A Formally Robust Time Series Distance Metric

Maximilian Toller
Know-Center GmbH
Graz, Austria
mtoller@know-center.at

Bernhard C. Geiger
Know-Center GmbH
Graz, Austria
bgeiger@know-center.at
geiger@ieee.org

Roman Kern
Graz University of Technology
Graz, Austria
rkern@tugraz.at

ABSTRACT

Distance-based classification is among the most competitive classification methods for time series data. The most critical component of distance-based classification is the selected distance function. Past research has proposed various different distance metrics or measures dedicated to particular aspects of real-world time series data, yet there is an important aspect that has not been considered so far: *Robustness against arbitrary data contamination.* In this work, we propose a novel distance metric that is robust against arbitrarily "bad" contamination and has a worst-case computational complexity of $O(n \log n)$. We formally argue why our proposed metric is robust, and demonstrate in an empirical evaluation that the metric yields competitive classification accuracy when applied in k-Nearest Neighbor time series classification.

CCS CONCEPTS

• **Mathematics of computing** → **Time series analysis**; • **Computing methodologies** → *Classification and regression trees*; • **Information systems** → Clustering.

KEYWORDS

time series, distance metric, robustness, classification, clustering

1 INTRODUCTION

Time series data classification is an important task in many domains such as data mining, machine learning and econometrics. Extensive past evaluations [6] have shown that k-Nearest Neighbor (k-NN) classification is among the most competitive classification approaches for time series data. In simple terms, k-NN classification assigns a query time series instance the class based on its k nearest neighbors in a labeled training set. As such, the k-NN classifier is a distance-based classifier, since its distance function is the only

component that discriminates between classes. The same applies to distance-based clustering algorithms [15, 21].

The data mining and machine learning communities have proposed numerous different distance functions for improving classification and clustering accuracies on benchmark datasets [1, 6] and for accelerating the practical computation [14, 16, 17, 20, 23]. However, there is an important aspect of distance measures that was not considered so far to the best of our knowledge: *Robustness against arbitrary data contamination.* While previous research has proposed distance measures that are "robust" against additive white Gaussian noise [13] or against temporal misalignment [22], we follow the definition used in the field of robust statistics. A crucial measure for determining the robustness of a distance function is breakdown point (BP) analysis [11]. The asymptotic BP describes the amount of contamination in the data that an estimator (in this case a distance function) can tolerate before it will be fully biased by the contamination in the worst case. For example, the Euclidean distance has an asymptotic BP of zero: If a single observation in one of the time series instances it compares is contaminated to (plus or minus) infinity, then the Euclidean distance becomes infinite as well, regardless of the remaining observations.

To address this issue, one may propose to use the raw Edit distance [18], as it is robust against arbitrary contamination at a few observations. However, Edit distance is susceptible to a different type of contamination that is routinely overlooked, as it is trivially fulfilled by most distance measures. If time series data are subject to a tiny contamination at every single data point, then the Edit distance will become very large. One may be tempted to address this issue by defining a small tolerance interval suggested by Chen et al. [3], yet this is difficult if the variance of the data is large or time-dependent, which is a well-known behavior of many econometric time series [8, 19]. Further, the time series classification accuracy of the raw Edit distance is poor for real time series data. If one extends Edit distance to an elastic (non-lockstep) variant thereof, such as Edit distance with real penalty [2], then the classification accuracy may increase, yet the asymptotic BP immediately drops to 0 .

Since all existing distance measures either have a low asymptotic BP or else yield a low classification accuracy, we aim to fill this gap. In this work, we propose a novel distance metric which is formally robust according to Huber's definition [11] against a small percentage of contaminated observations, and robust against tiny deviations at many observations. Additionally, we show that its classification accuracy is not significantly different from other distance metrics and that our metric has a worst-case computational complexity of $O(n \log n)$. The source code of our implementation and a script that reproduces all results can be found online.[1]

---

[1] https://github.com/mtoller/robust-distance-metric

2 NOTATION AND PROBLEM FORMULATION

2.1 Theoretical Concepts

Let $x = \{x_t, t \in 1, \ldots, n\}$ and $y = \{y_t, t \in 1, \ldots, n\}$ be two time series instances and $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ a distance function for comparing them. For an efficient distance-based classification, it is advantageous if $d(\cdot)$ is a *metric*, as this allows a variety of run-time acceleration techniques [9, 10]. To be a metric, $d(\cdot)$ has to fulfill the following properties for all $x, y \in \mathbb{R}^n$:

- $d(x, y) \geq 0$ *Non-negativity*
- $d(x, y) = 0 \Leftrightarrow x = y$ *Identity of Indiscernibles*
- $d(x, y) = d(y, x)$ *Symmetry*
- $d(x, z) \leq d(x, y) + d(y, z)$ *Triangle Inequality*

A typical example for a distance metric is the Euclidean distance

$$e(x, y) = \sqrt{\sum_{t=1}^{n} (x_t - y_t)^2}. \quad (1)$$

If a distance function $d(\cdot)$ fulfills all properties except the identity of indiscernibles, then it is called a *pseudometric*.

