Supplementary Material:
A Framework to Evaluate Early Time-Series Classification Algorithms

Charilaos Akasiadis   Evgenios Kladis   Petro-Foti Kamberi
Evangelos Michelioudakis   Elias Alevizos   Alexander Artikis

The source-code used for the experimental evaluation can be found at:

https://github.com/xarakas/ETSC/

A   Overview of Algorithm Results per Dataset Category

Algorithm	No. of Wins	Dataset Categories Won wrt. Accuracy
ECEC	7	Wide, Large, Unstable, Imbalanced, Common, Univariate, Multivariate
S-MINI	1	Multiclass

Table 1: Results overview per dataset category in terms of Accuracy

Algorithm	No. of Wins	Dataset Categories Won wrt. Earliness
ECO-K	1	Wide
S-MLSTM	7	Large, Unstable, Imbalanced, Multiclass, Common, Univariate, Multivariate

Table 2: Results overview per dataset category in terms of Earliness

Algorithm	No. of Wins	Dataset Categories Won wrt. Harmonic Mean
ECEC	2	Wide, Unstable
S-MINI	1	Common
S-MLSTM	5	Large, Imbalanced, Multiclass, Univariate, Multivariate

Table 3: Results overview per dataset category in terms of Harmonic Mean

Algorithm	No. of Wins	Dataset Categories Won wrt. Training Time
ECO-K	1	Wide
S-WEASEL	8	Wide, Large, Unstable, Imbalanced, Multiclass, Common, Univariate, Multivariate

Table 4: Results overview per dataset category in terms of Training Time

B Barplots of Results per Dataset

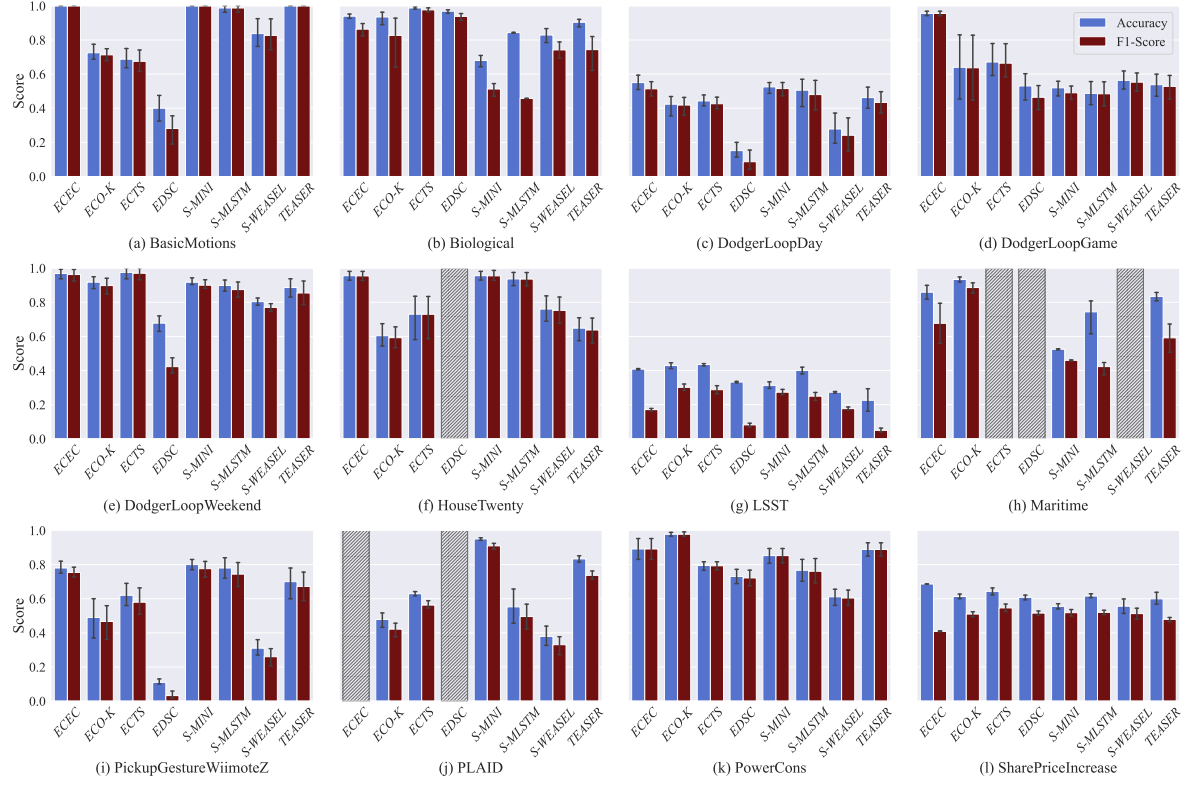In this section we present the barplots with the results from our evaluation.

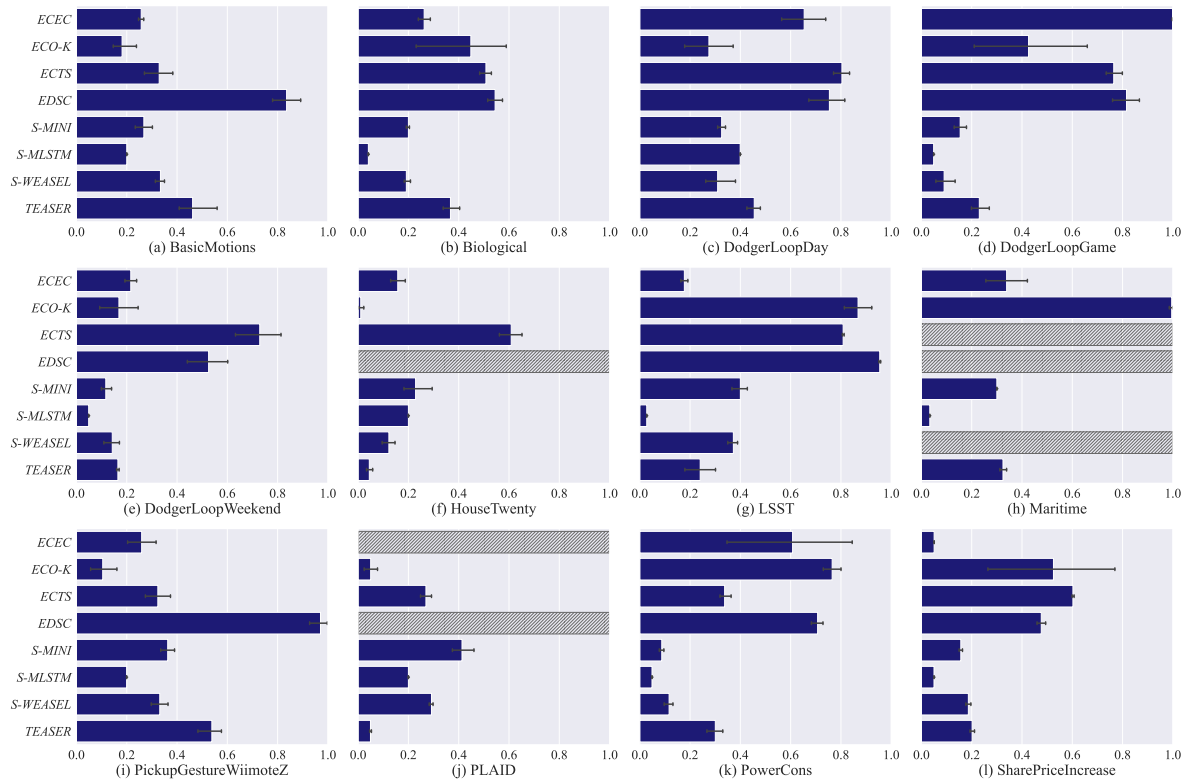Figure 1: Accuracy and F_1-Score.

