Black sheep
Detecting vehicles on the road that should not be there

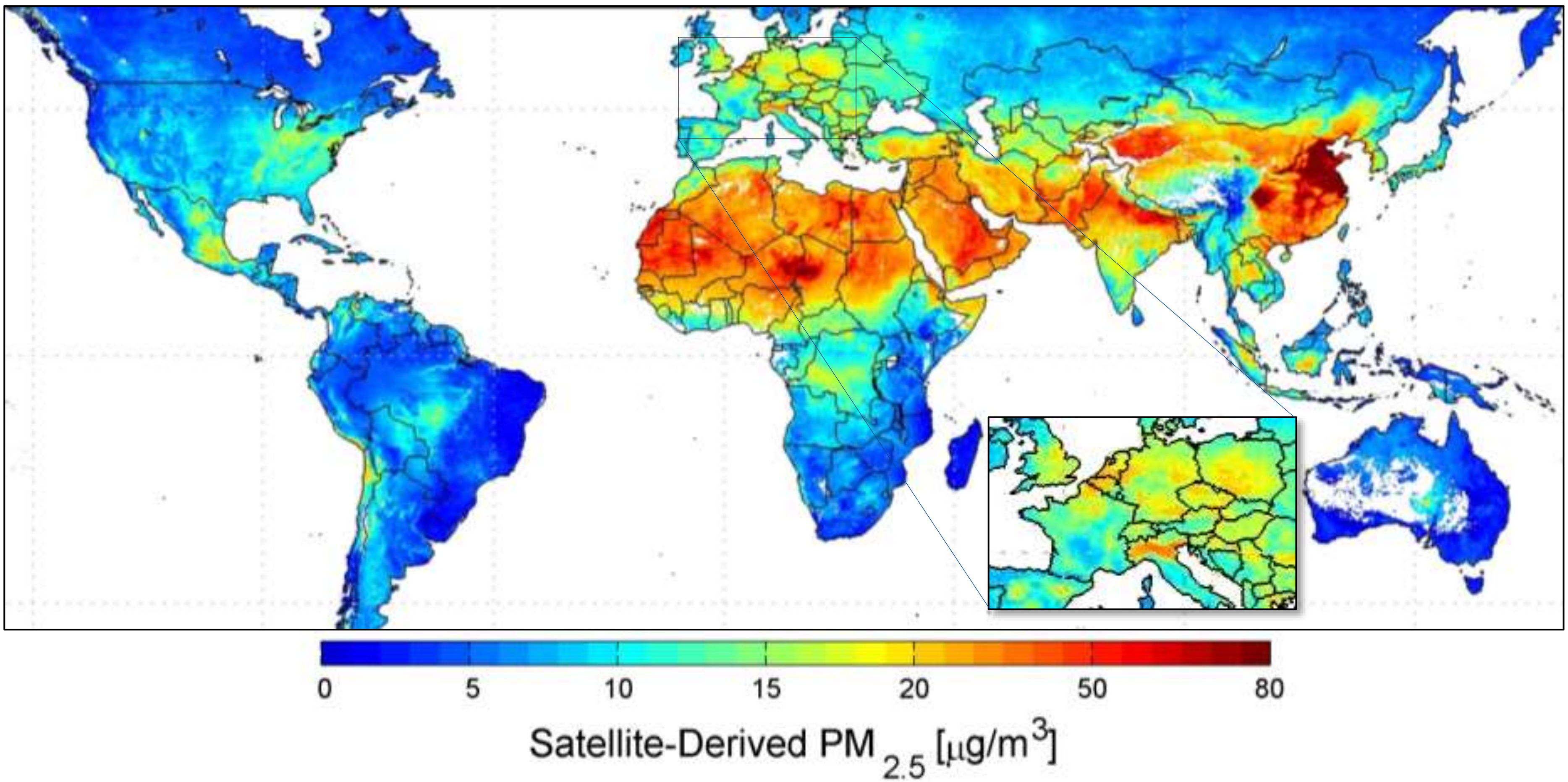
Jan Skácel
jan.skacel@fs.cvut.cz

December 2017

Outline

- Vehicle detection - Project Motivation
- Time series introduction
  - Dynamic time warping
  - Segmentation
  - Semi-supervised learning
- Project scope
- Measurement approach, data sets
- Data processing

Global pollution - Particulate matter
Satellite images of $PM_{2.5}$ averaged over 2001-2006



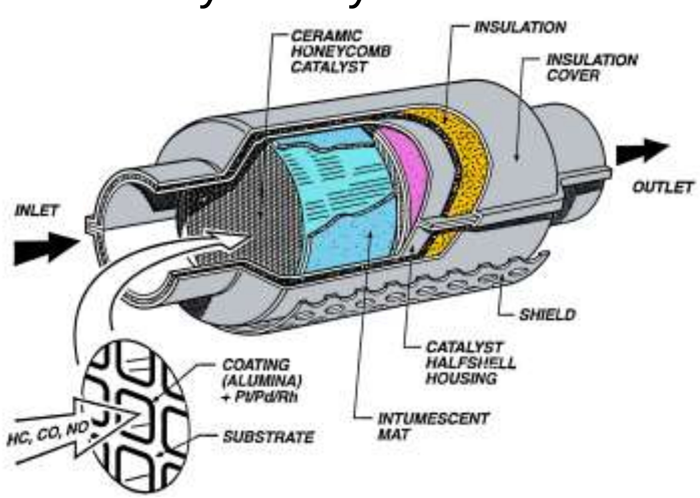
Global satellite-derived map of $PM_{2.5}$ averaged over 2001-2006.
Credit: Dalhousie University, Aaron van Donkelaar

Vehicle Emission Regulation
4 Key Pollutants, 3 Key Components

Key Pollutants

CO – carbon monoxide
NOx – nitric oxides
HC - hydrocarbons (unburnt fuel)
PM - particulate matter (soot)
PM – mass, PN - number

Three Way Catalyst



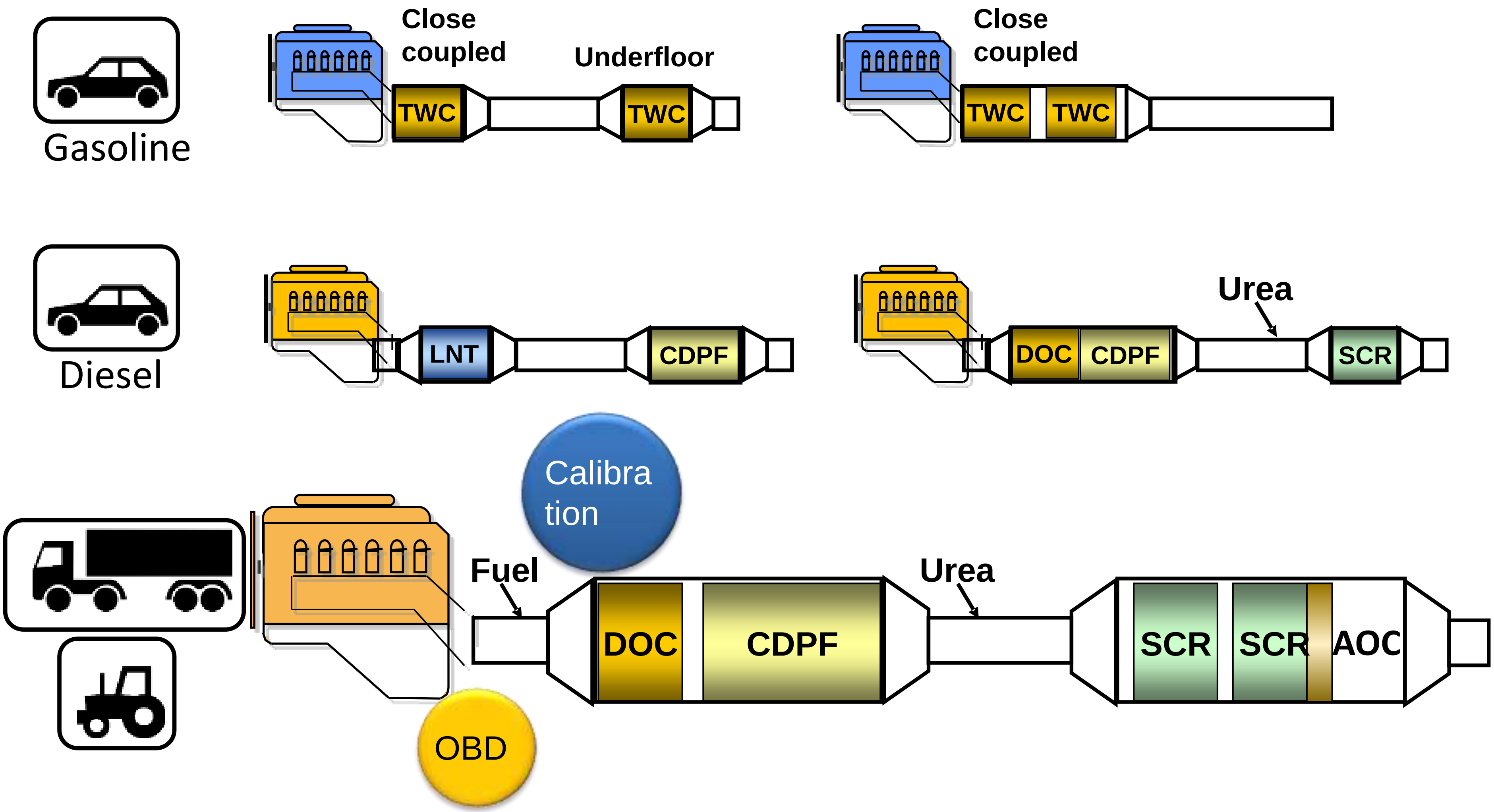
Regulation has 3 components

Emission Limits
- SULEV/PZEV
- EURO 6c
- EURO VI
- ARB 0.02g NOx
- post T4F
- N_2O/GHG

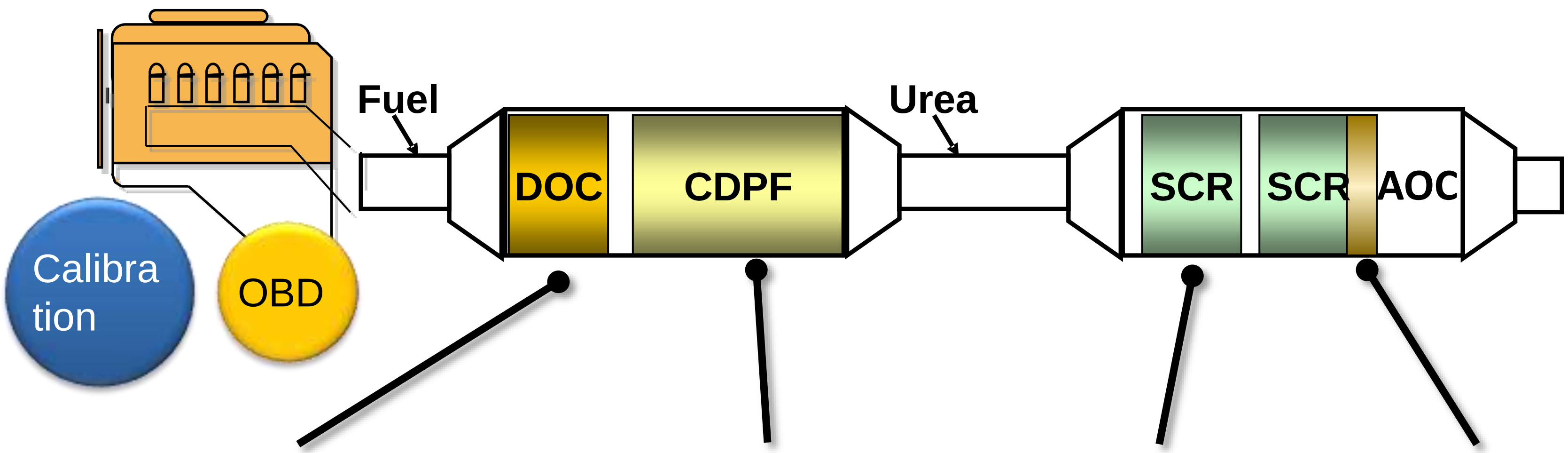
Drive Cycle
- NEDC
- WHTC
- FTP
- Vehicle/Engine
- Random/PEMS

Durability
- Age
- Mileage
- Deterioration factor

Engine Aftertreatment Systems
Today's Typical Layout

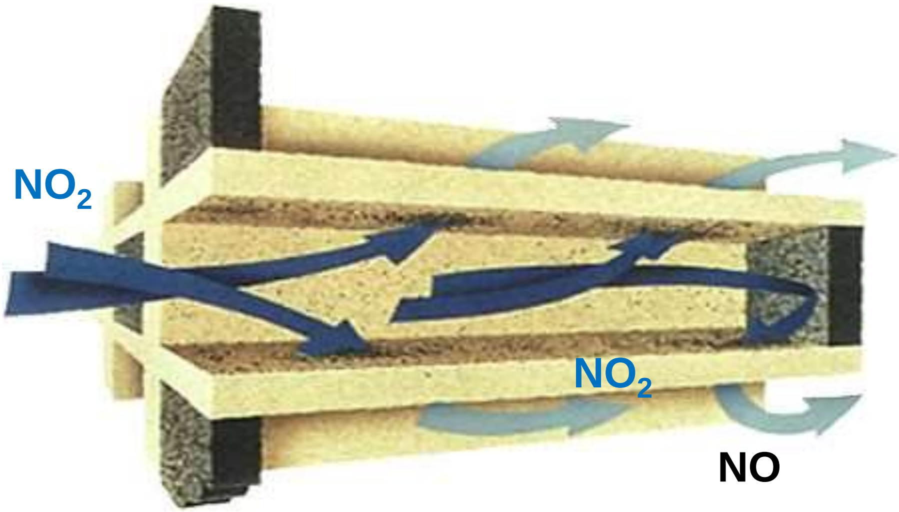


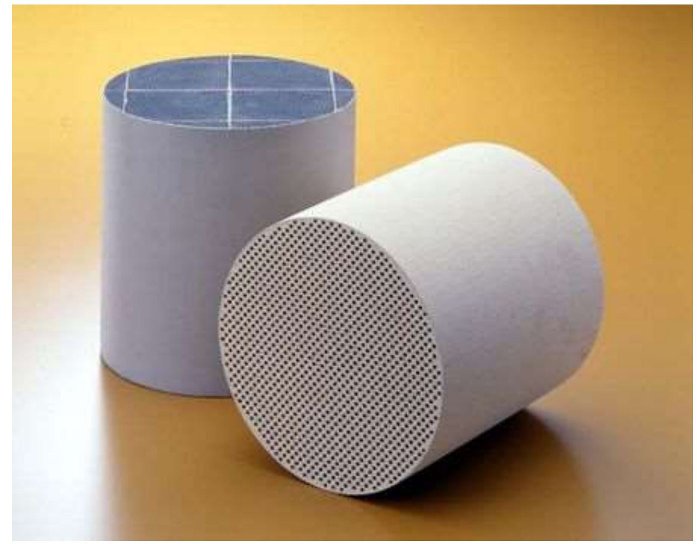
Diesel Engine Aftertreatment System
“Chemical Factory on wheels”



	Diesel Oxidation Catalyst	Catalyzed Diesel Particulate Filter	Selective Catalytic Reduction	Ammonia Oxidation Catalyst
Key Function:	• HC and CO removal • NO oxidation • Exothermic fuel light-off	• Diesel Particulate Matter Filtration • NO oxidation • HC and CO removal • Passive soot regeneration (NO_2) • Active soot regeneration (O_2)	• NOx abatement • Minimal N_2O make	• NH_3 abatement