To evaluate the robustness of a distance function, we adapt the definition of the breakdown point given in [11]. Specifically, let $d_{\sup} = \sup_{x,y \in \mathbb{R}^n} d(x, y)$ be the largest possible value the distance function can obtain theoretically. Then, the breakdown point $\beta_d^\star(n)$ is given by

$$\beta_d^\star(n) = \min \left\{ \frac{k}{n} \bigg| \sup d(x, x + K) = d_{\sup} \right\} \quad (2)$$

where the supremum is over all $x \in \mathbb{R}^n$ and over all contamination processes $K = \{K_t, t \in 1, \ldots, n\}$ that assume arbitrary non-zero values in at most k positions and zero otherwise. In simple terms, the breakdown point describes the highest percentage of contaminated observations that function $d(\cdot)$ can tolerate. For example, it is evident that for the Euclidean distance contaminating a single time point suffices, i.e., if $K_1 = \infty$ and $K_t = 0$ for $t = 2, \ldots, n$, then $e(x, x + K) = d_{\sup} = \infty$ for every $x \in \mathbb{R}^n$; thus, $\beta_e^\star(n) = 1/n$. For clarity, the *asymptotic* BP is obtained by evaluating $\beta_d^\star(n)$ as n tends to infinity.

2.2 Classification-Specific Aspects

To link the theoretical concept of breakdown points with practical classification, we formulate two classification-specific notions of robustness. To this end, let $C = \{C_1, \ldots, C_r\}$ denote a set of time series classes and $d(\cdot)$ a candidate distance function.

Definition 2.1. Contamination Tolerance: A distance function $d(\cdot)$ tolerates $\hat{k}$ contaminated observation w.r.t. C if

$$\forall i, j \neq i \in 1, \ldots, r : \forall x \in C_i : \forall y \in C_j : d(x, x + K) < d(x, y). \quad (3)$$

holds for every contamination processes $K = \{K_t, t \in 1, \ldots, n\}$ that assumes arbitrary non-zero values in at most $\hat{k}$ positions.

Intuitively, assume that a distance function $d(\cdot)$ ideally separates class C_i from other classes $C_{j, j \neq i}$. Function $d(\cdot)$ will tolerate up to $\hat{k}$ contaminated observations if the distance between an uncontaminated time series instance x and a contaminant variant thereof $x + K$ is smaller than the distance between x and an instance from a different class y.

Table 1: The components of the proposed ensemble metric $\mathcal{E}$. The top three members are metrics, while the bottom three are pseudometrics.

Member name	Definition
Euclidean distance	$e(x, y)$
Log-distance	$\ell(x, y)$
Raw Edit distance	$\text{Edit}(x, y)$
Robust Euclidean distance	$e(\vec{m}(x), \vec{m}(y))$
Robust Log-distance	$\ell(\vec{m}(x), \vec{m}(y))$
Robust Raw Edit distance	$\text{Edit}(\vec{m}(x), \vec{m}(y))$

To specify imprecision invariance as mentioned in Section 1, i.e. invariance to tiny changes, we introduce an imprecision process $\{\varepsilon_t\}$ that is negligibly small at all t. Specifically, we assume that, for all t, $|\varepsilon_t| \leq \varepsilon_{\max}$, where $\varepsilon_{\max}$ is much smaller than the standard deviation (or some norm) of the time series x.

Definition 2.2. Imprecision Invariance: A distance function $d(\cdot)$ is invariant to an imprecision of $\varepsilon_{\max}$ w.r.t. C if

$$\forall i, j \neq i \in 1, \ldots, r : \forall x \in C_i : \forall y \in C_j : d(x, x + \varepsilon) < d(x, y) \quad (4)$$

holds for every imprecision processes $\varepsilon = \{\varepsilon_t, t \in 1, \ldots, n\}$ that satisfies $|\varepsilon_t| \leq \varepsilon_{\max}$ for every t.

In other words, assume distance function $d(\cdot)$ perfectly discriminates class C_i from $C_{j, j \neq i}$. Function $d(\cdot)$ is invariant to an imprecision of $\varepsilon_{\max}$ if the distance between an instance x and almost the same instance $x + \varepsilon$ is smaller than the distance between x and an instance from another class y.

Contamination tolerance and imprecision invariance are very different properties. There are not many metrics that fulfill both simultaneously: For example, no metric induced by an $\mathcal{L}^p$ norm with a finite $p \geq 1$ is contamination tolerant for any non-zero $\hat{k}$. Also, while many popular metrics are imprecision invariant, some metrics such as the Edit distance are susceptible to it.

3 METHODS

In this section, we present a novel metric which can tolerate considerable contamination and is invariant to imprecision. The metric is obtained by aggregating an ensemble of metrics and pseudometrics in a way that preserves their discriminatory power while guaranteeing robust results.

3.1 Metric Ensemble Members

The ensemble consists of three metrics and three pseudometrics. The distances measured by these metrics are combined via a scaling function and an arbitrary $\mathcal{L}^p$ norm, with $p \geq 1$, to obtain the metric $\mathcal{E}(x, y)$. A summary of the ensemble members can be seen in Table 1.

Definition 3.1. Log-distance: Let $x, y \in \mathbb{R}^n$ be two real-valued n-dimensional observations. Then, the Log-distance $\ell(\cdot)$ between x and y is given by

$$\ell(x, y) = \sum_{t=1}^{n} \log(1 + |x_t - y_t|). \quad (5)$$

PROPOSITION 3.2. *The Log-distance $\ell(\cdot)$ is a metric.*

PROOF. Since $\log(x) : \mathbb{R}^+ \to \mathbb{R}$ is a strictly monotonic subadditive function, $\log(1 + x)$ is also a strictly monotonic subadditive function that is zero iff $x = 0$. Consequently, $\log(1 + |x - y|)$ also fulfills these properties and is a metric by Kelly's theorem [12, p. 131]. That the sum of metrics is a metric [7] completes the proof. □

For 1-dimensional data, the Log-distance is asymptotically smaller than any $\mathcal{L}^p$ metric with $p \geq 1$, since the logarithm grows slower than an arbitrary polynomial, i.e. $\lim_{z \to \infty} \frac{\log(z)}{P(z)} = 0, \quad z \in \mathbb{R}$. This property is beneficial when one expects a small number of large outliers in time series data x_t. $\mathcal{L}^p$ metrics such as the Euclidean distance will be much more influenced by a single large difference than several small deviations that sum up to the same value. The Log-distance will weight several small changes higher than one large change due to subadditivity of the logarithm.