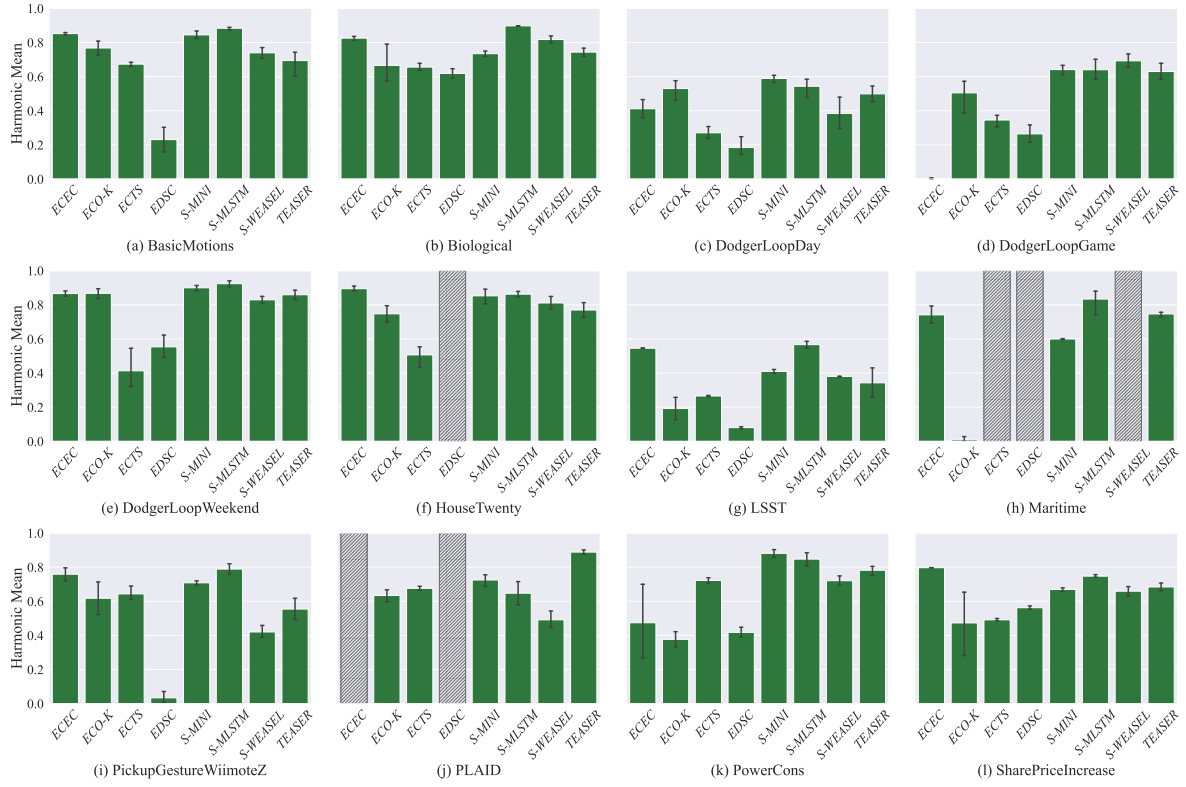Figure 2: Earliness (lower values are better).
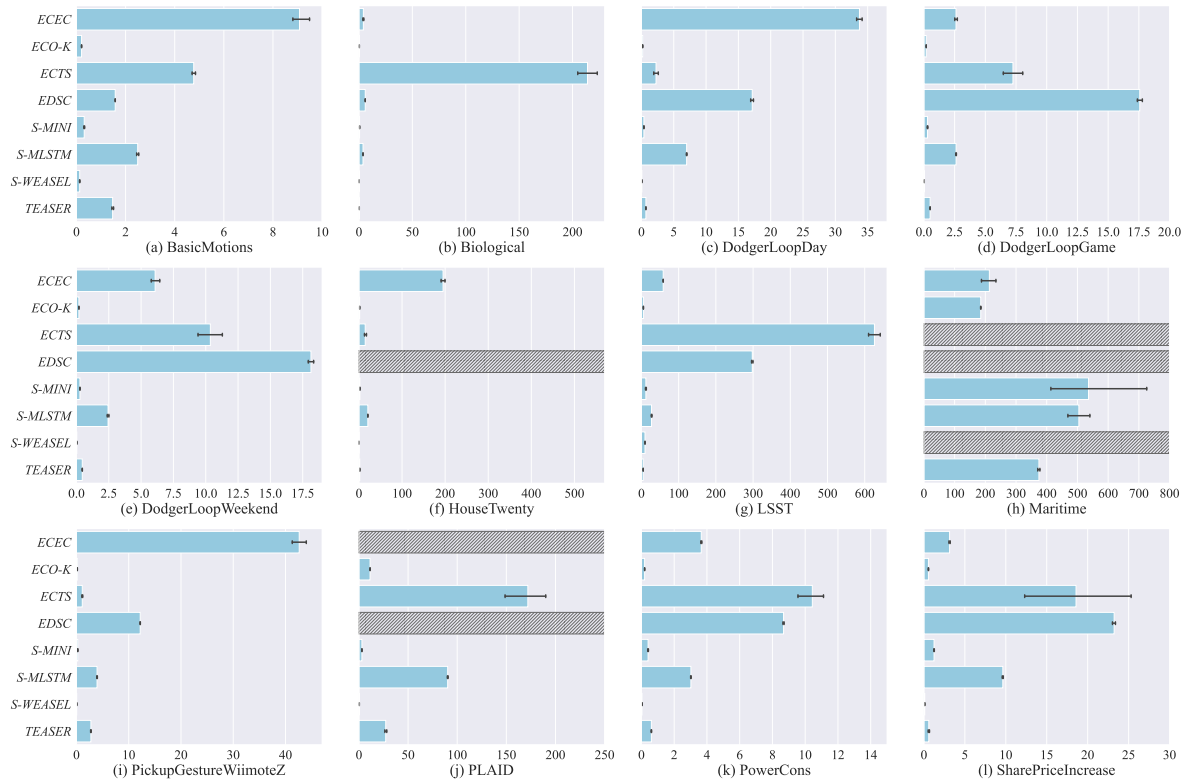
Figure 3: Harmonic Mean between earliness and accuracy.