Catalyzed Diesel Particulate Filter



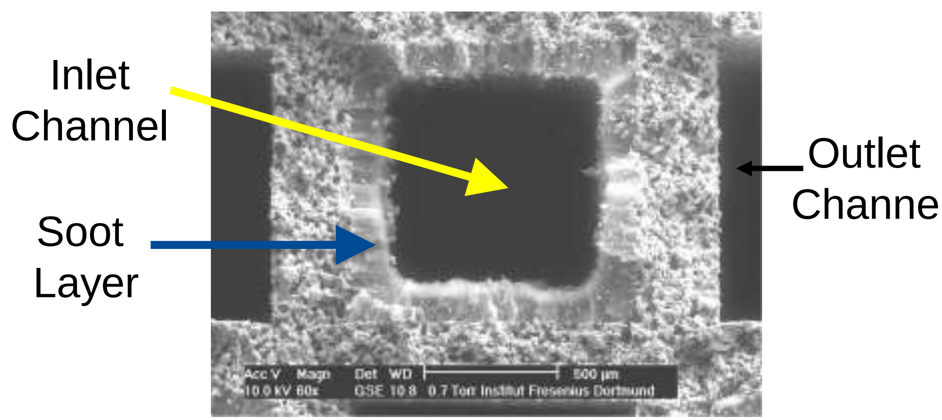


Left: silicon carbide; Right: cordierite

Passive Regeneration:

$$NO + O_2 \xrightarrow{Pt} NO_2$$

$$NO_2 + C\ (soot) \rightarrow NO + CO/CO_2$$



Active Regeneration:

$$O_2 + C\ (soot) \rightarrow CO/CO_2$$

Despite Tighter Vehicle Emission Regulation Air Quality is Not Always Improving

Problem:

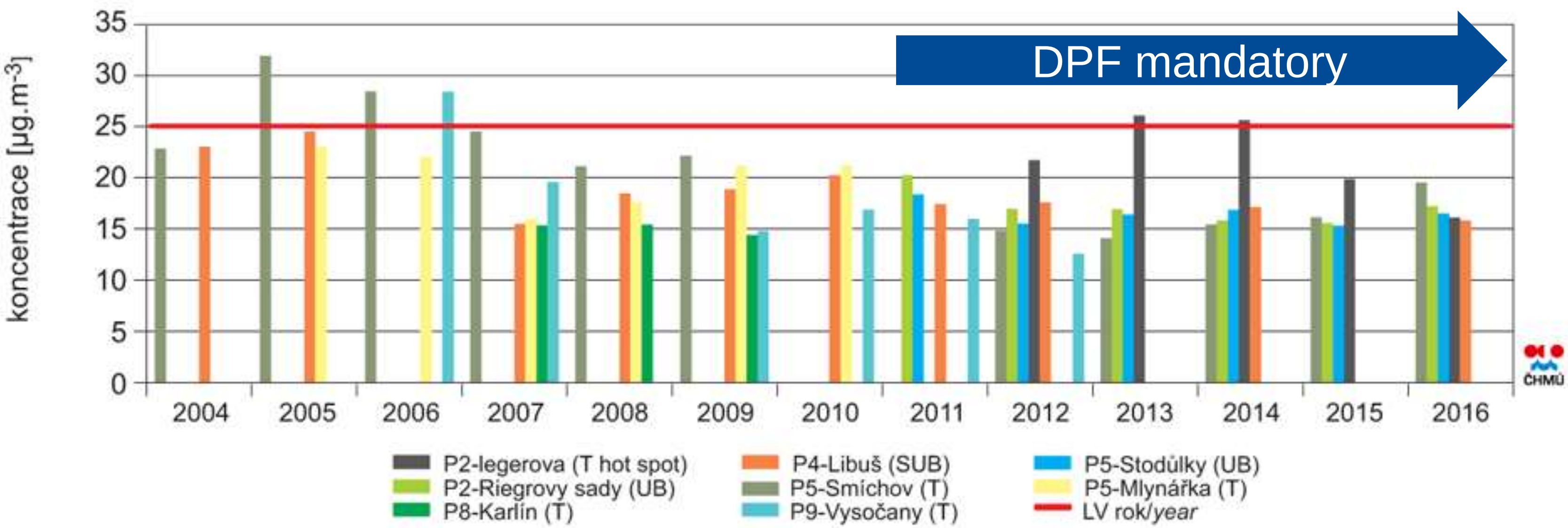
- Since 2011 all new diesel vehicles in EU must have diesel particulate filter
- Most harmful dust particles PM2,5 levels not improving

Hypothesis:

- Small number of non-compliant and/or older vehicles outweigh all clean ones

Verification:

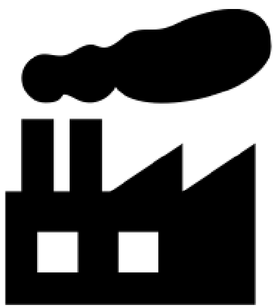
- Roadside emission measurement – real life data from affected locations
  – matching individual vehicle with its emission levels and technical data



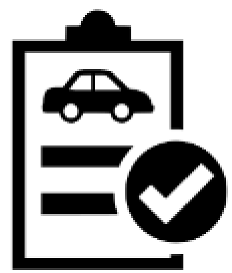
Obr. V.1.3 Průměrné roční koncentrace $PM_{2,5}$, aglomerace Praha, 2004–2016

There is Technology, Solution,
Issue – Regulation, Monitoring & Enforcement

Manufacturer
“Nominal”



In use inspection
“Durability”



On the road
Real world

On road emissions
the only thing that
really matters

Issues:

Meeting limits under
- laboratory conditions
- laboratory cycle

Unrealistic loading
Infrequent
“clients” - no motivation
Staged compliance

Real life on road
emissions can be
much higher

Roadside Particulate Matter Emission Measurement – Prague October 2017

Radar Speed detection

3 Cameras

5 Emission measurement instruments

Sampling line

Quick facts:

- 10 days of measurement

~ 12 500 vehicles

- Particulate data - count, size & mass -

- Gaseous emissions CO_2 CO, NO_x

Roadside Measurement
Particulate Matter Emissions – Time Series Task

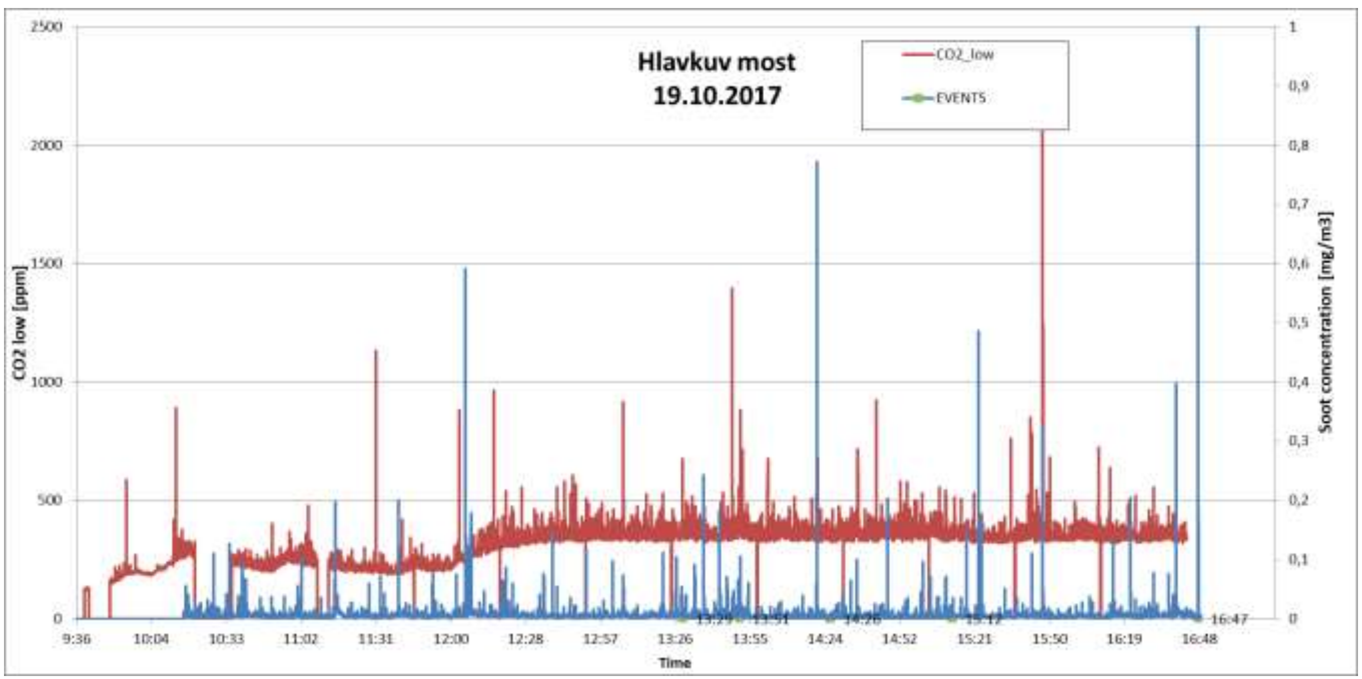
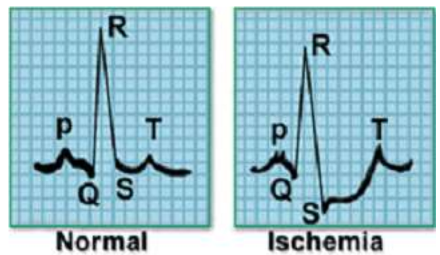


Plate No.	Capture Time	Country	Validity
2SL1722	19.10.2017 09:47:42	Czech Republic(CZE)	87
2Z68863	19.10.2017 09:47:46	Czech Republic(CZE)	78
4E78140	19.10.2017 09:47:56	Czech Republic(CZE)	84
3AY2703	19.10.2017 09:47:58	Czech Republic(CZE)	85
7A84858	19.10.2017 09:47:59	Czech Republic(CZE)	85
2SP4295	19.10.2017 09:48:02	Czech Republic(CZE)	82
5AS3754	19.10.2017 09:48:04	Czech Republic(CZE)	82
1BL8155	19.10.2017 09:48:05	Czech Republic(CZE)	82
4SC4081	19.10.2017 09:48:11	Czech Republic(CZE)	82
5AB3109	19.10.2017 09:48:38	Czech Republic(CZE)	78
3SC3641	19.10.2017 09:48:40	Czech Republic(CZE)	83
2AS0930	19.10.2017 09:48:41	Czech Republic(CZE)	78
9A81854	19.10.2017 09:48:43	Czech Republic(CZE)	86
3AE6666	19.10.2017 09:49:00	Czech Republic(CZE)	82
8U16611	19.10.2017 09:49:07	Czech Republic(CZE)	79
5AT2236	19.10.2017 09:49:08	Czech Republic(CZE)	86
1SH3899	19.10.2017 09:49:11	Czech Republic(CZE)	83
5AL9012	19.10.2017 09:49:23	Czech Republic(CZE)	82
4AM0962	19.10.2017 09:49:26	Czech Republic(CZE)	80
1AT2062	19.10.2017 09:49:31	Czech Republic(CZE)	71
6U97142	19.10.2017 09:49:34	Czech Republic(CZE)	84
5AZ7159	19.10.2017 09:49:41	Czech Republic(CZE)	87
1SL1820	19.10.2017 09:49:45	Czech Republic(CZE)	81
4AF1700	19.10.2017 09:49:48	Czech Republic(CZE)	83
5AV6231	19.10.2017 09:49:49	Czech Republic(CZE)	85
4AX3458	19.10.2017 09:49:53	Czech Republic(CZE)	84
6AL8823	19.10.2017 09:49:58	Czech Republic(CZE)	79
3SI1106	19.10.2017 09:50:06	Czech Republic(CZE)	81
5AH8016	19.10.2017 09:50:15	Czech Republic(CZE)	80
4AV6937	19.10.2017 09:50:19	Czech Republic(CZE)	85
5A97015	19.10.2017 09:50:22	Czech Republic(CZE)	82
4L94455	19.10.2017 09:50:23	Czech Republic(CZE)	87
9T81	19.10.2017 09:50:26	Europe Region(NON)	78
5AN0880	19.10.2017 09:50:29	Czech Republic(CZE)	81
4AX0304	19.10.2017 09:51:06	Czech Republic(CZE)	84
2AL5261	19.10.2017 09:51:18	Czech Republic(CZE)	84
5AD9034	19.10.2017 09:51:47	Czech Republic(CZE)	82
3AE9666	19.10.2017 09:51:50	Czech Republic(CZE)	81
9A58720	19.10.2017 09:51:55	Czech Republic(CZE)	83
9A81042	19.10.2017 09:51:57	Czech Republic(CZE)	85
5H56777	19.10.2017 09:52:20	Czech Republic(CZE)	82
5M20388	19.10.2017 09:52:26	Czech Republic(CZE)	83
5AE6680	19.10.2017 09:52:36	Czech Republic(CZE)	83
2SK9114	19.10.2017 09:52:37	Czech Republic(CZE)	80
6AK4443	19.10.2017 09:52:39	Czech Republic(CZE)	86
4AR6740	19.10.2017 09:52:41	Czech Republic(CZE)	81
5AX7696	19.10.2017 09:52:42	Czech Republic(CZE)	85
MVJ832	19.10.2017 09:52:47	Hungary(HUN)	85
99H9999	19.10.2017 09:52:51	Europe Region(NON)	82