The remaining two metrics of the ensemble are the Euclidean distance $e(\cdot)$ as defined in Equation (1) and the raw Edit distance

$$\text{Edit}(x, y) = \sum_{t=1}^n \phi_t, \quad \phi_t = \begin{cases} 0 & x_t = y_t \\ 1 & x_t \neq y_t \end{cases} \tag{6}$$

which is equivalent to the number of observations where x_t and y_t differ. While the Edit distance tolerates up to $n - 1$ contaminated observations and is sensitive to imprecision, the Euclidean distance $e(\cdot)$ is invariant to imprecision but sensitive to contamination. The Log-distance $\ell(\cdot)$ aims to present a middle-ground between the two. Compared to the Euclidean distance it is "more" sensitive to imprecision and "less" sensitive to contamination, while the inverse holds when it is compared against the Edit distance. However, in terms of robustness, the Euclidean distance and the Log-distance are asymptotically equivalent, since they have the same BP $\beta_e^\star(n) = \beta_\ell^\star(n) = \frac{1}{n}$. Hence, when confronted with arbitrary contamination, both metrics become equally useless in the worst case.

3.2 Pseudometric Ensemble Members

To raise the BP of the metrics in the ensemble $\mathcal{E}$, one can introduce a function composition with a function that has a high BP while preserving metric properties. Let $m(x)$ be the median of x. As a measure of central tendency, the median has a BP of $\beta_m^\star(n) = 0.5 + \frac{1}{n}$ according to Huber's definition [11]. However, computing the median of time series data x_t is meaningless, since it disregards the temporal structure of x_t by treating it like an unordered data set. To exploit the asymptotic robustness of the median in the context of time series, one can instead apply the median via a sliding window:

Definition 3.3. Let x be a time series instance and let w, an odd integer in $[3; n]$, be the size of a sliding window. The sliding median $\vec{m} : \mathbb{R}^n \to \mathbb{R}^{n-w+1}$ of x is then defined as

$$\vec{m}(x) = \{m(x_1, \ldots, x_w), m(x_2, \ldots, x_{w+1}), \ldots, m(x_{n-w+1}, \ldots, x_n)\}. \tag{7}$$

If one computes the Euclidean, Log and Edit distance of $\vec{m}$, then the result is no longer a metric — the identity of indiscernibles becomes violated since the median is not an injective function. However, the remaining metric properties are preserved:

PROPOSITION 3.4. *Let $d(\cdot)$ be a metric. Then the sliding median distance $M_d(x, y) = d(\vec{m}(x), \vec{m}(y))$ is a pseudometric.*

PROOF. Non-negativity, symmetry and triangle inequality follow trivially from the application of $d(\cdot)$. □

The sliding median distance $M_d(\cdot)$ has a BP of $\frac{w}{2n}$, since, if all contamination occurred at $\frac{w+1}{2}$ consecutive observations, the median of all $\frac{w-1}{2}$ windows containing these observations could be contaminated to an arbitrary value.

3.3 Combining the Members

Since all six ensemble members operate on different scales, it is desirable to convert them to the same scale without loss of generality. Therefore we propose the following scaling function $S(\cdot)$ that is applied after distance computation and that preserves all metric properties:

Definition 3.5. Let $d(\cdot)$ be an arbitrary distance function. The metric-preserving scaling $S : \mathbb{R}^+ \to [0; 1]$ of this metric is then defined as

$$S(d(x, y)) = 1 - \frac{1}{1 + d(x, y)}. \tag{8}$$

LEMMA 3.6. *Metrics are closed under scaling with $S(\cdot)$.*

PROOF. $S(\cdot)$ is a concave, monotonically increasing function with $S(d(x, y)) = 0 \Leftrightarrow d(x, y) = 0$. Hence, it is a metric by Kelly's theorem [12]. □

After scaling, the ensemble members can be combined into a single metric $\mathcal{E}(\cdot)$ via an arbitrary $\mathcal{L}^p$ norm with $p \geq 1$. Specifically, we suggest the $\mathcal{L}^2$ norm. The resulting function is a metric, since the sum of a pseudometric and a metric is a metric. In particular, we propose the following ensemble:

$$\mathcal{E}(x, y) := \sqrt{\begin{aligned} & S(e(x, y))^2 + S(\ell(x, y))^2 + S(\text{Edit}(x, y))^2 \\ & + S(e(\vec{m}(x)), \vec{m}(y))^2 + S(\ell(\vec{m}(x), \vec{m}(y))^2 \\ & + S(\text{Edit}(\vec{m}(x), \vec{m}(y)))^2 \end{aligned}} \tag{9}$$

The ensemble $\mathcal{E} : \mathbb{R}^n \to [0; \sqrt{6}]$ has a BP of $\beta_{\mathcal{E}}^\star = 1$. This follows from the fact that the measurements of the non-robust metrics $e(\cdot)$ and $\ell(\cdot)$ are mapped onto the interval $[0, \sqrt{2}]$ and thus their influence on the ensemble is restricted. The remaining members have a BP of $\frac{w}{2n}$ or higher, and the inclusion of the Edit distance raises the total BP to 1.