Figure 4: Training times (x-axis scale varies, lower values are better).

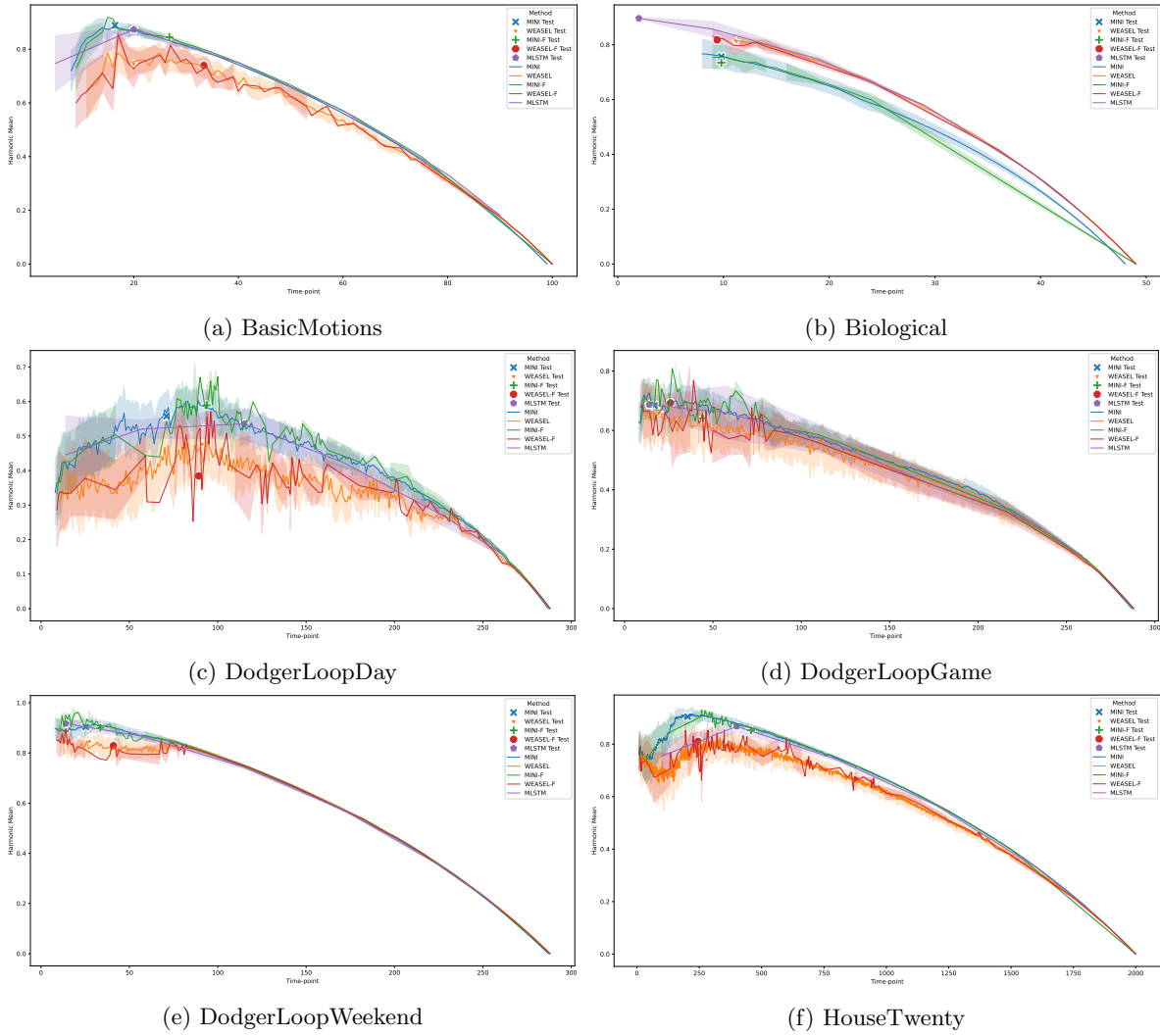# C STRUT: Harmonic-Mean vs Earliness for Each Dataset

(a) BasicMotions

(b) Biological

(c) DodgerLoopDay

(d) DodgerLoopGame

(e) DodgerLoopWeekend

(f) HouseTwenty

Figure 5: Harmonic Mean vs. Earliness for each dataset (i).