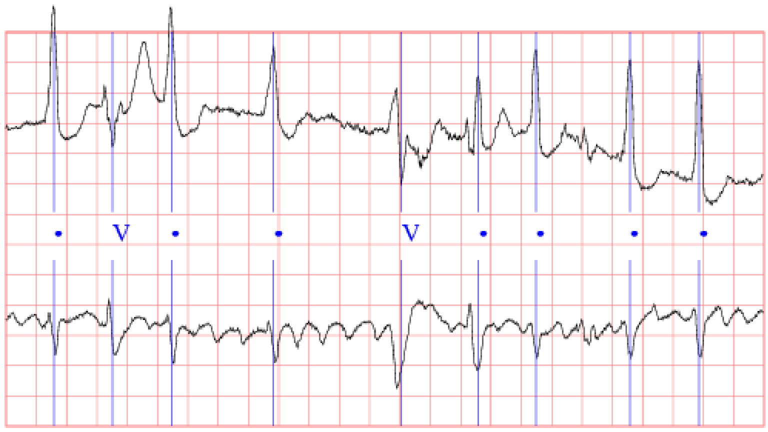
The Ubiquity of Time Series

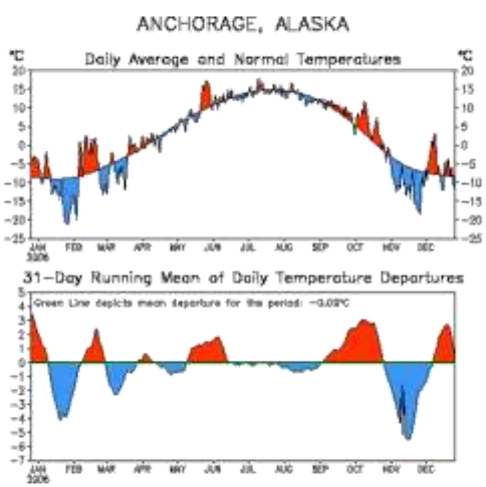
Stock

Medicinal Applications

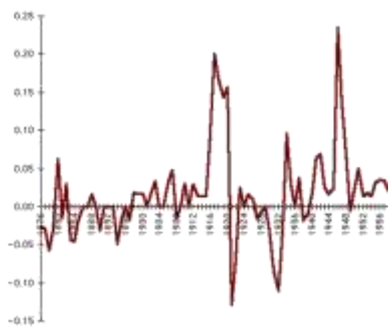


ECG Heartbeat

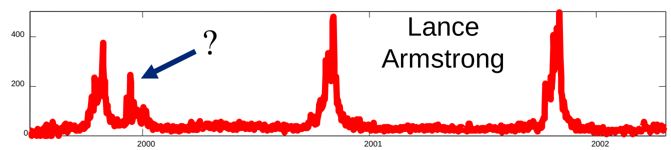




Industrial Measurement



Web search history

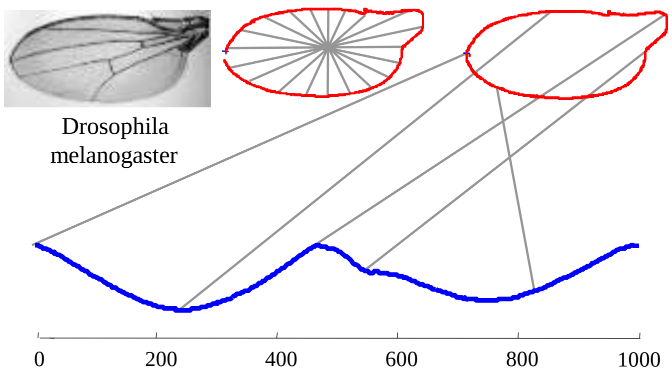


Music

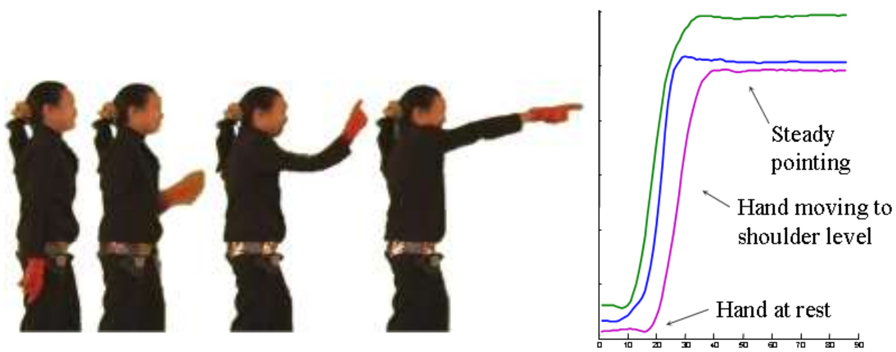
Eamonn Keogh, eamonn@cs.ucr.edu

Time Series in Less Intuitive Domains

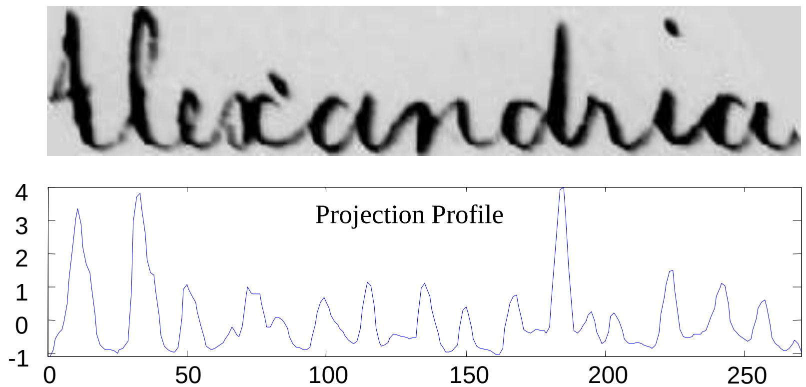
Image



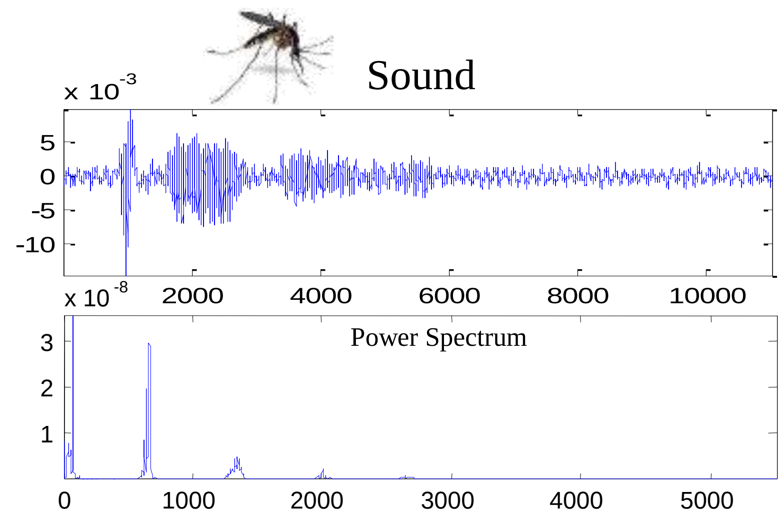
Motion Capture



Text

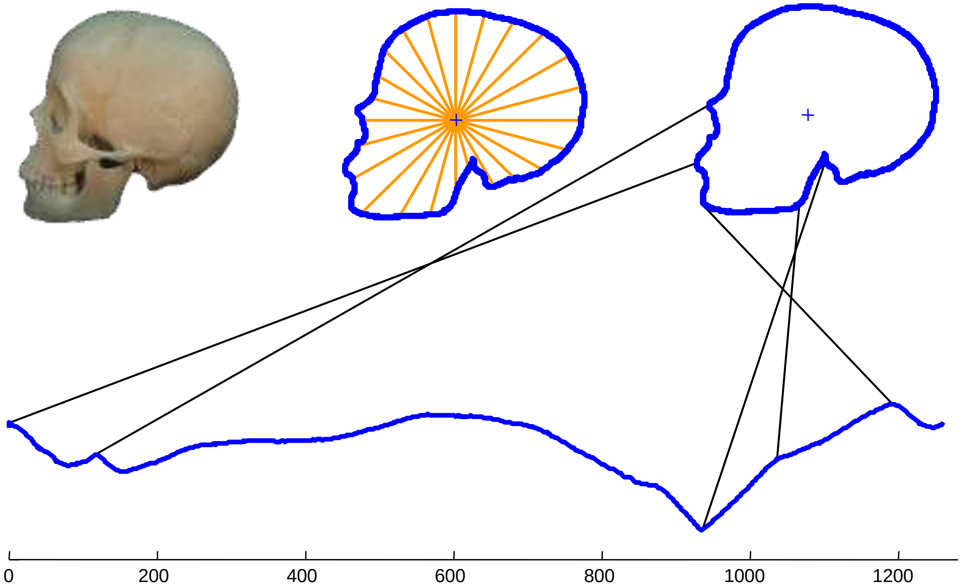


Sound



Eamonn Keogh, eamonn@cs.ucr.edu

Shape Representations



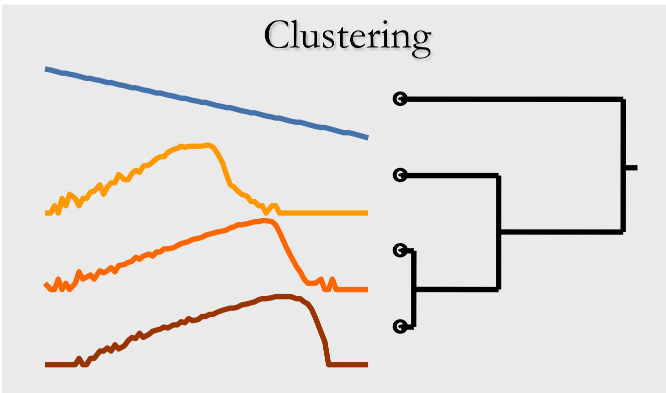
We can convert shapes into a 1D signal. By doing this we remove information about *scale* and *offset*. But we must deal with *rotation* in our algorithms …

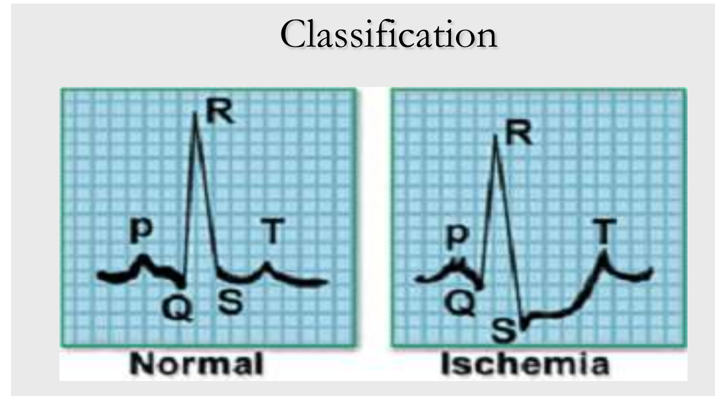
There are three ways to be rotation invariant:

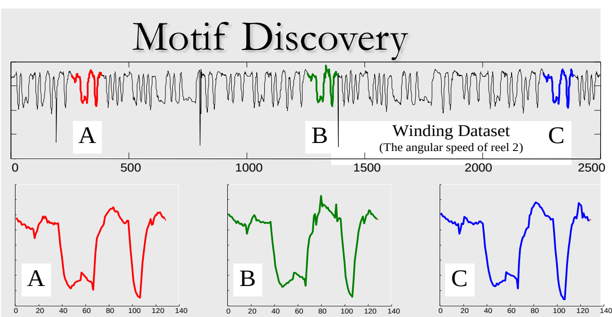
Landmarking, Rotation Invariant Features, Brute Force Rotation Alignment

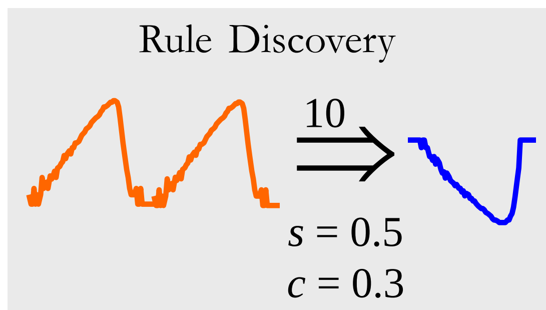
Eamonn Keogh, eamonn@cs.ucr.edu

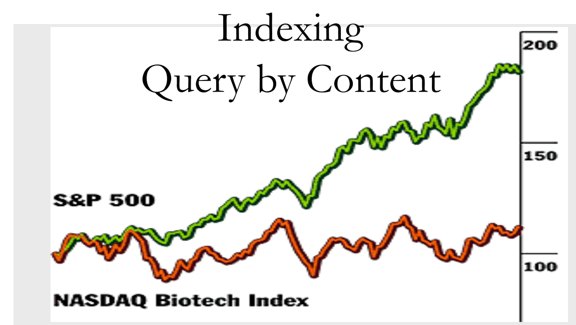
Time Series Data Mining Tasks

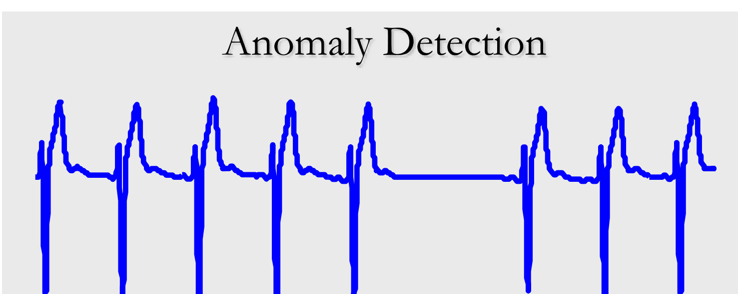


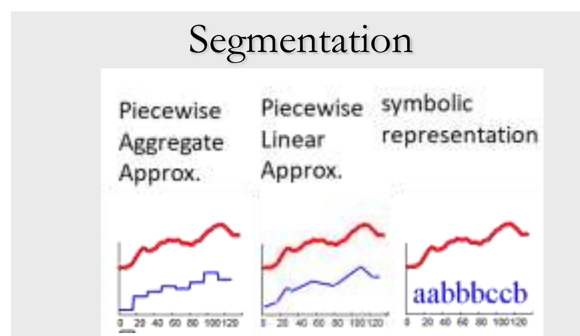












Time Series Data Mining Tasks

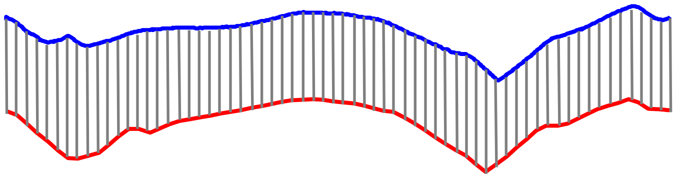
Indexing (Query by Content): Given a query time series Q, and some similarity/dissimilarity measure D(Q,C), find the nearest matching time series in database DB.

- **Clustering:** Find natural groupings of the time series in database DB under some similarity/dissimilarity measure D(Q,C).
- **Classification:** Given an unlabeled time series Q, assign it to one of two or more predefined classes.
- **Segmentation:** Given a time series Q containing n datapoints, construct a model Q, from K piecewise segments (K << n) such that Q closely approximates Q.

Shape Distance Measures

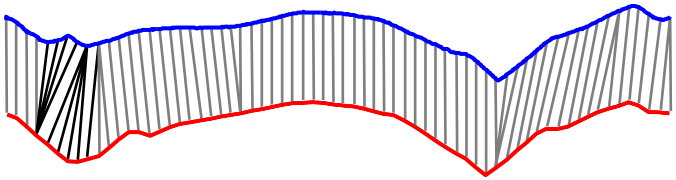
Euclidean Distance:

Does "one to one" matching.



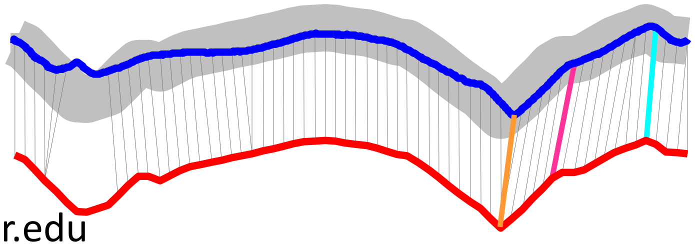
Dynamic Time Warping:

Allows elastic shifting of the time axis.



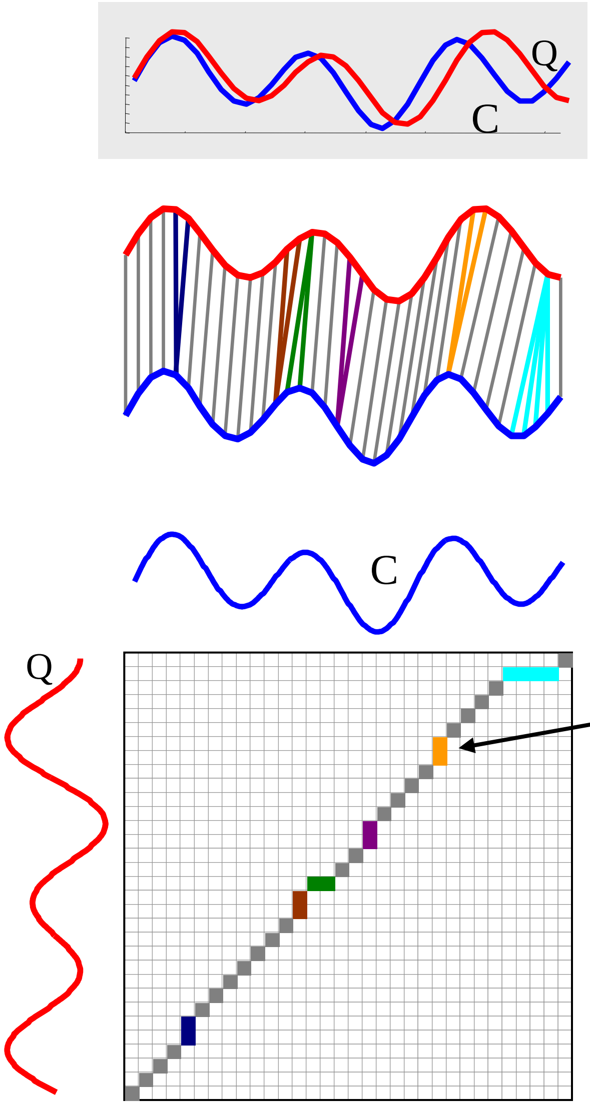
Longest Common Subsequence:

Omits missing parts.