The ensemble has a worst-case computational complexity of $O(n \log n)$ under the assumption that $w = O(n)$. This arises from the ensemble's most expensive step, which is the computation of the sliding median $\vec{m}$. A more detailed explanation can be found in the Appendix.

4 PRACTICAL EVALUATION

In this section, we describe the experiments we conducted to show that the proposed ensemble metric $\mathcal{E}$ has competitive classification accuracy. Further, we validate that $\mathcal{E}$ tolerates contamination and is imprecision invariant. We compared $\mathcal{E}$ with Euclidean distance (Euc), Dynamic Time Warping [22] (DTW) with window size $w =$

100, Log-distance (Log), raw Edit distance (ED), and edit distance with a tolerance interval (EDR) set to 10% of the median absolute deviation.

4.1 Setup

In our practical evaluation, we conducted three experiments to assess the following properties of the ensemble $\mathcal{E}$:

- 1-NN classification error rate
- Contamination tolerance (cf. Equation (3))
- Imprecision invariance (cf. Equation (4))

For all three experiments we used 83 selected benchmark data from the UCR Time Series Classification Archive [4]. All datasets which contained non-real data such as missing values were omitted, which was necessary since otherwise the behavior of the ensemble would be undefined. Further, we were forced to omit all datasets in which either training or test datasets contained more that 1000 instances due to our limited computational resources.

For the classification accuracy experiment we computed the raw accuracy of a 1-NN classifier based on the ensemble metric and subtracted the resulting value from 1 to obtain the error rate. To determine statistical significance, we then performed a Friedman's rank test [5].

The dataset-dependent contamination tolerance was computed with the following procedure:

- i) Assume $d(\cdot)$ perfectly separates all classes in the dataset.
- ii) For every instance $x \in C_i$, count for how many $y \in C_{j,j \neq i}$ Equation (3) holds when $k = 0.05 \times n$, i.e. 5% of the observations are contaminated to be $\pm\infty$
- iii) Compute the ratio of this count and the number of instances in $C_{j \neq i}$.
- iv) Compute the mean over all instance-based ratios.

The window size of the ensemble $\mathcal{E}$ was set to $w = 0.1 \times n + 1$ observations, which ensures that the window is always large enough to be resilient against 5% contamination. The imprecision invariance was computed similarly, only with Equation (3) replaced by Equation (4), k set to n and $\varepsilon_t \sim \mathcal{U}(-10^{-10}, 10^{-10})$. Theoretically, ε_t should be as close to zero as possible. Yet the results below suggest that the above interval is sufficiently small for asserting the imprecision invariance property of the distance functions under consideration.

4.2 Results

In this subsection we present a summary of the results of the three experiments we conducted. The complete results can be found in the Appendix.

Our first experiment showed that, in terms of classification error rate, there is no significant difference between the ensemble $\mathcal{E}$ and Euc, DTW or Log, but that ED and EDR are significantly worse than $\mathcal{E}$. A visual representation of this result is depicted in Figure 1.

The second and third experiment revealed that the only distance functions which are both contamination tolerant and imprecision invariant are the ensemble $\mathcal{E}$ and EDR. An overview of these results can be found in Table 2.
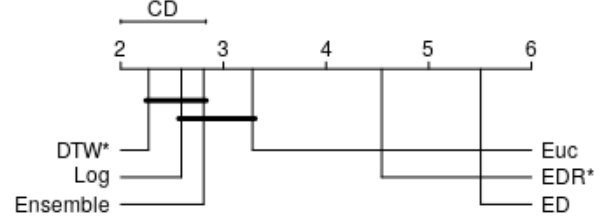
Figure 1: Critical distance plot for the classification error rate. Average ranks are depicted in order, where lower is better and the horizontal bars highlight no significant difference. The ensemble $\mathcal{E}$ is not significantly different from either Euc, DTW or Log, but significantly better than ED and EDR. Functions labeled with an asterisk (*) are not metrics.

Table 2: Summary of the second and third experiments. The ensemble $\mathcal{E}$ and EDR tolerate contamination on 56 and 79 data sets, respectively, while the other distance measures never tolerate contamination. In terms of imprecision invariance, the first four distances are perfectly invariant, while ED is susceptible and EDR is almost perfectly invariant.

	$\mathcal{E}$	Euc	DTW	Log	ED	EDR
Is a metric?	✓	✓	×	✓	✓	×
Contam. tol. on # datasets	56	0	0	0	83	81
Imprec. Invar. on # datasets	83	83	83	83	0	79
Both on # datasets	56	0	0	0	0	79

5 DISCUSSION AND CONCLUSION

The goal of this work was to propose a distance function that

- is robust against arbitrary contamination
- is invariant to imprecision
- fulfills all metric properties
- has a competitive classification accuracy
- is computationally efficient.

The combined results of our theoretical analysis and of the practical evaluation suggest that the ensemble $\mathcal{E}$ has all of these properties. One might argue that the ensemble $\mathcal{E}$ is no improvement over EDR, since both methods depend on one parameter which influences their classification accuracy. However, correctly choosing an appropriate tolerance interval for time series with large or time-dependent variance is difficult, while tuning the ensemble's window size is simpler — it should be just larger than the expected amount of contamination. Additionally, $\mathcal{E}$ has a significantly better classification accuracy than EDR. The same holds for the Log-distance, which we believe should be seen as a natural alternative to the Euclidean distance.

Future work might consider less extreme cases of contamination and determine precisely how this affects classification accuracy. Further, evaluating robust distance functions on a clustering task

with arbitrarily contaminated data seems a promising avenue for the future.