(a) LSST

(b) Maritime

(c) PickupGestureWiimoteZ

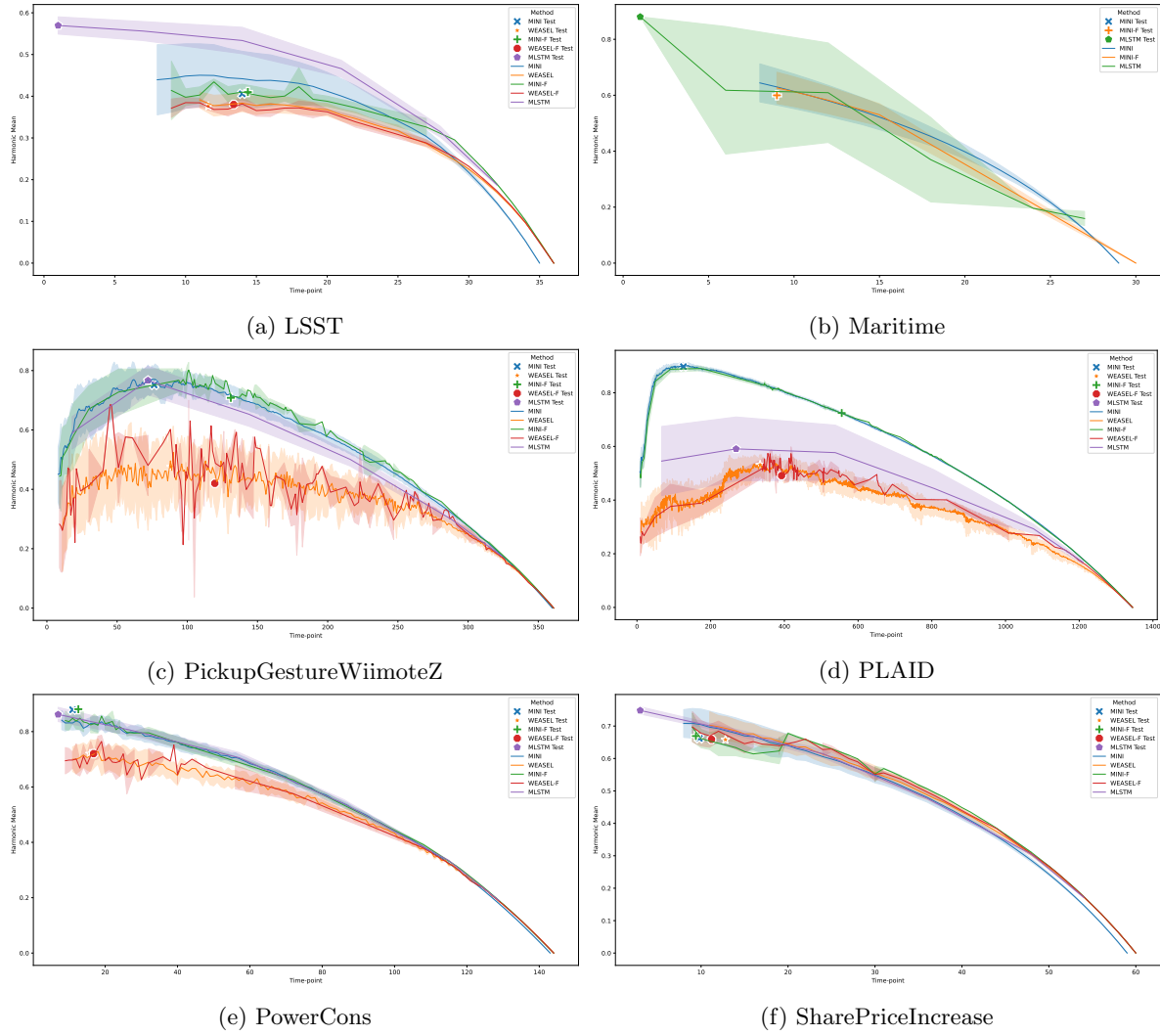(d) PLAID

(e) PowerCons

(f) SharePriceIncrease

Figure 6: Harmonic Mean vs. Earliness for each dataset (ii).

7

D   Details of Incorporated Datasets

We hereby provide a brief overview of the contents of each dataset.

BasicMotions   This dataset contains measurements from smart watches wore by four students performing different tasks. The measurements correspond to accelerometer and gyroscope 3-dimensional values. The labels correspond to a specific type of motion, in particular walking, resting, running, and playing badminton. The frequency of the measurements in this dataset is 10 Hz, and the duration of the time-series is 10 seconds. Figure 7 illustrates the mean and standard deviation of the measurements at each time-point, for the different variables categorized by the class labels assigned.
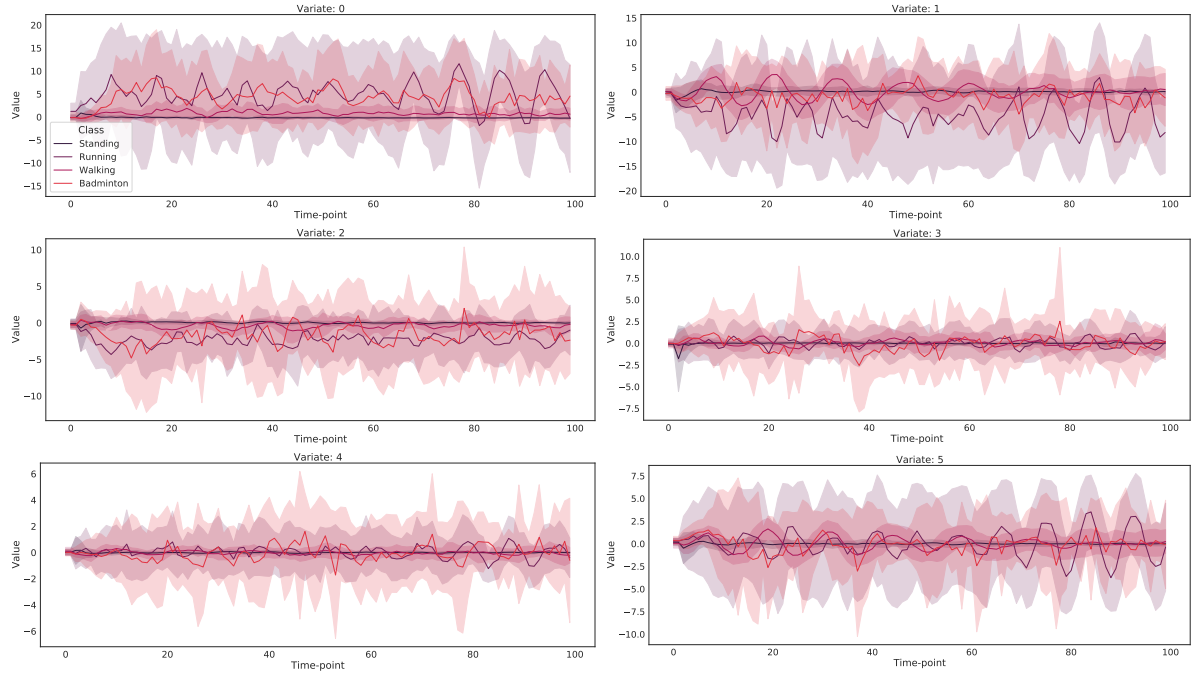

Figure 7: BasicMotions dataset: Mean and standard deviation of each variable's time-points per class.