Eamonn Keogh, eamonn@cs.ucr.edu

Dynamic Time Warping



Warping path w

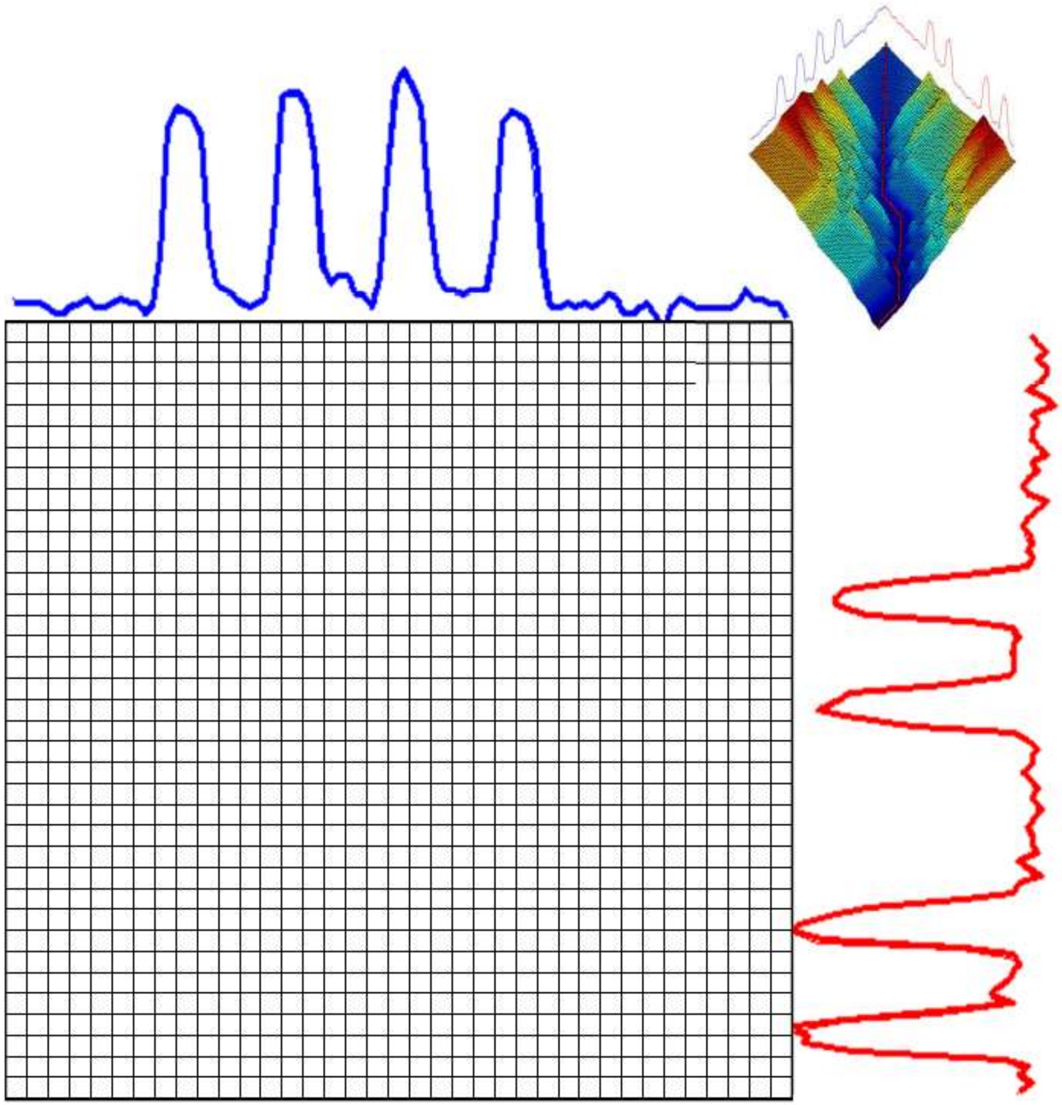
$$DTW(Q,C) = \min\left\{ \sqrt{\sum_{k=1}^{K} w_k} \middle/ K \right.$$

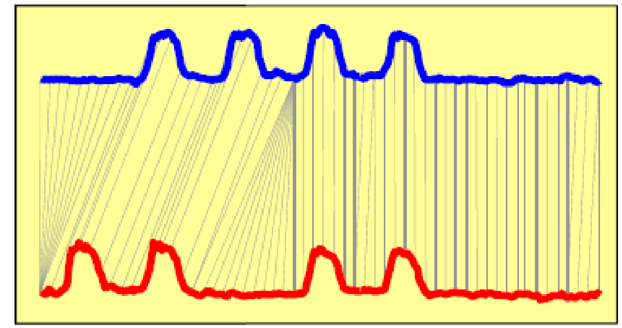
This recursive function gives us the minimum cost path

$$\gamma(i,j) = d(q_i,c_j) + \min\{ \gamma(i\text{-}1,j\text{-}1), \gamma(i\text{-}1,j), \gamma(i,j\text{-}1) \}$$

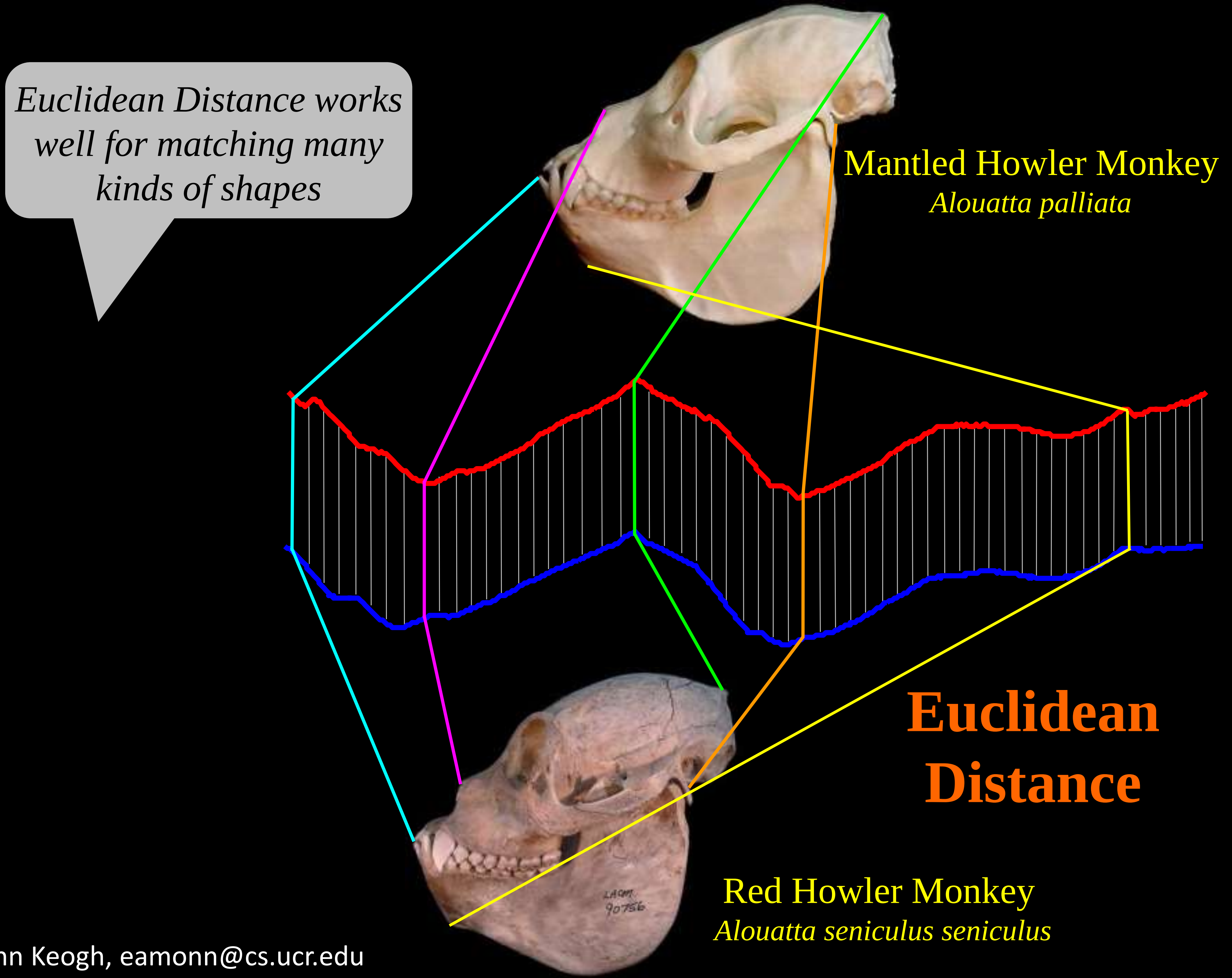
Eamonn Keogh, eamonn@cs.ucr.edu

Example: visualize the process on a real world problem



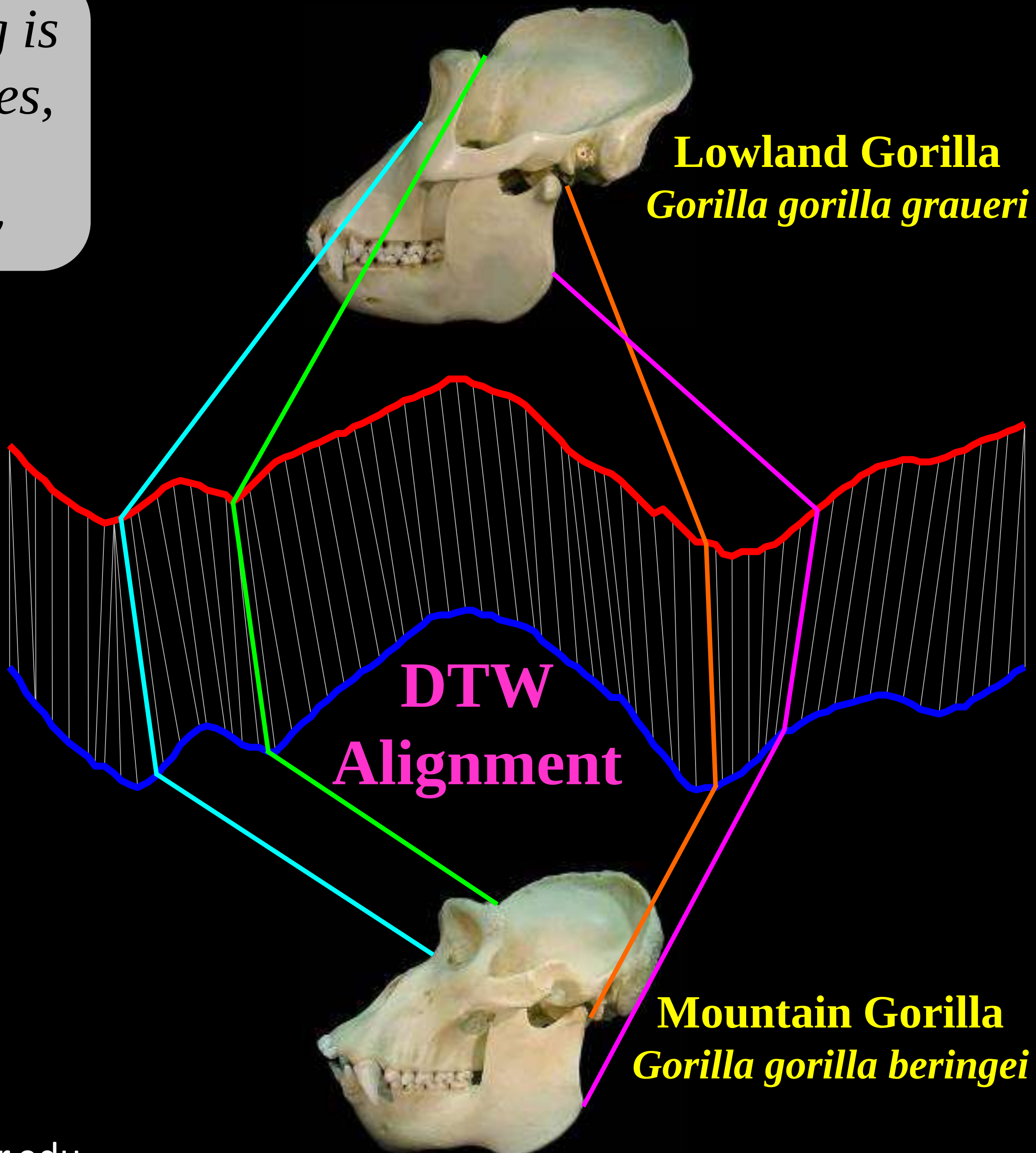


Weekly power demand
Monday to Sunday
4 day work weeks
Blue – Monday is holiday
Red – Wednesday is holiday