6 ACKNOWLEDGMENTS

REFERENCES

[1] Amaia Abanda, Usue Mori, and Jose A Lozano. 2019. A review on distance based time series classification. *Data Mining and Knowledge Discovery* 33, 2 (2019), 378–412.
[2] Lei Chen and Raymond Ng. 2004. On the marriage of lp-norms and edit distance. In *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*. VLDB Endowment, 792–803.
[3] Lei Chen, M Tamer Özsu, and Vincent Oria. 2005. Robust and fast similarity search for moving object trajectories. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*. ACM, 491–502.
[4] Hoang Anh Dau, Eamonn Keogh, Kaveh Kamgar, Chin-Chia Michael Yeh, Yan Zhu, Shaghayegh Gharghabi, Chotirat Ann Ratanamahatana, Yanping, Bing Hu, Nurjahan Begum, Anthony Bagnall, Abdullah Mueen, and Gustavo Batista. 2018. The UCR Time Series Classification Archive. (October 2018). https://www.cs.ucr.edu/~eamonn/time_series_data_2018/.
[5] Janez Demšar. 2006. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine learning research* 7, Jan (2006), 1–30.
[6] Hui Ding, Goce Trajcevski, Peter Scheuermann, Xiaoyue Wang, and Eamonn Keogh. 2008. Querying and mining of time series data: experimental comparison of representations and distance measures. *Proceedings of the VLDB Endowment* 1, 2 (2008), 1542–1552.
[7] Jozef Doboš. 1998. *Metric preserving functions*. Štroffek Košice.
[8] Robert F Engle. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society* (1982), 987–1007.
[9] Christos Faloutsos, Mudumbai Ranganathan, and Yannis Manolopoulos. 1994. *Fast subsequence matching in time-series databases*. Vol. 23. ACM.
[10] Gisli R Hjaltason and Hanan Samet. 2003. Index-driven similarity search in metric spaces (survey article). *ACM Transactions on Database Systems (TODS)* 28, 4 (2003), 517–580.
[11] Peter J Huber. 2011. Robust statistics. In *International Encyclopedia of Statistical Science*. Springer, 1248–1251.
[12] John L Kelley. 2017. *General topology*. Courier Dover Publications.
[13] Eamonn Keogh. 1997. A fast and robust method for pattern matching in time series databases. *Proceedings of WUSS* 97, 1 (1997), 99.
[14] Eamonn Keogh, Kaushik Chakrabarti, Michael Pazzani, and Sharad Mehrotra. 2001. Dimensionality reduction for fast similarity search in large time series databases. *Knowledge and information Systems* 3, 3 (2001), 263–286.
[15] Eamonn Keogh and Shruti Kasetty. 2003. On the need for time series data mining benchmarks: a survey and empirical demonstration. *Data Mining and knowledge discovery* 7, 4 (2003), 349–371.
[16] Eamonn J Keogh and Michael J Pazzani. 2000. Scaling up dynamic time warping for datamining applications. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 285–289.
[17] Eamonn J Keogh and Padhraic Smyth. 1997. A probabilistic approach to fast pattern matching in time series databases.. In *Kdd*, Vol. 1997. 24–30.
[18] Karen Kukich. 1992. Techniques for automatically correcting words in text. *Acm Computing Surveys (CSUR)* 24, 4 (1992), 377–439.
[19] Benoit Mandelbrot. 1967. The variation of some other speculative prices. *The Journal of Business* 40, 4 (1967), 393–413.
[20] Abdullah Mueen and Eamonn Keogh. 2016. Extracting optimal performance from dynamic time warping. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2129–2130.
[21] Sangeeta Rani and Geeta Sikka. 2012. Recent techniques of clustering of time series data: a survey. *International Journal of Computer Applications* 52, 15 (2012).
[22] Hiroaki Sakoe, Seibi Chiba, A Waibel, and KF Lee. 1990. Dynamic programming algorithm optimization for spoken word recognition. *Readings in speech recognition* 159 (1990), 224.
[23] Xiaopeng Xi, Eamonn Keogh, Christian Shelton, Li Wei, and Chotirat Ann Ratanamahatana. 2006. Fast time series classification using numerosity reduction. In *Proceedings of the 23rd international conference on Machine learning*. ACM, 1033–1040.

A DETAILED TIME COMPLEXITY ANALYSIS

A procedural description of the ensemble $\mathcal{E}$ is listed in Algorithm 1. Since there are no directly listed loops, the ensemble is linear in the complexity of the functions it applies. Hence, its worst-case time complexity is equal to that of the function that requires the most expensive computation. When looking at the distance functions, it quickly becomes evident that these can be computed in $O(n)$, since these lockstep methods look at each observation only once. The scaling function can be computed in $O(1)$, so the only non-trivial step is the computation of the sliding median.

Algorithm 1 Ensemble Metric $\mathcal{E}$

Require: x_t, y_t, w
  if $w \cong 0 \pmod 2$ **then**
    $w \leftarrow w + 1$
  end if
  $\text{med}_x \leftarrow \vec{m}(x)$
  $\text{med}_y \leftarrow \vec{m}(y)$
  $\text{dist}_1 \leftarrow S_e(x, y)$
  $\text{dist}_2 \leftarrow S_\ell(x, y)$
  $\text{dist}_3 \leftarrow S_{\text{Edit}}(x, y)$
  $\text{dist}_4 \leftarrow S_e(\text{med}_x, \text{med}_y)$
  $\text{dist}_5 \leftarrow S_\ell(\text{med}_x, \text{med}_y)$
  $\text{dist}_6 \leftarrow S_{\text{Edit}}(\text{med}_x, \text{med}_y)$
  return $\sqrt{\sum_i^6 \text{dist}_i^2}$

Directly computing the sliding median requires one to sort the observations in all windows. Since the most efficient sorting algorithm requires $O(w \log w)$ steps, sorting all $n - w + 1$ windows would take $(n - w + 1) \times O(w \log w) = n \times O(w \log w)$ steps. We assume that a certain small percentage of the n observations is contaminated, so we must conclude that $w = O(n)$, which results in a total complexity of $O(n^2 \log n)$.

