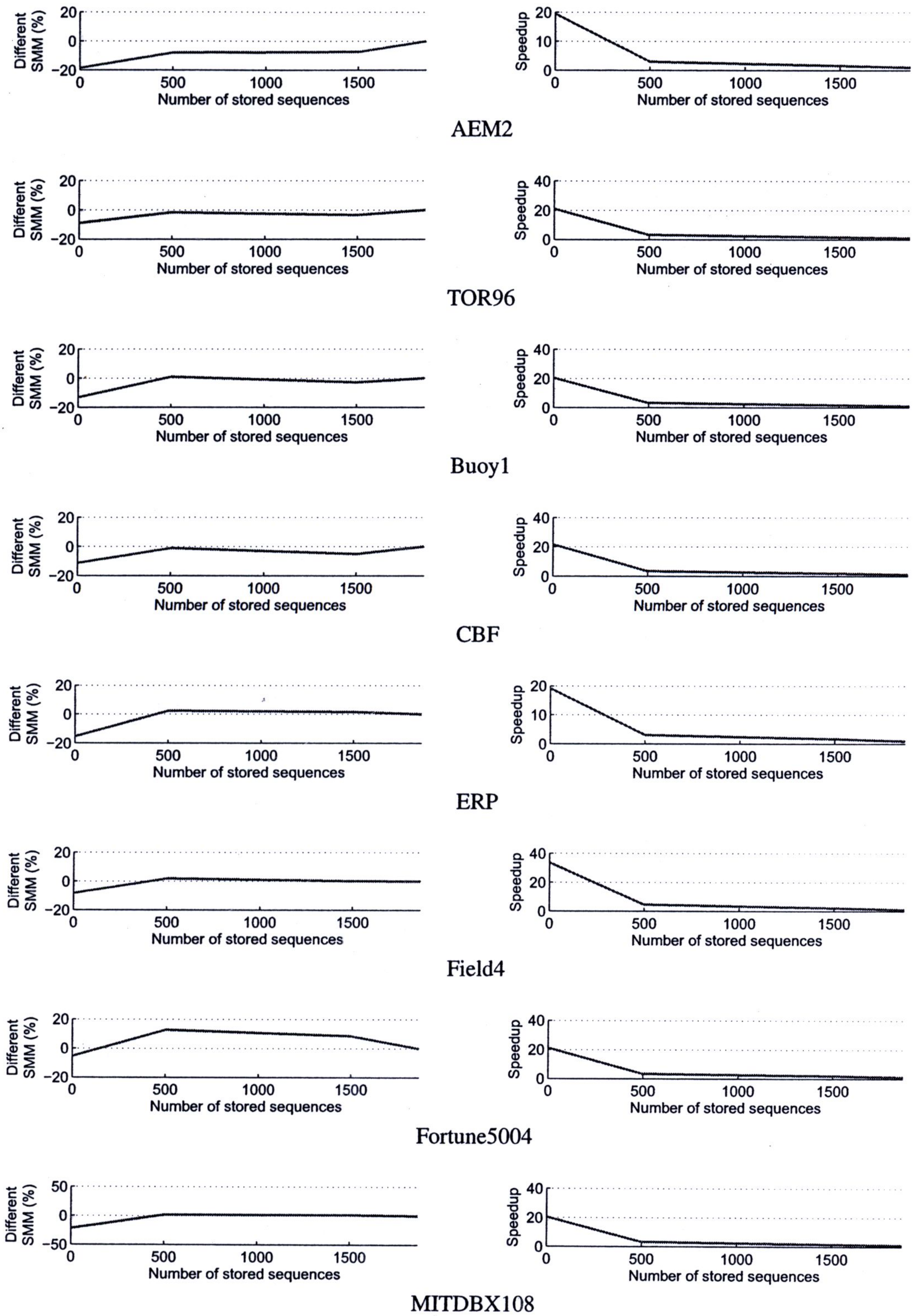


Figure H.20: Percentage difference of SMM and speedup of 3STSC with ICDTW function and average linkage when $k = 3$, $w = 128$, and number of stored sequences are varied.

Biography

Vit Niennattrakul was born in Bangkok, Thailand, on September 8, 1984. He received his B.Eng. in Computer Engineering from Chulalongkorn University in 2006. His doctorate has been under supervision of Asst. Prof. Dr. Chotirat Ann Ratanamahatana. During his Ph.D. study, he was a junior specialist at the University of California, Riverside under supervision of Prof. Eamonn J. Keogh for one year (September 2009 to August 2010), and he was granted scholarships from the Thailand Research Fund through the Royal Golden Jubilee Ph.D. Program (June 2007 to May 2011), Chulalongkorn University Graduate Scholarship to Commemorate the 72nd Anniversary of His Majesty King Bhumibol Adulyadej (June 2006 to May 2007), and the 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund) (September 2009). He also was reviewers of many well-known journals including Knowledge and Information System (KAIS) and Data Mining and Knowledge Discovery (DMKD). His research interests include but not limited to time series data mining, machine learning, and natural language processing.