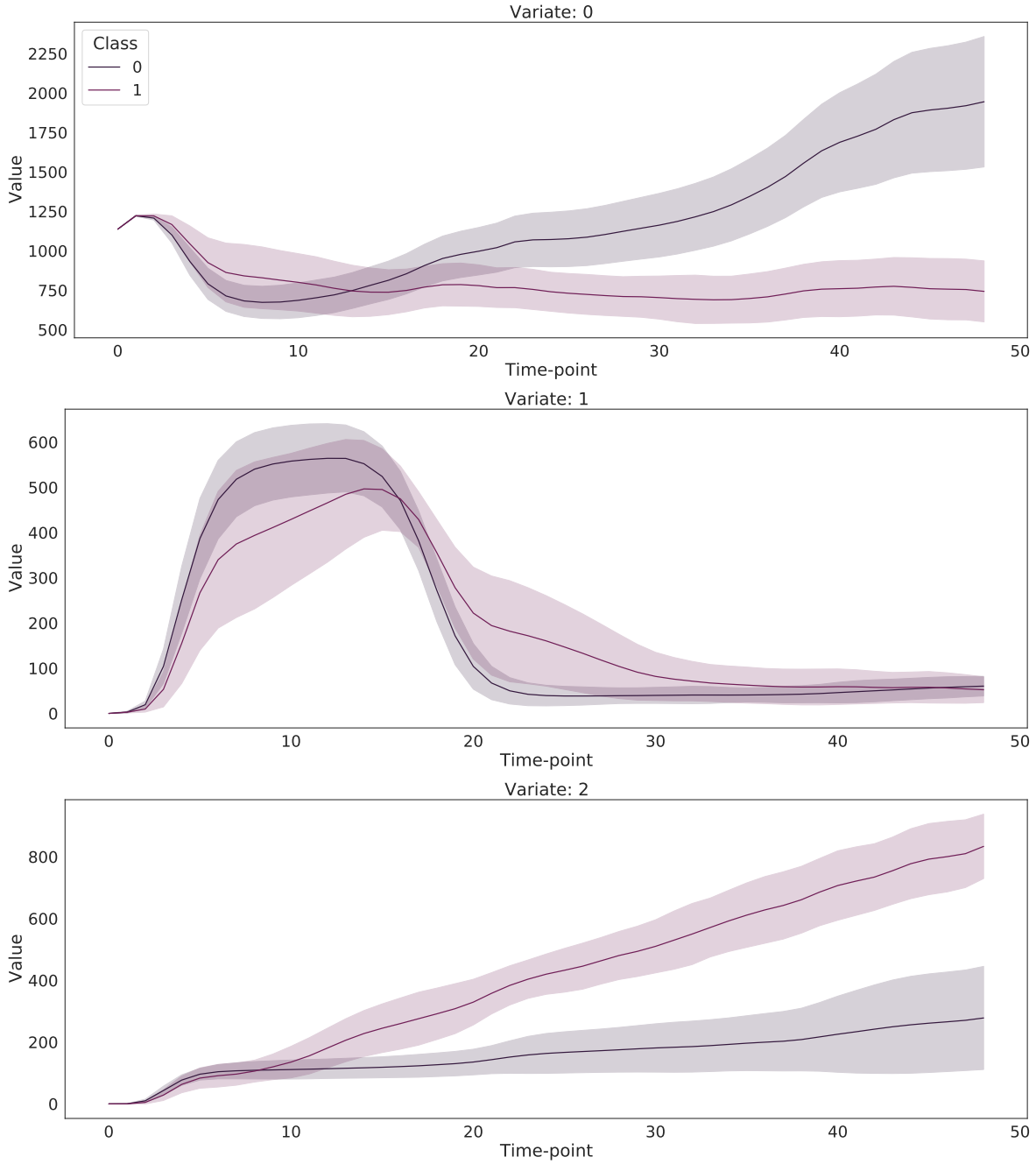**Biological**   The summary is given in Figure 8.

Figure 8: Biological dataset: Mean and standard deviation of each variable's time-points per class.

DodgerLoopDay This dataset contains measurements regarding the number of cars passing by a freeway ramp in Los Angeles, which is located close to a baseball stadium and where traffic is impacted when a game is scheduled. The frequency of measurements is 5 minutes and each class corresponds to a day of the week, Sunday to Saturday, respectively. Figure 9 illustrates the mean and standard deviation of the measurements at each time-point, for the different variables categorized by the class labels assigned.
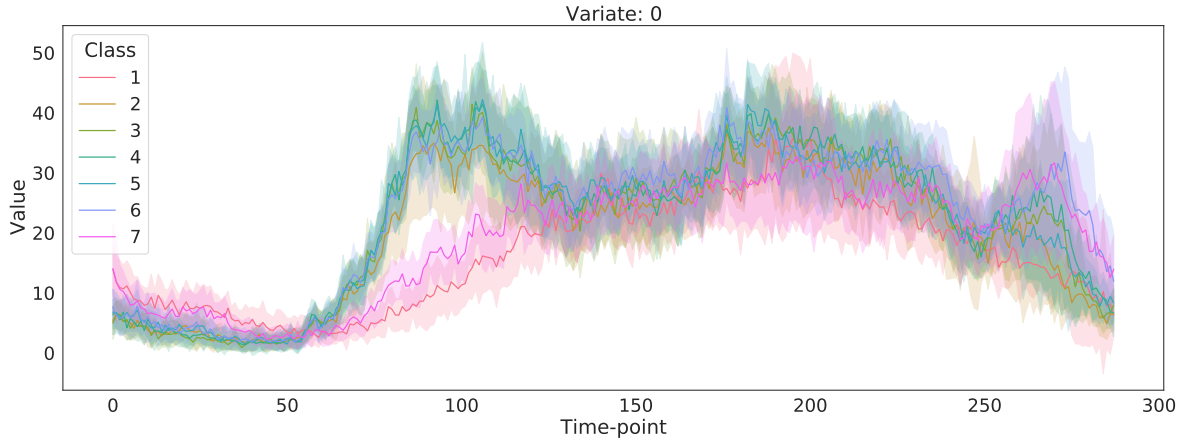
Figure 9: DodgerLoopDay dataset: Mean and standard deviation of each variable's time-points per class.

DodgerLoopGame This dataset's measurements are the same as the previous case, however here we have only two classes that correspond to normal and game days. Figure 10 summarizes the measurements contained in this dataset.
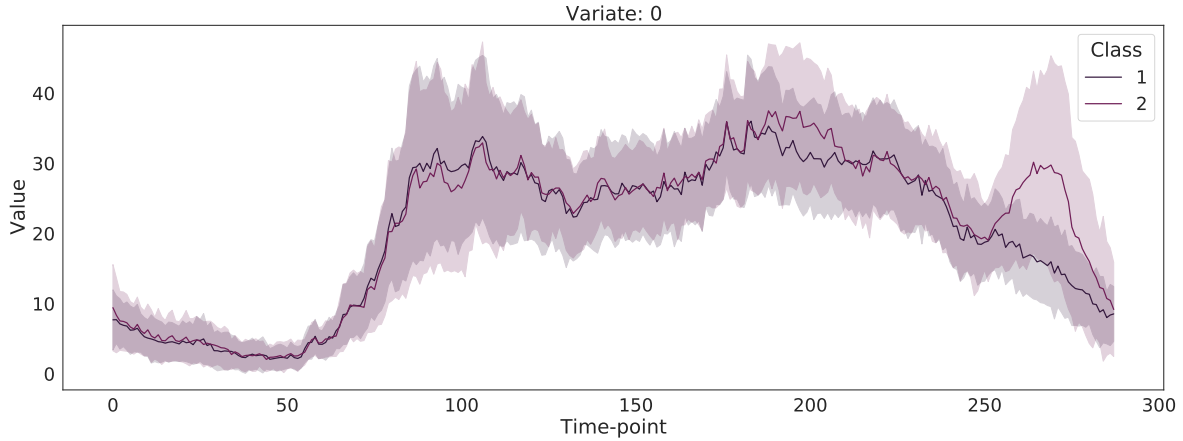

Figure 10: DodgerLoopGame dataset: Mean and standard deviation of each variable's time-points per class.