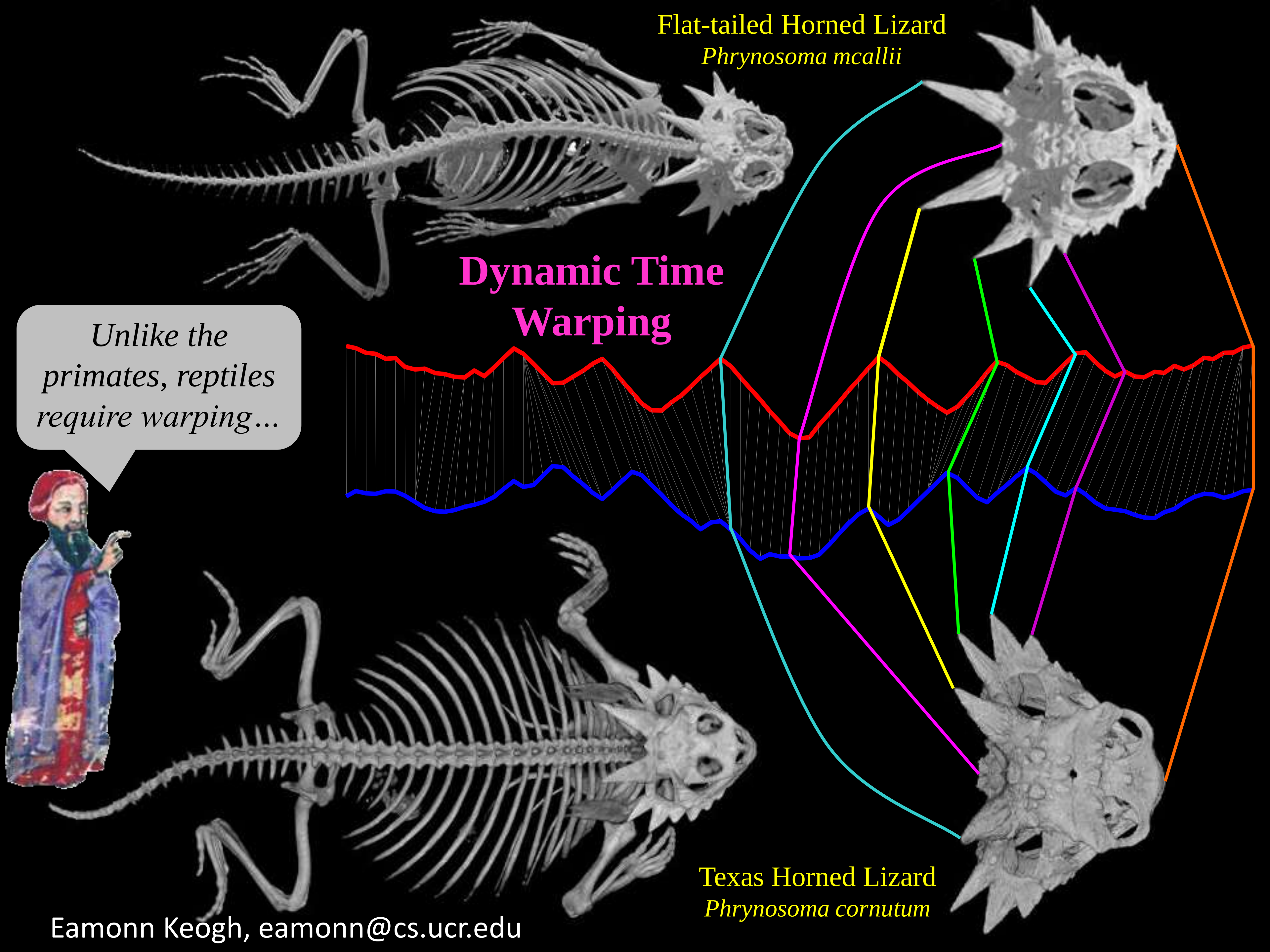


Eamonn Keogh, eamonn@cs.ucr.edu

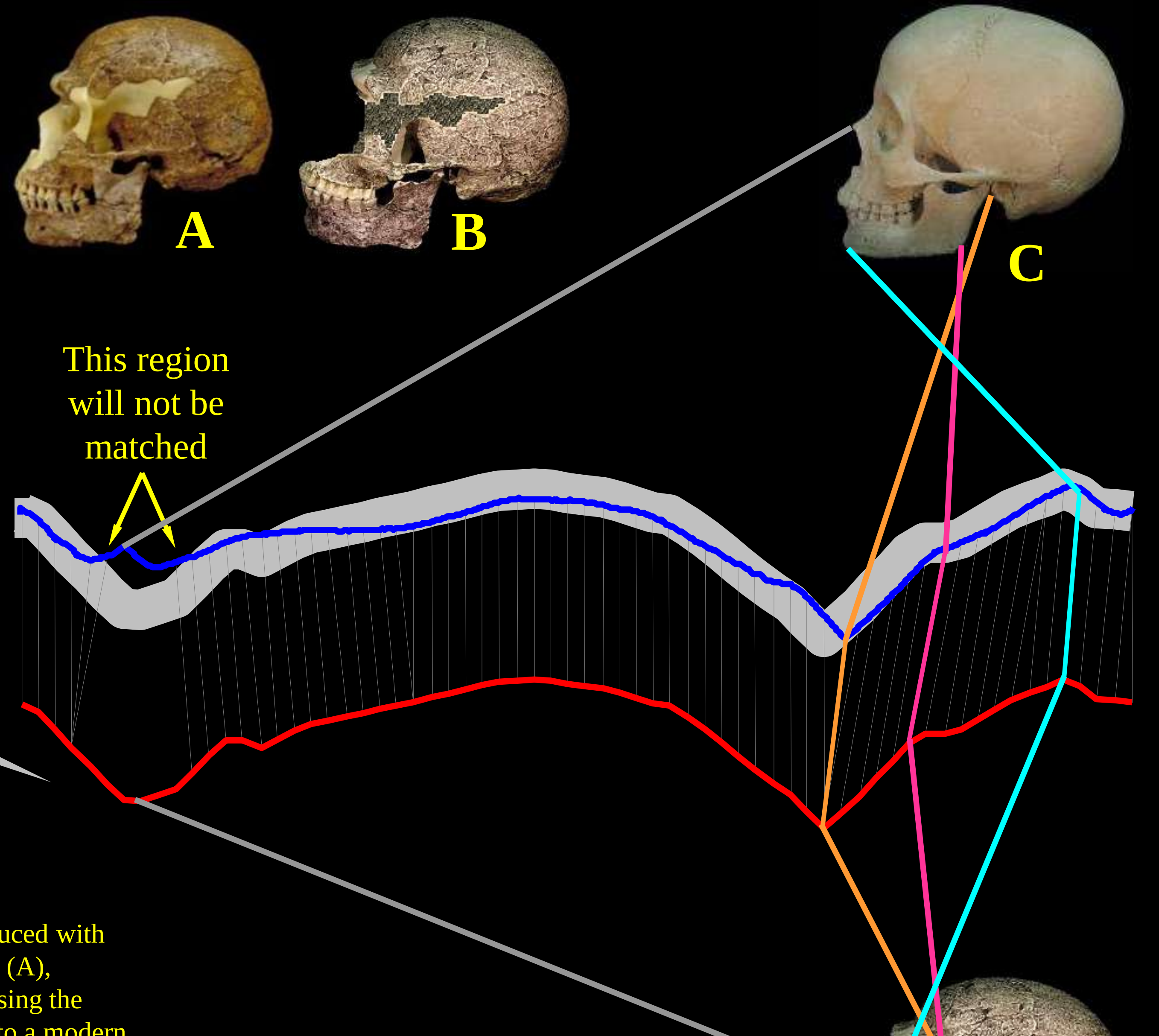
Dynamic Time Warping is useful for natural shapes, which often exhibit intraclass variability



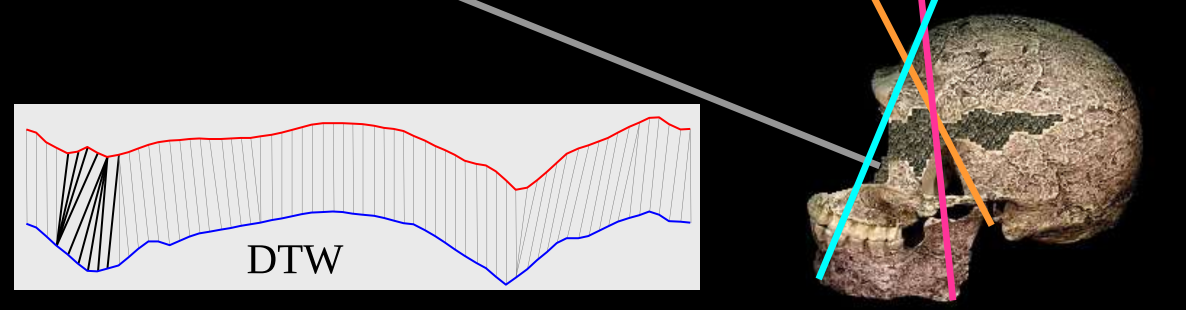
Eamonn Keogh, eamonn@cs.ucr.edu


Flat-tailed Horned Lizard
Phrynosoma mcallii
Dynamic Time Warping
Unlike the primates, reptiles require warping…
Texas Horned Lizard
Phrynosoma cornutum
Eamonn Keogh, eamonn@cs.ucr.edu

Longest Common Subsequence



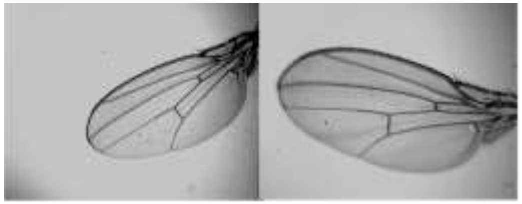
The famous Skhul V is generally reproduced with the missing bones extrapolated in epoxy (A), however the original Skhul V (**B**) is missing the nose region, which means it will match to a modern human (**C**) poorly, even after DTW alignment (inset). In contrast, LCSS alignment will not attempt to match features that are outside a "matching envelope" (heavy gray line) created from the other sequence.



The Ubiquity of Shape Matching

Skulls

Fruit fly wings

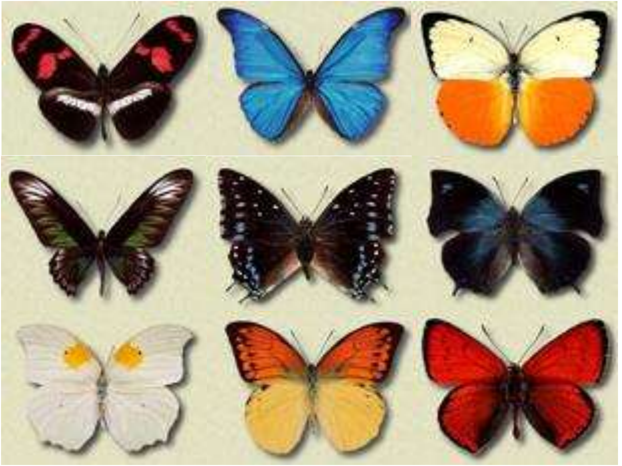


Sea animals

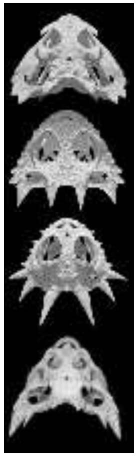
Leaves

Petroglyphs

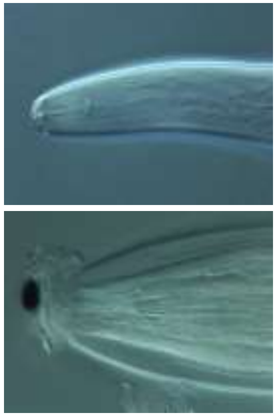
Butterflies



Lizards



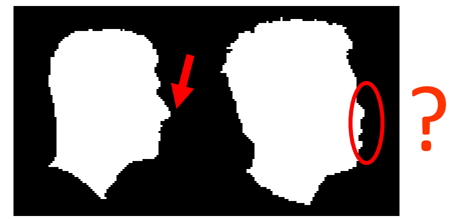
Nematodes



Arrowheads

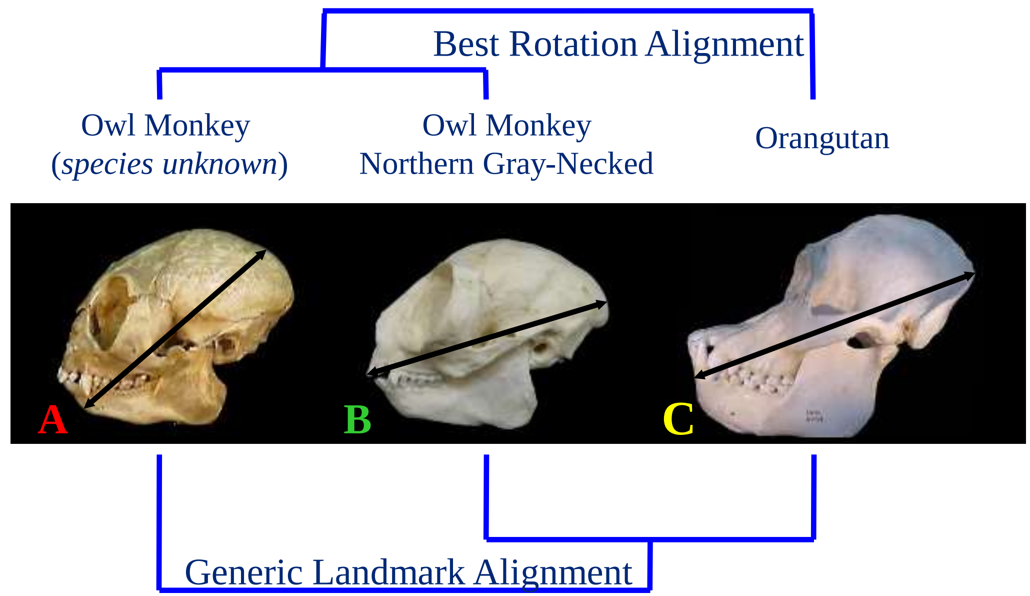
Landmarking*: Find the "True" Rotation

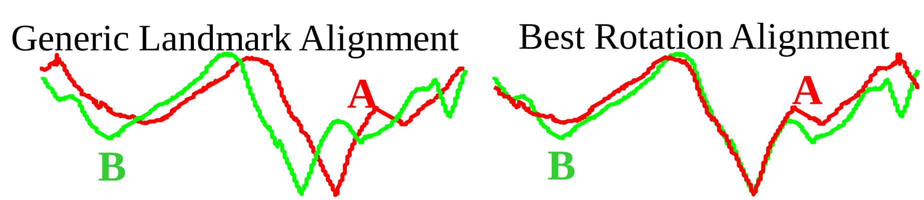
- **Domain Specific Landmarking**
Find some fixed point in your domain, eg. the nose on a face, the stem of leaf, the tail of a fish …



- **Generic Landmarking**
Find the major axis of the shape and use that as the canonical alignment

- **Problem**
It does not work in many cases.

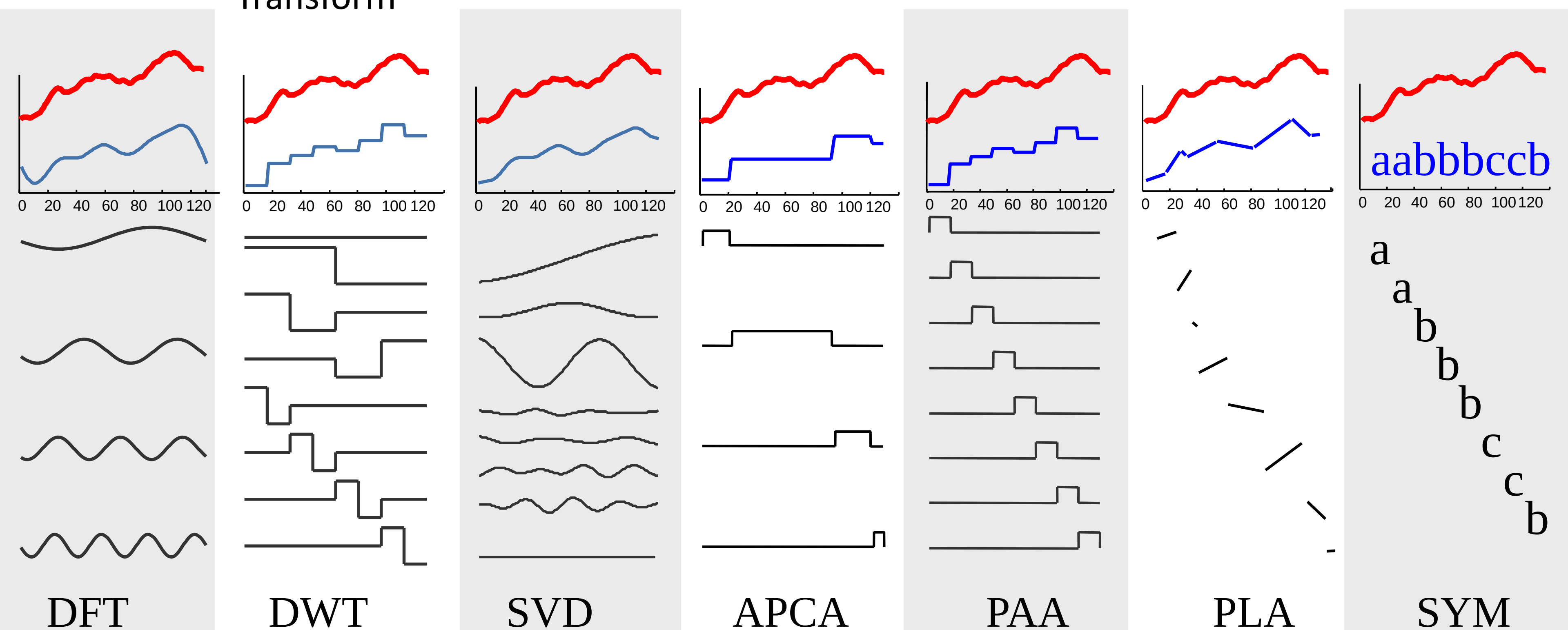




* Xie, J. AND Heng P. Shape Modeling Using Automatic Landmarking. MICCAI 2005.