However, there exist more efficient algorithms for computing the sliding median, and several implementations are available in C libraries. If one starts by sorting the initial window, one consumes $O(n \log n)$ time as argued above. After this, if one keeps the computed window and an index list in the memory, the next window can be computed by removing the oldest observation in $O(1)$ time and sorting in the new observation in $O(\log n)$ time. So, for all windows, this algorithms requires $O(n \log n) + (n - w) \times O(\log n) = O(n \log n)$ steps.

B PROOFS

This section contains alternate proofs of the propositions and lemmas presented in the main article. These proofs do no rely on Kelly's theorem and are easy to verify.

B.1 Proof of Proposition 3.2

PROOF. *Non-negativity*: $\ell(x, y) \geq 0$

$$\ell(x, y) = \log(1 + |x - y|) \geq \log(1) = 0 \geq 0$$

Identity of Indiscernibles: $\ell(x, y) = 0 \Leftrightarrow x = y$

This is trivial, since the log function has exactly one zero at $\log(1 + |x - x|) = log(1) = 0$

Symmetry
Trivial, due to absolute value.

Triangle Inequality: $\ell(x, z) \leq \ell(x, y) + \ell(y, z)$

$$\ell(x, z) = \log(1 + |x - z|) \leq \log(1 + |x - y|) + \log(1 + |y - z|)$$

$$\log(1 + |x - z|) \leq \log\big((1 + |x - y|)(1 + |y - z|)\big)$$

$$1 + |x - z| \leq (1 + |x - y|)(1 + |y - z|)$$

$$|x - z| \leq |x - y| + |y - z| + |x - y||y - z|$$

$$|x - z| \leq |x - y| + |y - z|$$

□

B.2   Proof of Proposition 3.4

PROOF. *Non-negativity*: $M_d(x, y) \geq 0$
This follows trivially from the fact that $d(\cdot) \geq 0$.

Symmetry: $M_d(x, y) = M_d(y, x)$
This also follows trivially from the symmetry of $d(\cdot)$.

Triangle Inequality: $M_d(x, z) \leq M_d(x, y) + M_d(y, z)$

$$d(\vec{m}(x), \vec{m}(z)) \leq d(\vec{m}(x), \vec{m}(y)) + d(\vec{m}(y), \vec{m}(z))$$

$$a := \vec{m}(x), \quad b := \vec{m}(y), \quad c := \vec{m}(z)$$

$$d(a, c) \leq d(a, b) + d(b, c)$$

□

B.3   Proof of Lemma 3.6

PROOF. *Non-negativity*: $S(x, y) \geq 0$

$$1 - \frac{1}{1 + d(x, y)} \geq 0$$

$$\frac{1}{1 + d(x, y)} \leq 1$$

$$1 + d(x, y) \geq 1$$

$$d(x, y) \geq 0$$

Identity of Indiscernibles: $S(x, y) = 0 \Leftrightarrow x = y$
First we show $S(x, y) = 0 \implies x = y$

$$1 - \frac{1}{1 + d(x, y)} = 0$$

$$1 + d(x, y) = 1$$

$$d(x, y) = 0$$

Now we show $S(x, y) = 0 \impliedby x = y$

$$S(x, x) = 1 - \frac{1}{1 + d(x, x)}$$

$$= 1 - \frac{1}{1}$$

$$= 0$$

Symmetry: $S(x, y) = S(y, x)$
This is trivial, since $d(x, y)$ is symmetric.

Triangle Inequality: $S(x, z) \leq S(x, y) + S(y, z)$

$$1 - \frac{1}{1 + d(x, z)} \leq 1 - \frac{1}{1 + d(x, y)} + 1 - \frac{1}{1 + d(y, z)}$$

$$-\frac{1}{1 + d(x, z)} \leq 1 - \frac{1}{1 + d(x, y)} - \frac{1}{1 + d(y, z)}$$

$$-1 \leq (1 + d(x, z)) - \frac{1 + d(x, z)}{1 + d(x, y)} - \frac{1 + d(x, z)}{1 + d(y, z)}$$

$$-\big(1 + d(x, y)\big)\big(1 + d(y, z)\big) \leq \big(1 + d(x, z)\big)\Big(\big(1 + d(x, y)\big)\big(1 + d(y, z)\big)$$

$$- \big(1 + d(x, y)\big) - \big(1 + d(y, z)\big)\Big)$$

$$-\big(1 + d(x, y)\big)\big(1 + d(y, z)\big) \leq \big(1 + d(x, z)\big)\big(d(x, y)d(y, z) - 1\big)$$

$$-1 - d(x, y)d(y, z) - d(x, y) - d(y, z) \leq d(x, y)d(y, z) - 1 - d(x, z)$$

$$+ d(x, y)d(y, z)d(x, z)$$

$$d(x, z) - d(x, y) - d(y, z) \leq 0 \leq d(x, y)d(y, z)d(x, z)$$

□

C   FULL EMPIRICAL RESULTS

This section contains tables with the complete empirical results. The classification accuracy of Dynamic Time Warping was taken from the results published in the UCR archive [4]. The names of the used datasets and the complete results can be found below. The first two tables list the classification error rate per distance function. The second two table compare the contamination tolerance and the imprecision invariance per distance function.