DodgerLoopWeekend In this case, the measurements are the same as the two previous datasets, but now the class labelling is performed according to if the day is a week day or a weekend (see, Figure 11).
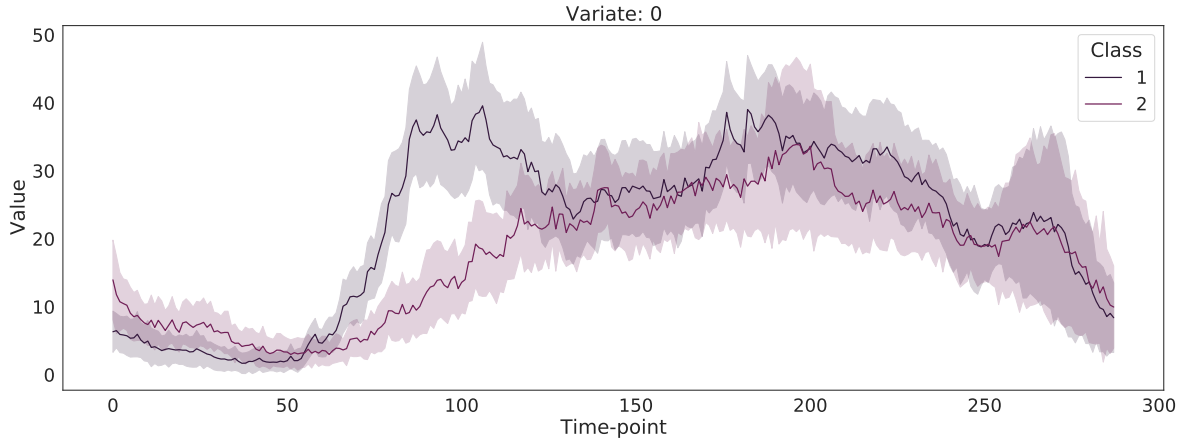
Figure 11: DodgerLoopWeekend dataset: Mean and standard deviation of each variable's time-points per class.

HouseTwenty   This dataset includes electricity consumption measurements originating from a house in a UK area. The two classes correspond to aggregate electricity consumption (label "1") and to washing machine and tumble dryer aggregate (label "2") with an 8 second frequency of measurements. The summary of this dataset's values is given by Figure 12
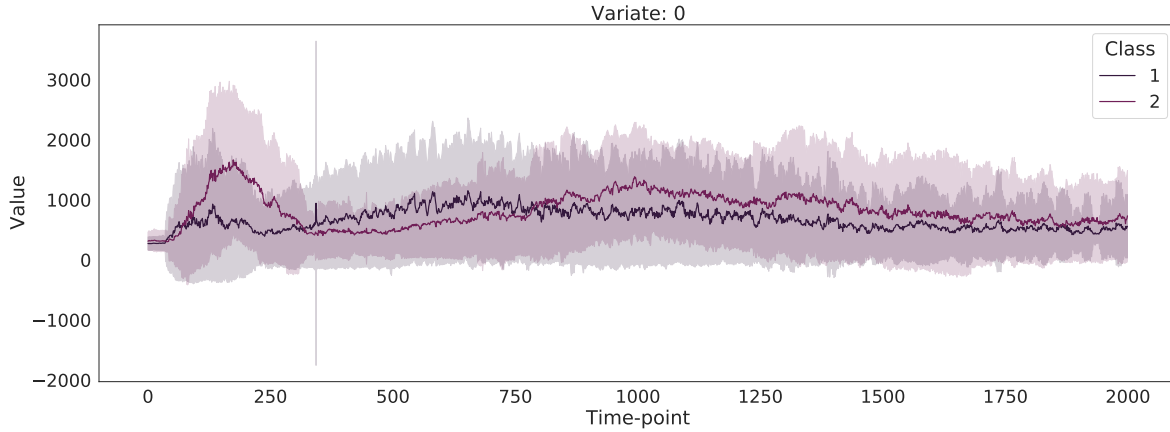

Figure 12: HouseTwenty dataset: Mean and standard deviation of each variable's time-points per class.

LSST   In this case the measurements originate from simulations of astronomical time-series observations that are expected to be captured by the Large Synoptic Survey Telescope. These values describe the brightness of an object in terms of 6 photometric passbands. The class labels indicate the type of the astronomical object. The summary is given in Figure 13.
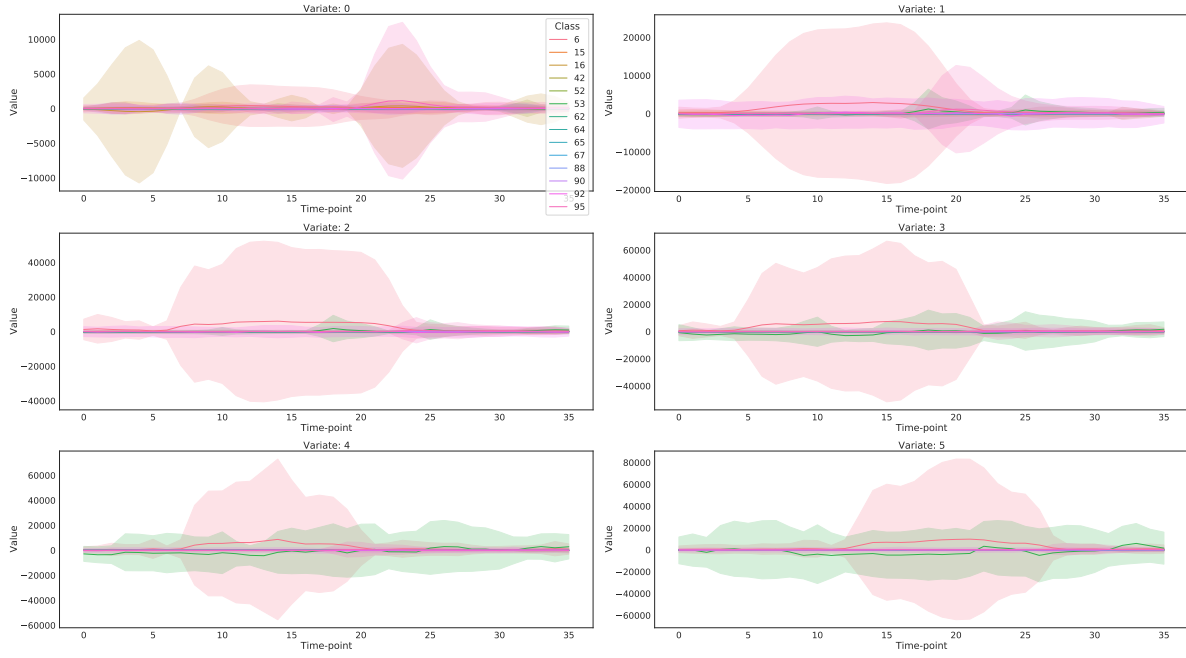
Figure 13: LSST dataset: Mean and standard deviation of each variable's time-points per class.