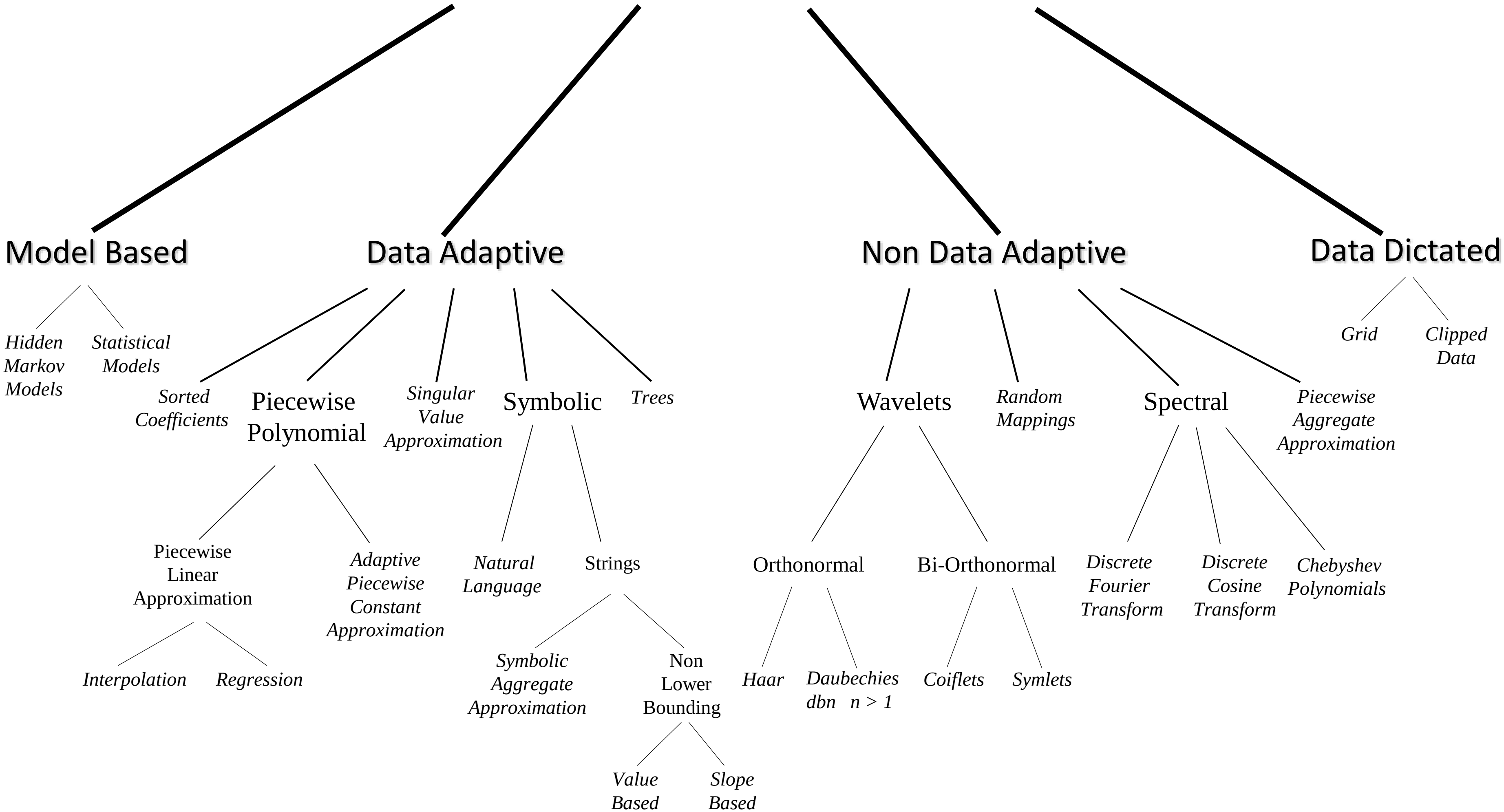
Time Series approximations

Discrete Fourier Transform | Discrete Wavelet Transform | Singular Value Decomp. | Adaptive Piecewise Constant Approx. | Piecewise Aggregate Approx. | Piecewise Linear Approx. | symbolic representation



Eamonn Keogh, eamonn@cs.ucr.edu

Time Series Representations

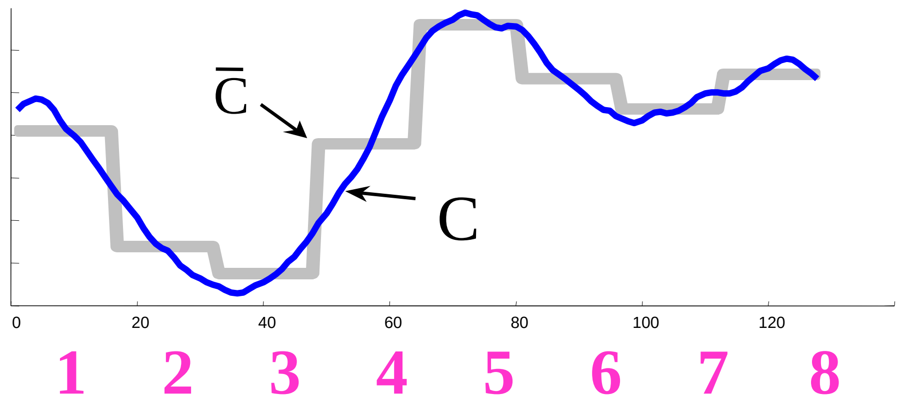


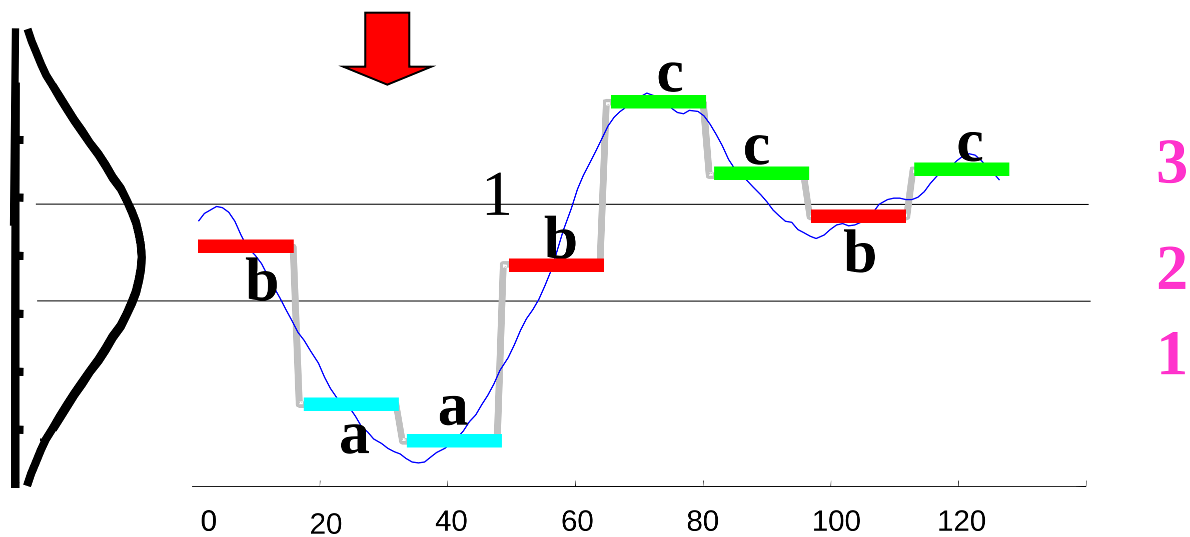
Eamonn Keogh, eamonn@cs.ucr.edu

Symbolic Aggregate ApproXimation

First convert the time series to PAA representation, then convert the PAA to symbols

It takes linear time





The *word size*, in this case 8.

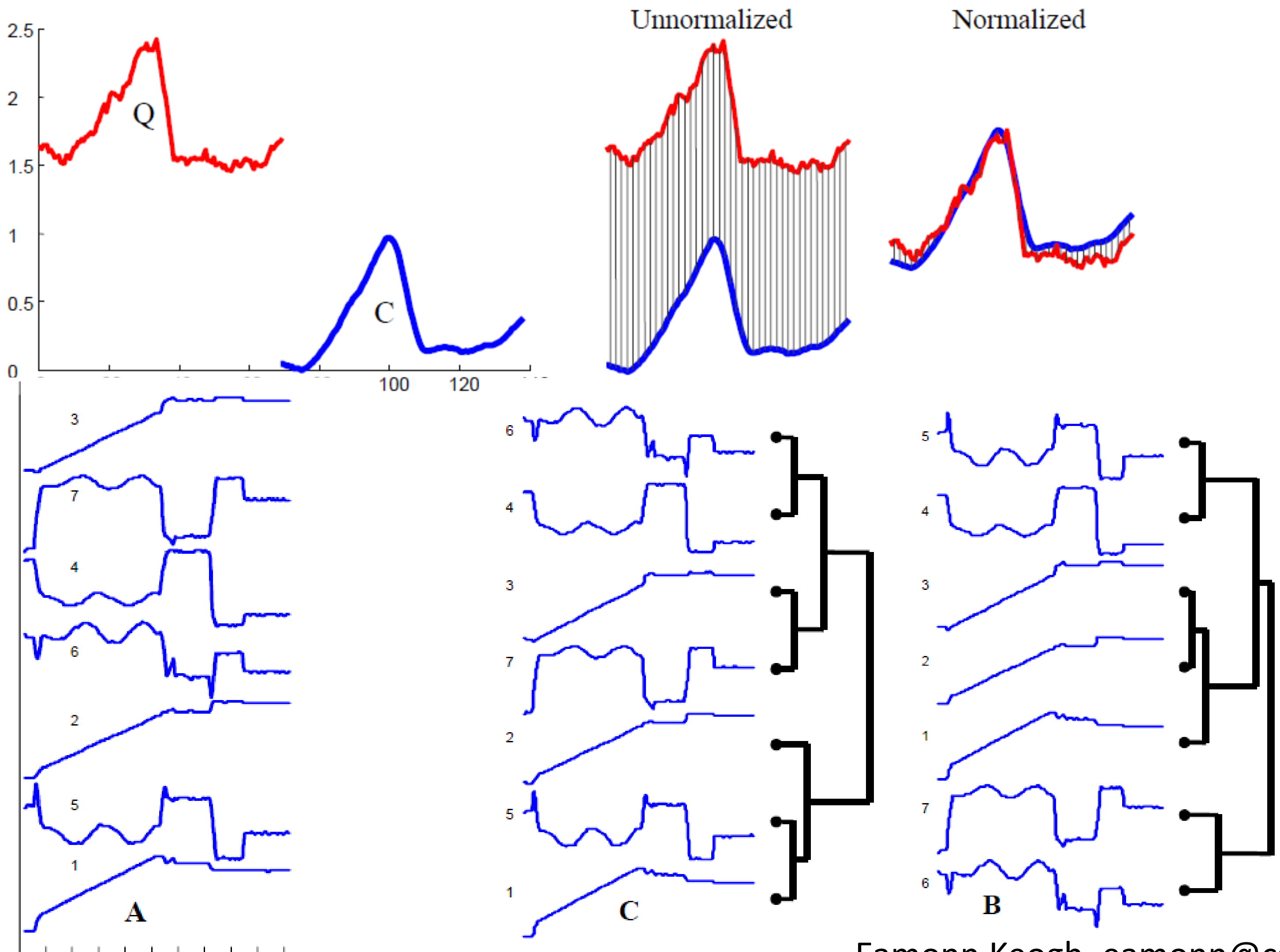
The *alphabet size* (cardinality), in this case 3.

baabccbc

Eamonn Keogh, eamonn@cs.ucr.edu

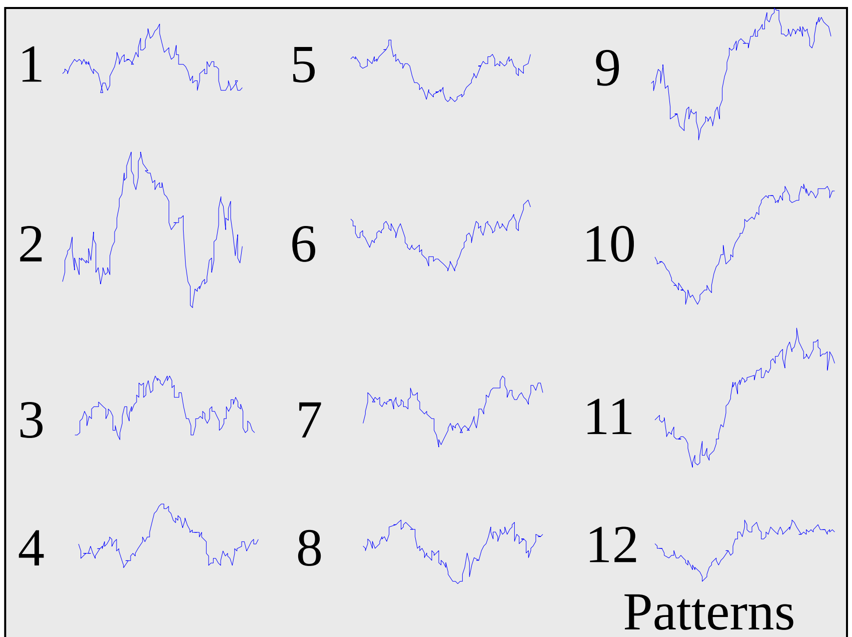
Clustering

Importance of data preprocessing/normalizing – offset & scaling



Eamonn Keogh, eamonn@cs.ucr.edu

Time Series Filtering



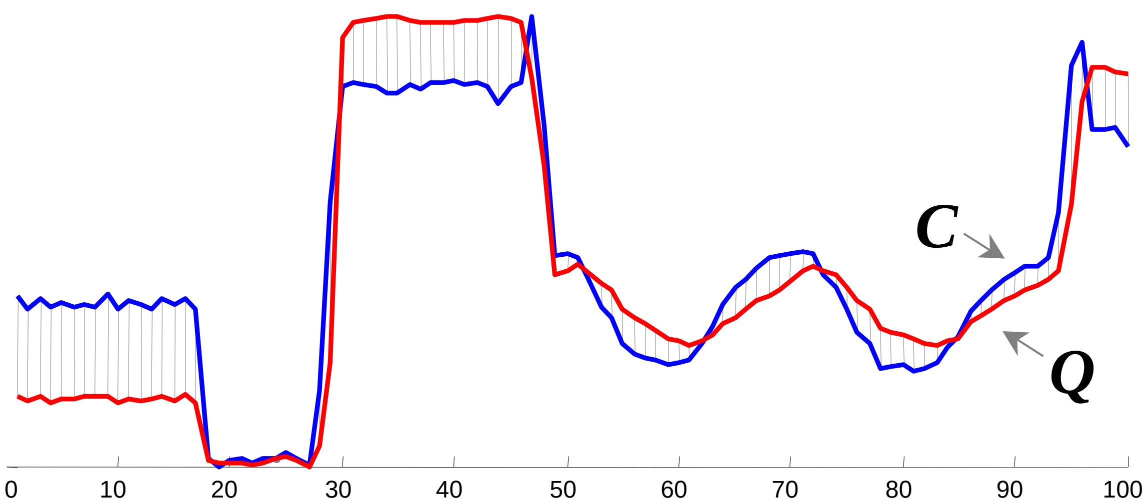
Given a time series ***T***, a set of patterns ***C*** and a distance threshold ***r***, find all subsequences in ***T*** that are within ***r*** distance to any of the patterns in ***C***.