Dataset	Error rate					
	$\mathcal{E}$	Euc	DTW	Log	ED	EDR
ACSF1	0.28	0.46	0.36	0.17	0.90	0.48
Adiac	0.42	0.39	0.40	0.40	0.97	0.54
ArrowHead	0.21	0.20	0.30	0.20	0.61	0.31
Beef	0.33	0.33	0.37	0.40	0.80	0.50
BeetleFly	0.25	0.25	0.30	0.35	0.50	0.40
BirdChicken	0.45	0.45	0.25	0.35	0.50	0.35
BME	0.25	0.17	0.10	0.20	0.65	0.59
Car	0.27	0.27	0.27	0.28	0.77	0.30
CBF	0.06	0.15	0.00	0.11	0.67	0.38
Chinatown	0.06	0.05	0.04	0.05	0.45	0.03
Coffee	0.04	0.00	0.00	0.07	0.46	0.11
Computers	0.47	0.42	0.30	0.42	0.50	0.42
CricketX	0.37	0.42	0.25	0.37	0.92	0.71
CricketY	0.38	0.43	0.26	0.34	0.92	0.70
CricketZ	0.38	0.41	0.25	0.36	0.93	0.72
DiatomSizeReduction	0.07	0.07	0.03	0.08	0.70	0.08
DistalPhalanxOutlineAgeGroup	0.35	0.37	0.23	0.33	0.58	0.26
DistalPhalanxOutlineCorrect	0.29	0.28	0.28	0.26	0.42	0.30
DistalPhalanxTW	0.37	0.37	0.41	0.37	0.70	0.34
Earthquakes	0.35	0.29	0.28	0.33	0.75	0.75
ECG200	0.12	0.12	0.23	0.11	0.64	0.20
ECGFiveDays	0.18	0.20	0.23	0.21	0.50	0.36
EOGHorizontalSignal	0.60	0.58	0.50	0.67	0.83	0.83
EOGVerticalSignal	0.68	0.56	0.55	0.73	0.86	0.86
EthanolLevel	0.72	0.73	0.72	0.69	0.75	0.69
FaceFour	0.22	0.22	0.17	0.15	0.70	0.26
FiftyWords	0.34	0.37	0.31	0.31	0.97	0.61
Fish	0.22	0.22	0.18	0.23	0.88	0.30
Fungi	0.12	0.18	0.16	0.08	0.96	0.51
GunPoint	0.07	0.09	0.09	0.05	0.51	0.24
GunPointAgeSpan	0.03	0.10	0.08	0.00	0.49	0.17
GunPointMaleVersusFemale	0.01	0.03	0.00	0.01	0.47	0.21
GunPointOldVersusYoung	0.00	0.05	0.16	0.00	0.52	0.01
Ham	0.42	0.40	0.53	0.50	0.51	0.43
HandOutlines	0.14	0.14	0.12	0.14	0.46	0.18
Haptics	0.63	0.63	0.62	0.64	0.79	0.64
Herring	0.47	0.48	0.47	0.41	0.41	0.55
HouseTwenty	0.14	0.34	0.08	0.18	0.32	0.37
InlineSkate	0.66	0.66	0.62	0.64	0.83	0.69
InsectEPGRegularTrain	0.00	0.32	0.13	0.00	0.00	0.00
InsectEPGSmallTrain	0.00	0.34	0.27	0.00	0.00	0.00
LargeKitchenAppliances	0.47	0.51	0.21	0.42	0.67	0.67
Lightning2	0.18	0.25	0.13	0.18	0.46	0.51
Lightning7	0.33	0.42	0.27	0.26	0.74	0.71
Meat	0.07	0.07	0.07	0.07	0.65	0.07

Table 3: Classification error per distance function, Part 1. The ensemble $\mathcal{E}$ is not significantly different from DTW, Log or Euc. ED and EDR frequently have a higher classification error than the remaining functions. Unsurprisingly, DTW has the lowest overall error rate. Its elastic nature likely superior classification accuracy over lockstep distance functions.