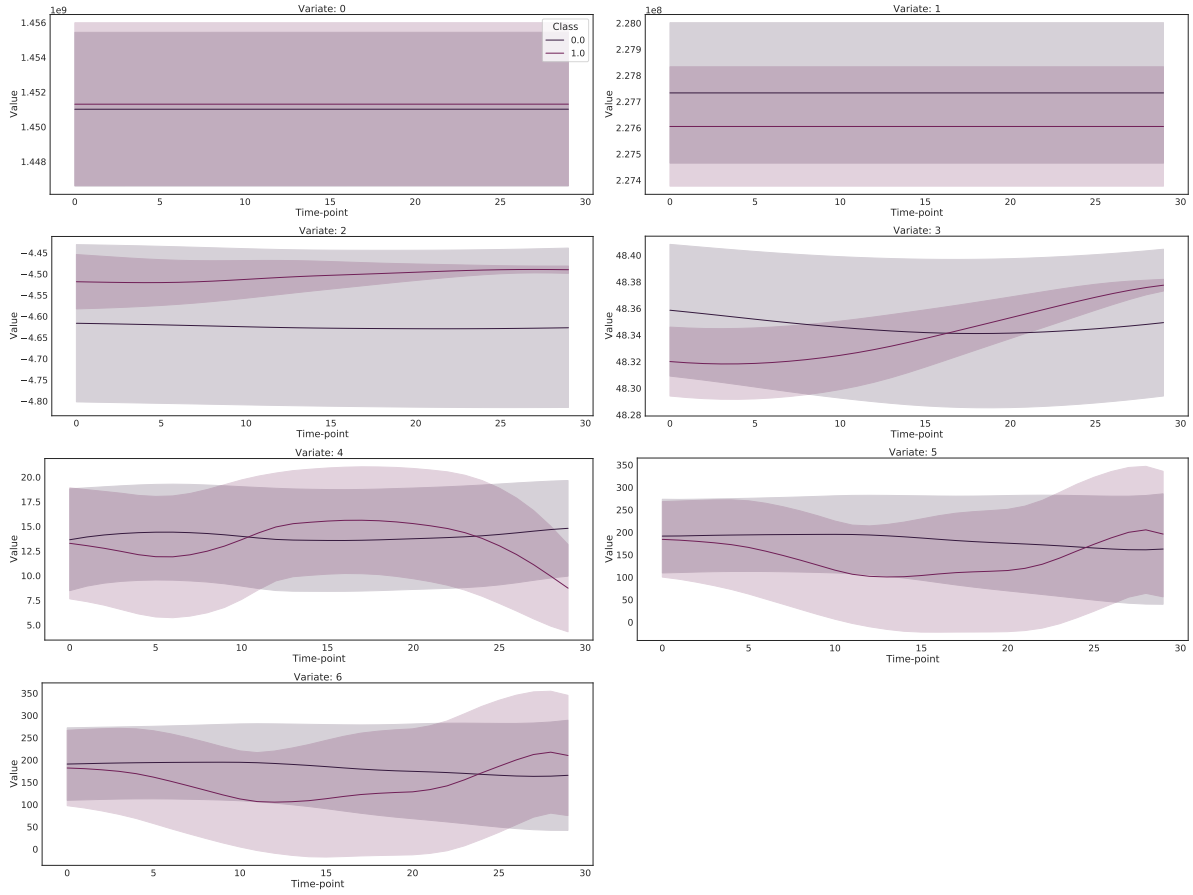**Maritime** The summary of this dataset is given in Figure 14.

Figure 14: Maritime dataset: Mean and standard deviation of each variable's time-points per class.

PickupGestureWiimoteZ This dataset contains Z-axis acceleration measurements from a video game controller, as 10 persons pick it up a number of times. The labels correspond to each different person, and the frequency of measurements is 100 per second. In this case the length of the time-series varies (see Figure 15).
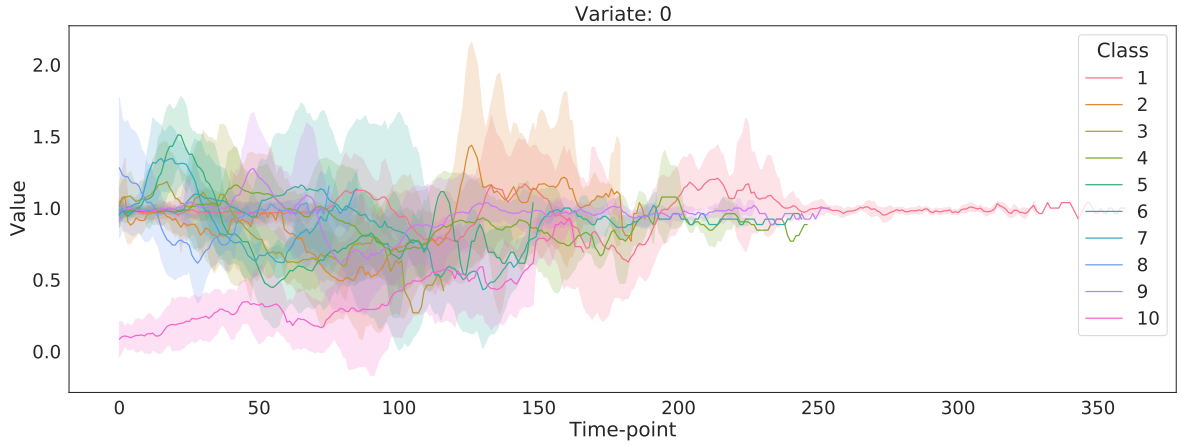
Figure 15: PickupGestureWiimoteZ dataset: Mean and standard deviation of each variable's time-points per class.

PLAID This case includes high frequency (30kHz) current and voltage measurements from different types of electrical appliances operating at different states. Each class label encodes the type of the appliance, and in this case two, the length of the instances varies. The dataset summary is given in Figure 16.
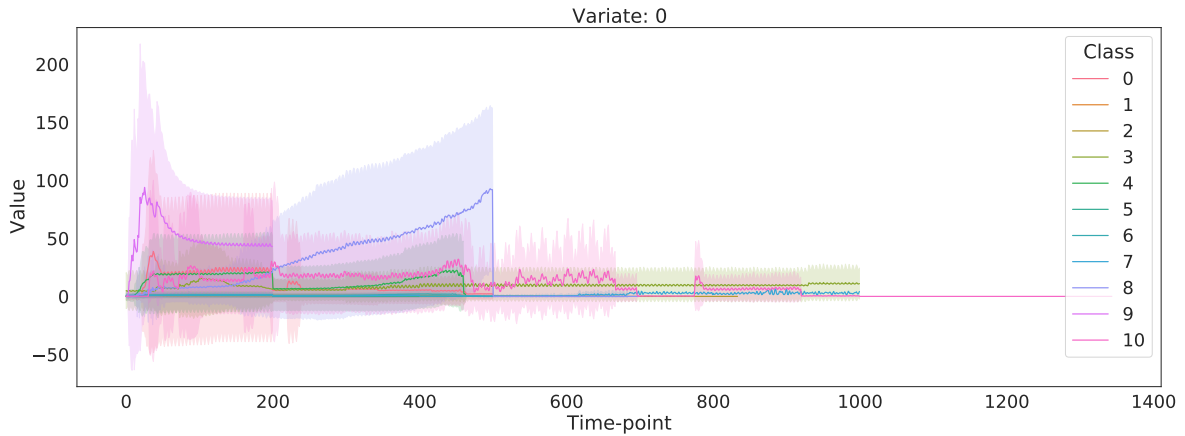

Figure 16: PLAID dataset: Mean and standard deviation of each variable's time-points per class.