Eamonn Keogh, eamonn@cs.ucr.edu

Euclidean Distance Metric

Given two time series $Q = q_1 \ldots q_n$ and $C = c_1 \ldots c_n$, the Euclidean distance between them is defined as:

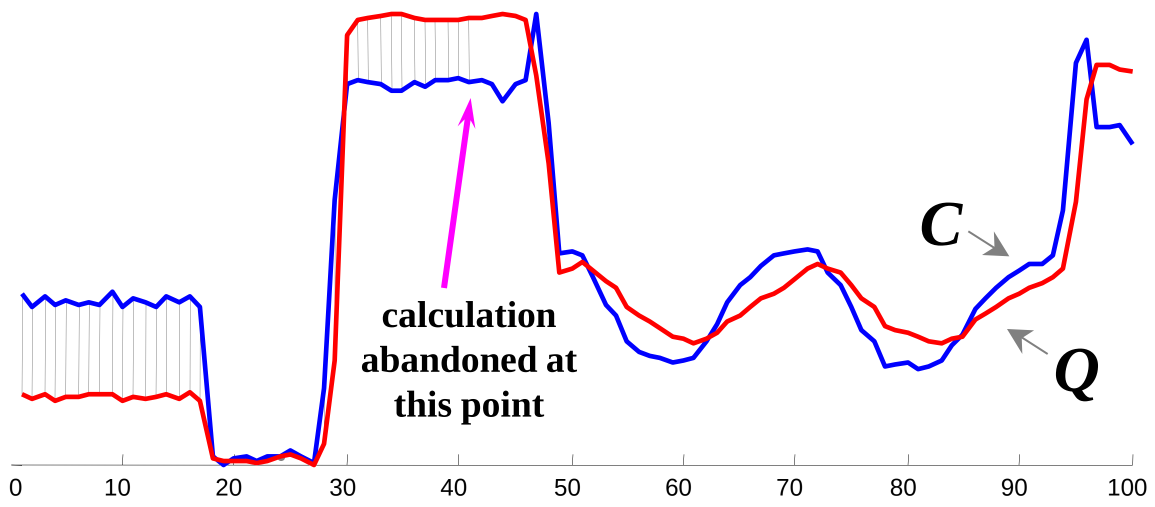
$$ED(Q,C) \equiv \sqrt{\sum_{i=1}^{n} (q_i - c_i)^2}$$



Eamonn Keogh, eamonn@cs.ucr.edu

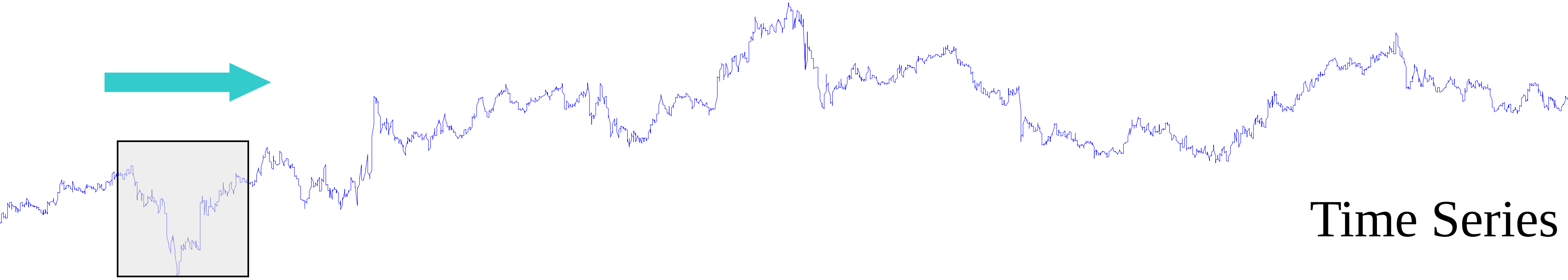
Early Abandon Euclidean Distance

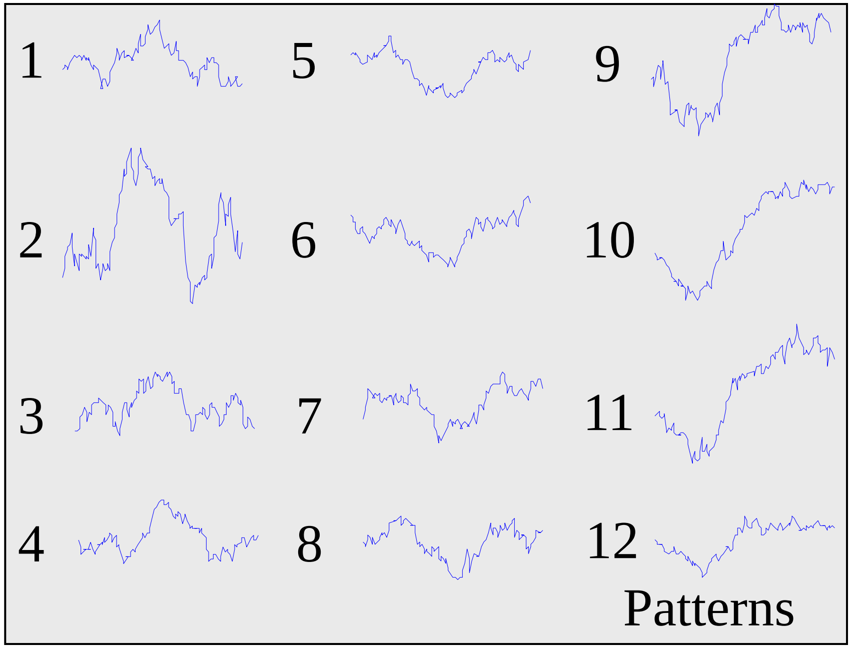
During the computation, if current sum of the squared differences between each pair of corresponding data points exceeds r^2, we can safely **abandon** the calculation.



Eamonn Keogh, eamonn@cs.ucr.edu

Classic Approach





Individually compare each pattern sequence to the time series subsequence using the early abandoning Euclidean distance measure.

Eamonn Keogh, eamonn@cs.ucr.edu

Logic Unit: Wedge

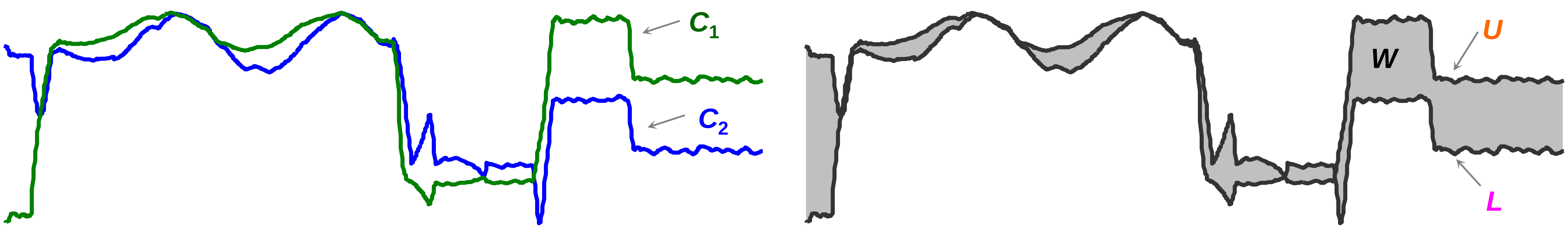
Having pattern sequences $C_1, .., C_k$, we can form two new sequences U and L :

$U_i = \max(C_{1i}, .., C_{ki})$

$L_i = \min(C_{1i}, .., C_{ki})$

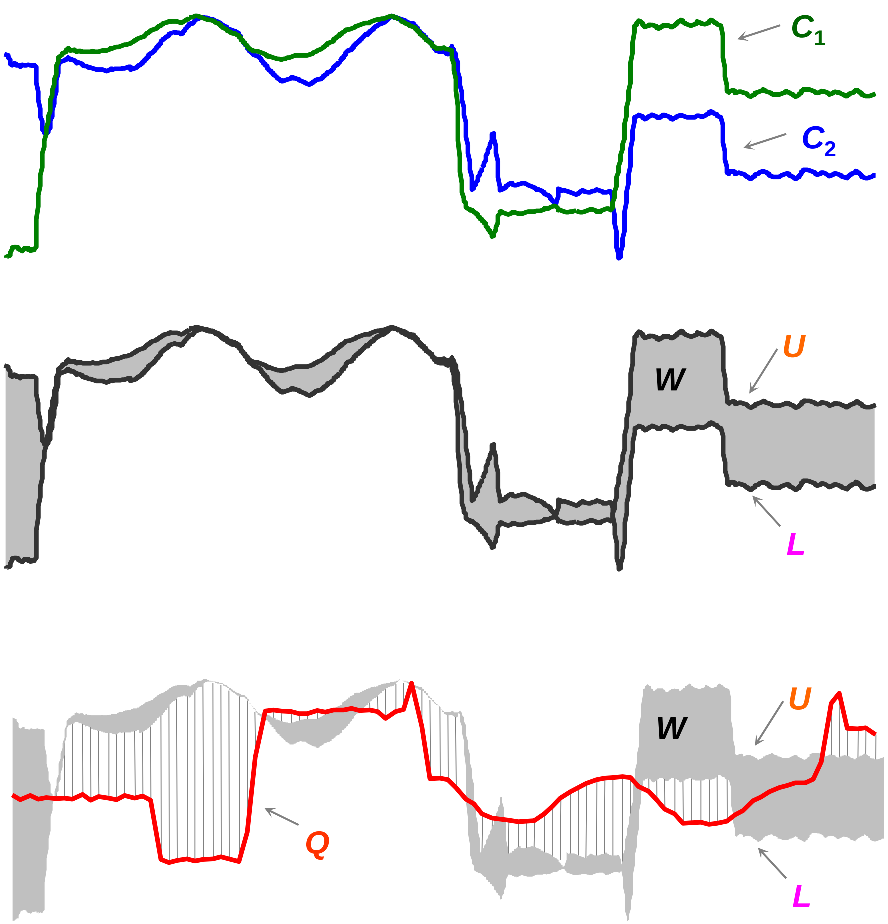
Wedge $W = \{U, L\}$:

the smallest possible bounding envelope that encloses sequences $C_1, .., C_k$



Eamonn Keogh, eamonn@cs.ucr.edu

Lower Bound Distance: LB_Keogh

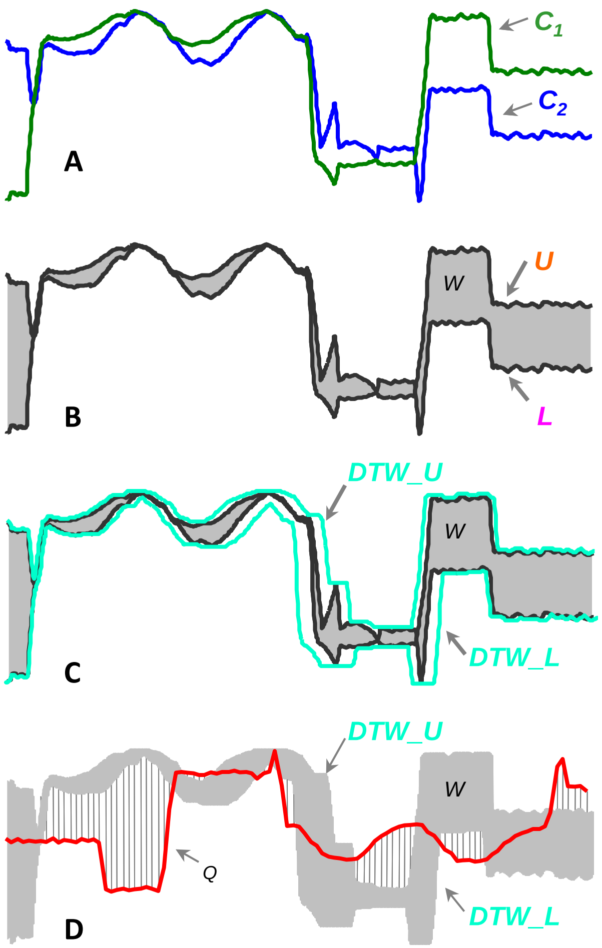


A lower bounding measure between an arbitrary subsequence Q and the entire set of pattern sequences contained in a wedge W:

$$LB\_Keogh(Q,W) = \sqrt{\sum_{i=1}^{n} \begin{cases} (q_i - U_i)^2 & \text{if } q_i > U_i \\ (q_i - L_i)^2 & \text{if } q_i < L_i \\ 0 & \text{otherwise} \end{cases}}$$

Eamonn Keogh, eamonn@cs.ucr.edu

Lower Bound Dynamic Time Warping



$U_i = \max(C_{1i}, .., C_{ki})$
$L_i = \min(C_{1i}, .., C_{ki})$

Wedge $W = \{U, L\}$

An additional envelope

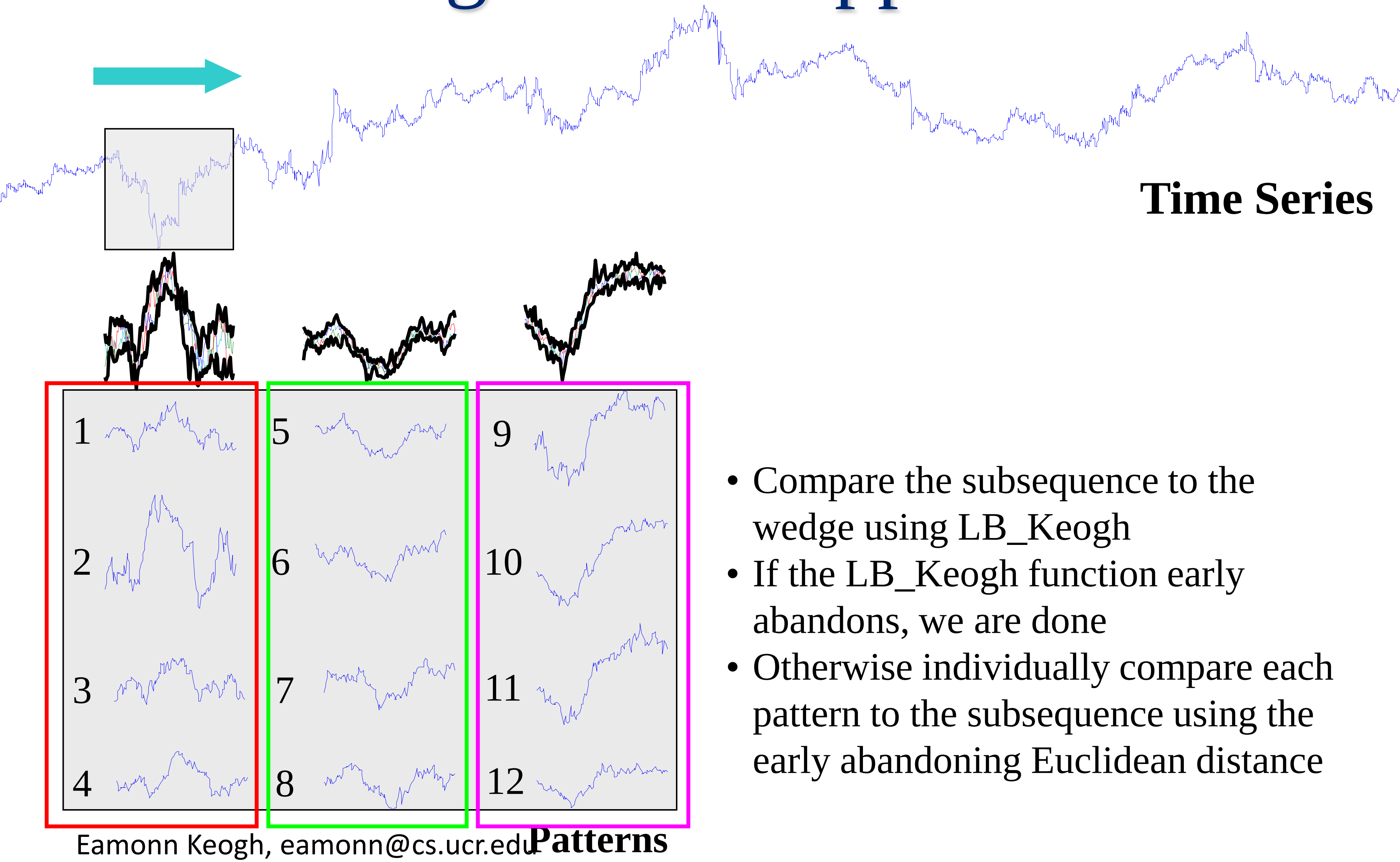
$DTW\_U_i = \max(U_i\text{-}R : U_i\text{+}R)$
$DTW\_L_i = \min(L_i\text{-}R : L_i\text{+}R)$
where R is the allowed warping range.