Dataset	Error rate					
	$\mathcal{E}$	Euc	DTW	Log	ED	EDR
MedicalImages	0.31	0.32	0.26	0.29	0.49	0.53
MiddlePhalanxOutlineAgeGroup	0.45	0.48	0.50	0.47	0.44	0.49
MiddlePhalanxOutlineCorrect	0.25	0.23	0.30	0.25	0.43	0.24
MiddlePhalanxTW	0.49	0.49	0.49	0.45	0.73	0.47
OliveOil	0.13	0.13	0.17	0.17	0.83	0.60
OSULeaf	0.48	0.48	0.41	0.45	0.90	0.55
PigAirwayPressure	0.88	0.94	0.89	0.90	0.90	0.94
PigArtPressure	0.73	0.88	0.75	0.72	0.93	0.86
PigCVP	0.87	0.92	0.85	0.87	0.91	0.88
Plane	0.04	0.04	0.00	0.04	0.85	0.01
PowerCons	0.03	0.07	0.12	0.04	0.20	0.32
ProximalPhalanxOutlineAgeGroup	0.23	0.21	0.20	0.22	0.57	0.22
ProximalPhalanxOutlineCorrect	0.24	0.19	0.22	0.22	0.32	0.24
ProximalPhalanxTW	0.28	0.29	0.24	0.30	0.98	0.27
RefrigerationDevices	0.57	0.61	0.54	0.52	0.67	0.69
Rock	0.44	0.16	0.40	0.34	0.62	0.46
ScreenType	0.65	0.64	0.60	0.62	0.67	0.72
SemgHandGenderCh2	0.13	0.24	0.20	0.22	0.35	0.36
SemgHandMovementCh2	0.24	0.63	0.42	0.55	0.83	0.82
SemgHandSubjectCh2	0.16	0.60	0.27	0.42	0.80	0.77
ShapeletSim	0.52	0.46	0.35	0.49	0.50	0.50
ShapesAll	0.24	0.25	0.23	0.24	0.98	0.37
SmallKitchenAppliances	0.46	0.66	0.36	0.51	0.66	0.74
SmoothSubspace	0.02	0.09	0.17	0.00	0.67	0.15
SonyAIBORobotSurface1	0.24	0.30	0.27	0.31	0.57	0.37
SonyAIBORobotSurface2	0.13	0.14	0.17	0.12	0.38	0.21
Strawberry	0.06	0.05	0.06	0.05	0.36	0.05
SwedishLeaf	0.22	0.21	0.21	0.22	0.93	0.32
Symbols	0.10	0.10	0.05	0.10	0.84	0.20
SyntheticControl	0.07	0.12	0.01	0.13	0.83	0.35
ToeSegmentation1	0.29	0.32	0.23	0.27	0.46	0.36
ToeSegmentation2	0.15	0.19	0.16	0.12	0.18	0.21
Trace	0.31	0.24	0.00	0.21	0.76	0.32
UMD	0.25	0.24	0.01	0.24	0.48	0.51
Wine	0.33	0.39	0.43	0.35	0.50	0.48
WordSynonyms	0.38	0.38	0.35	0.34	0.91	0.63
Worms	0.48	0.55	0.42	0.56	0.44	0.64
WormsTwoClass	0.36	0.39	0.38	0.42	0.44	0.43

Table 4: Classification error per distance function, Part 2.

Dataset	Contamination Tolerance						Imprecision Invariance					
	$\mathcal{E}$	Euc	DTW	Log	ED	EDR	$\mathcal{E}$	Euc	DTW	Log	ED	EDR
ACSF1	0.68	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.98
Adiac	0.49	0.00	0.00	0.00	1.00	0.94	1.00	1.00	1.00	1.00	0.00	0.98
ArrowHead	0.98	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Beef	0.85	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
BeetleFly	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
BirdChicken	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
BME	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Car	0.99	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
CBF	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Chinatown	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Coffee	0.30	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Computers	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	0.99
CricketX	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
CricketY	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
CricketZ	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
DiatomSizeReduction	0.95	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
DistalPhalanxOutlineAgeGroup	0.95	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
DistalPhalanxOutlineCorrect	0.66	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
DistalPhalanxTW	0.94	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Earthquakes	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
ECG200	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
ECGFiveDays	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
EOGHorizontalSignal	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.16	1.00
EOGVerticalSignal	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.15	1.00
EthanolLevel	0.85	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
FaceFour	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
FiftyWords	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Fish	0.97	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Fungi	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
GunPoint	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
GunPointAgeSpan	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.05	1.00
GunPointMaleVersusFemale	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.07	1.00
GunPointOldVersusYoung	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.04	1.00
Ham	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
HandOutlines	0.57	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Haptics	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Herring	0.62	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
HouseTwenty	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.18	1.00
InlineSkate	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
InsectEPGRegularTrain	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
InsectEPGSmallTrain	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
LargeKitchenAppliances	0.96	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Lightning2	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Lightning7	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Meat	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00

Table 5: Contamination tolerance and imprecision invariance per distance function, Part 1. The numbers indicate the percentage of non-class members which have a larger distance to the considered in-class time series, averaged over all classes. ED perfectly tolerates contamination, while $\mathcal{E}$ and EDR commonly, but not always achieve perfect scores. In terms of imprecision invariance, all measures besides ED appear to fulfill this property. Altogether, EDR with a median absolute deviation-based tolerance interval appears to have the highest combined robustness.

Dataset	Contamination Tolerance						Imprecision Invariance					
	$\mathcal{E}$	Euc	DTW	Log	ED	EDR	$\mathcal{E}$	Euc	DTW	Log	ED	EDR
MedicalImages	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
MiddlePhalanxOutlineAgeGroup	0.69	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
MiddlePhalanxOutlineCorrect	0.54	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
MiddlePhalanxTW	0.84	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
OliveOil	0.00	0.00	0.00	0.00	1.00	0.01	1.00	1.00	1.00	1.00	0.00	0.51
OSULeaf	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
PigAirwayPressure	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
PigArtPressure	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.09	1.00
PigCVP	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Plane	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
PowerCons	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
ProximalPhalanxOutlineAgeGroup	0.88	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
ProximalPhalanxOutlineCorrect	0.52	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
ProximalPhalanxTW	0.71	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
RefrigerationDevices	0.97	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Rock	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.07	1.00
ScreenType	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
SemgHandGenderCh2	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.02	1.00
SemgHandMovementCh2	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.02	1.00
SemgHandSubjectCh2	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.02	1.00
ShapeletSim	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
ShapesAll	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
SmallKitchenAppliances	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
SmoothSubspace	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
SonyAIBORobotSurface1	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
SonyAIBORobotSurface2	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Strawberry	0.50	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
SwedishLeaf	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Symbols	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
SyntheticControl	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
ToeSegmentation1	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
ToeSegmentation2	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Trace	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
UMD	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Wine	0.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
WordSynonyms	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
Worms	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00
WormsTwoClass	1.00	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00	1.00

Table 6: Contamination tolerance and imprecision invariance per distance function, Part 2.