PowerCons This is another dataset that contains household electricity consumption measurements, divided into two classes, during warm (label "1"), and cold (label "2") periods. The frequency of observations is 10 minutes. The summary is given in Figure 17.
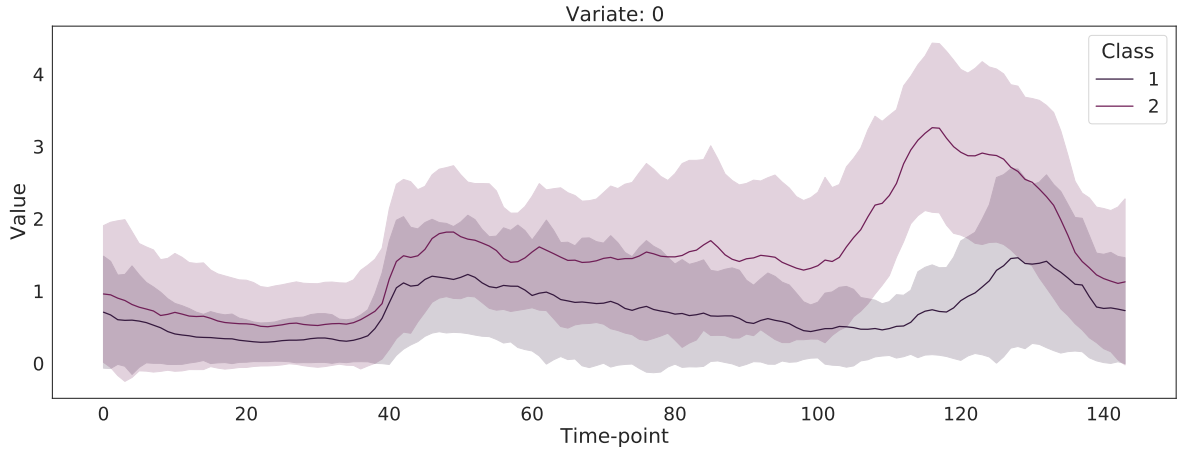
Figure 17: PowerCons dataset: Mean and standard deviation of each variable's time-points per class.

SharePriceIncrease The last dataset that we incorporate from the UEA & UCR repository includes signals of share price evolution over time in the form of close price percentage change during a day for the past 60 days. The class label indicates if the share price did not increase after Earning Per Share report releases more than 5% (label "0"), or if it increased more (label "1"). The summary is given in Figure 18.
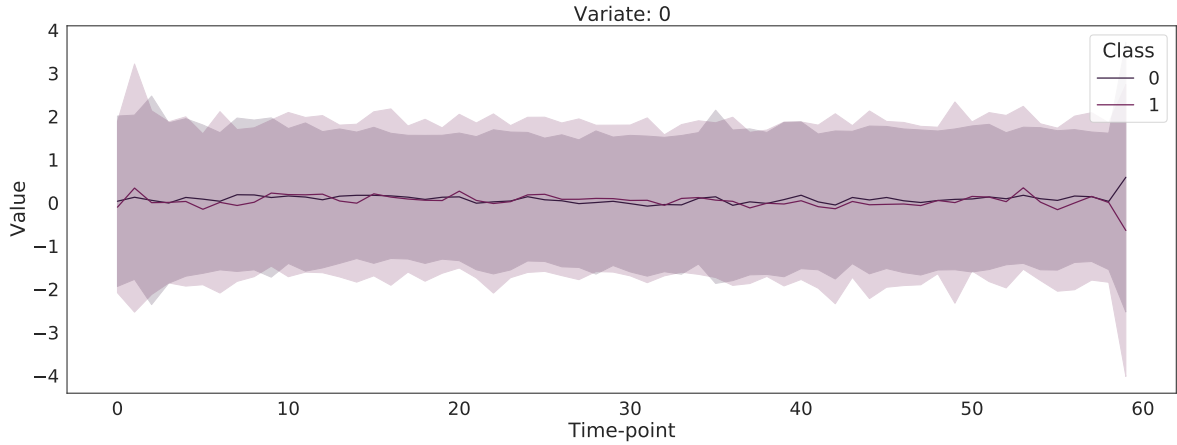

Figure 18: SharePriceIncrease dataset: Mean and standard deviation of each variable's time-points per class.

Table 5 presents the characteristics of all datasets used in our experimental evaluation, while Table 6 shows the dataset categorization.

15

Dataset Name	No. of Variables	No. of Instances (Height)	No. of Timepoints (Length)	No. of Classes	Class Imbalance Ratio	Coefficient of Variation
BasicMotions	6	80	100	4	1	608.91
Biological	3	644	49	2	5.37	1.06
DodgerLoopDay	1	158	288	7	1.25	0.63
DodgerLoopGame	1	158	288	2	1.07	0.63
DodgerLoopWeekend	1	158	288	2	2.43	0.63
HouseTwenty	1	159	2,000	2	1.27	1.10
LSST	6	4,925	36	14	111	69.16
Maritime	7	80,591	30	2	4.21	2.08
PickupGestureWiimoteZ	1	100	361	10	1	0.41
PLAID	1	1,074	1,345	11	6.73	3.01
PowerCons	1	360	144	2	1	0.97
SharePriceIncrease	1	1,931	60	2	2.19	24.21

Table 5: Dataset characteristics.

Group	Specifications	Datasets
Wide	Length > 1,300	HouseTwenty, PLAID
Large	Height > 1,000	LSST, Maritime, PLAID, SharePriceIncrease
Unstable	Coeff. of Var. > 1.08	BasicMotions, HouseTwenty, LSST, Maritime, PLAID, SharePriceIncrease
Imbalanced	Class Imb. Ratio > 1.73	Biological, DodgerLoopWeekend, LSST, Maritime, PLAID, SharePriceIncrease
Multiclass	Number of Classes > 2	BasicMotions, DodgerLoopDay, LSST, PickupGestureWiimoteZ, PLAID
Common	None of the above	DodgerLoopGame, PowerCons
Univariate	One variable	DodgerLoopDay, DodgerLoopGame, DodgerLoopWeekend, HouseTwenty, PickupGestureWiimoteZ, PLAID, PowerCons, SharePriceIncrease
Multivariate	> 1 variables	BasicMotions, Biological, LSST, Maritime

Table 6: Categorization of datasets.