DTW lower bounding function:

$$LB\_Keogh_{DTW}(Q,W) = \sqrt{\sum_{i=1}^{n} \begin{cases} (q_i - DTW\_U_i)^2 & \text{if } q_i > DTW\_U_i \\ (q_i - DTW\_L_i)^2 & \text{if } q_i < DTW\_L_i \\ 0 & \text{otherwise} \end{cases}}$$

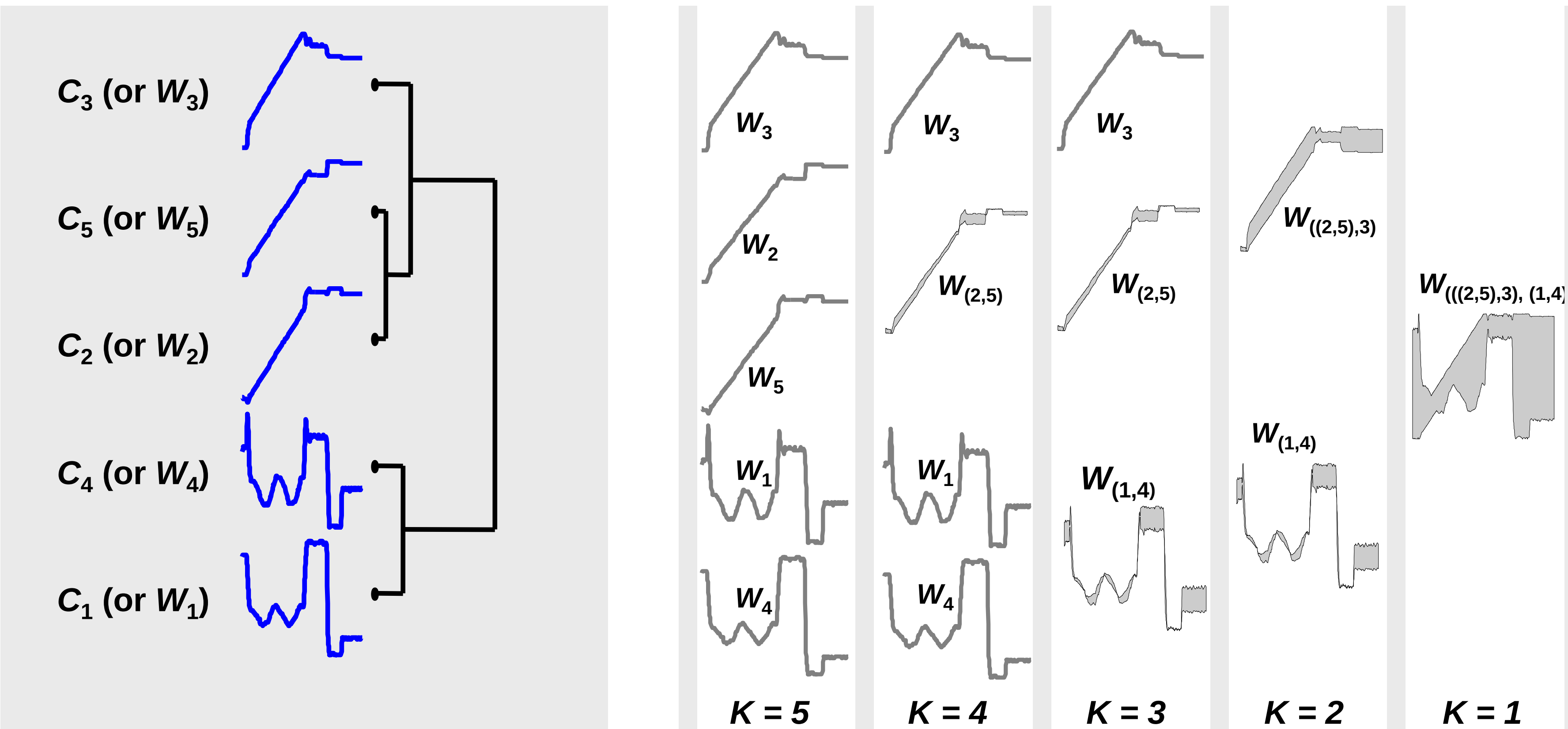
Eamonn Keogh, eamonn@cs.ucr.edu

Wedge Based Approach



- Compare the subsequence to the wedge using LB_Keogh
- If the LB_Keogh function early abandons, we are done
- Otherwise individually compare each pattern to the subsequence using the early abandoning Euclidean distance

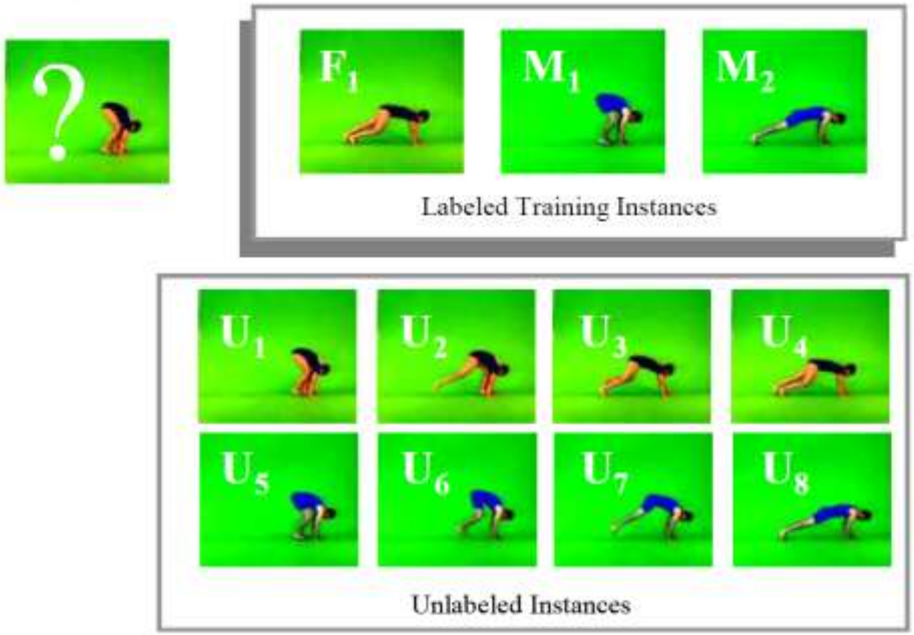
Hierarchal Clustering

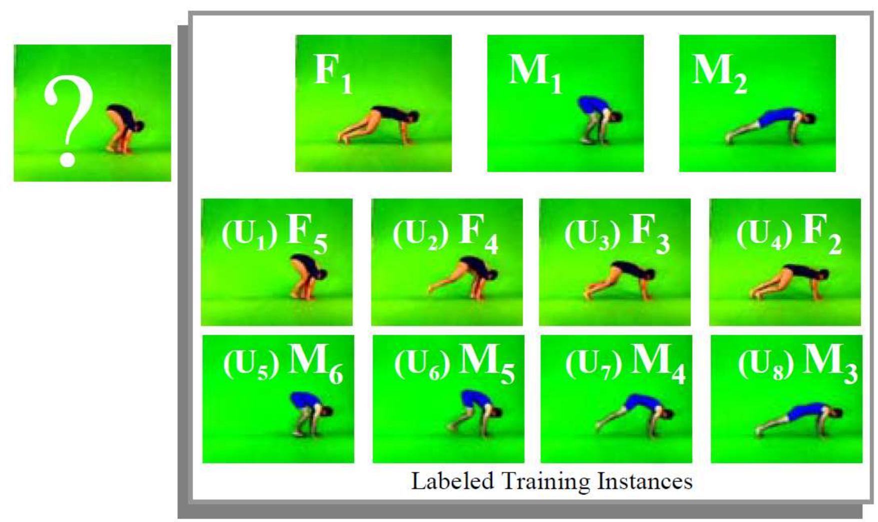


Which wedge set to choose ?

Eamonn Keogh, eamonn@cs.ucr.edu

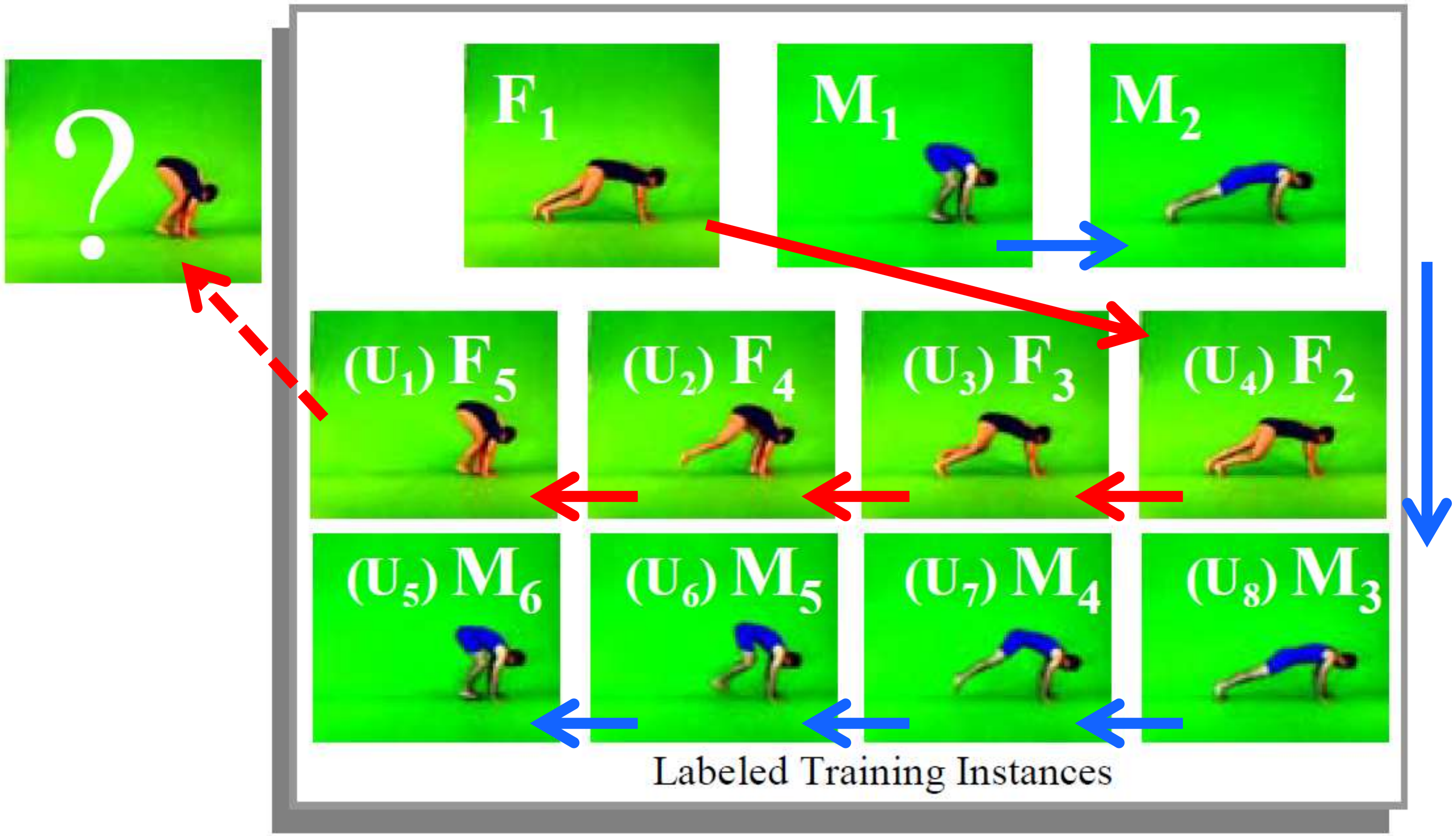
Classification – Semi-supervised
…Unlabeled data have value too…





Eamonn Keogh, eamonn@cs.ucr.edu

Classification – Semi-supervised



Li Wei Eamonn Keogh, **Semi-Supervised Time Series Classification,** KDD '06 Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining
Pages 748-753

Training the Classifier

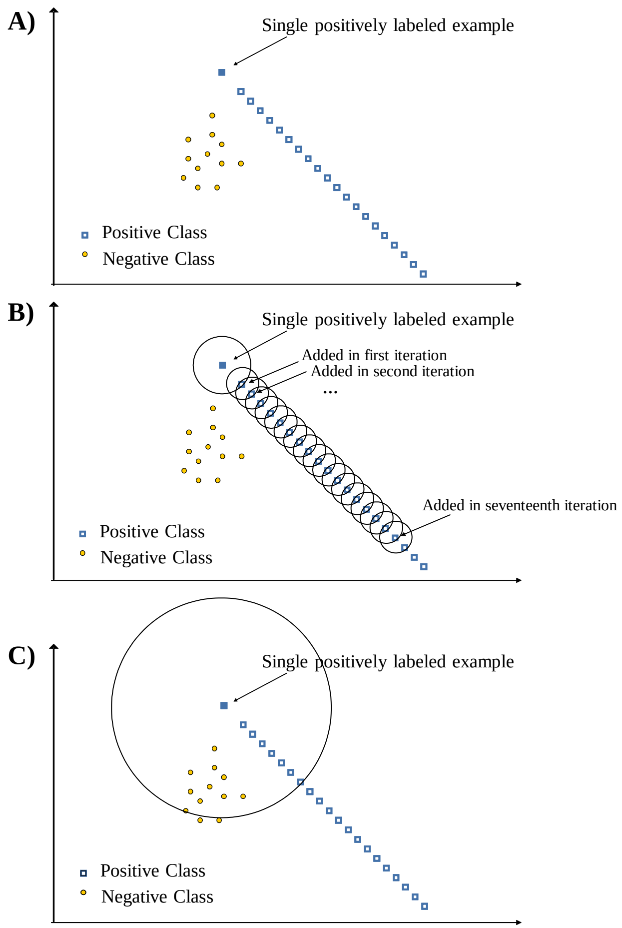


Figure 4: Semi-supervised training on a simple two-class dataset yields much higher accuracy than a naive *k*-nearest-neighbor classifier

We let the classifier teach itself by its own predication.

For example in Figure 4 we have a two-class dataset, where initially only one example is known as positive (the solid square in subplot A). In subplot B, we can see the chaining effect of semi-supervised learning: a positive example is labeled which helps labeling other positive examples and so on. Eventually all positive examples are correctly classified. In contrast, if we simply put the seventeen nearest neighbors of the single labeled example to the positive class, we will get very poor accuracy (see subplot C).

Eamonn Keogh, eamonn@cs.ucr.edu

Indexing of Handwritten Documents

A)
B)
C)

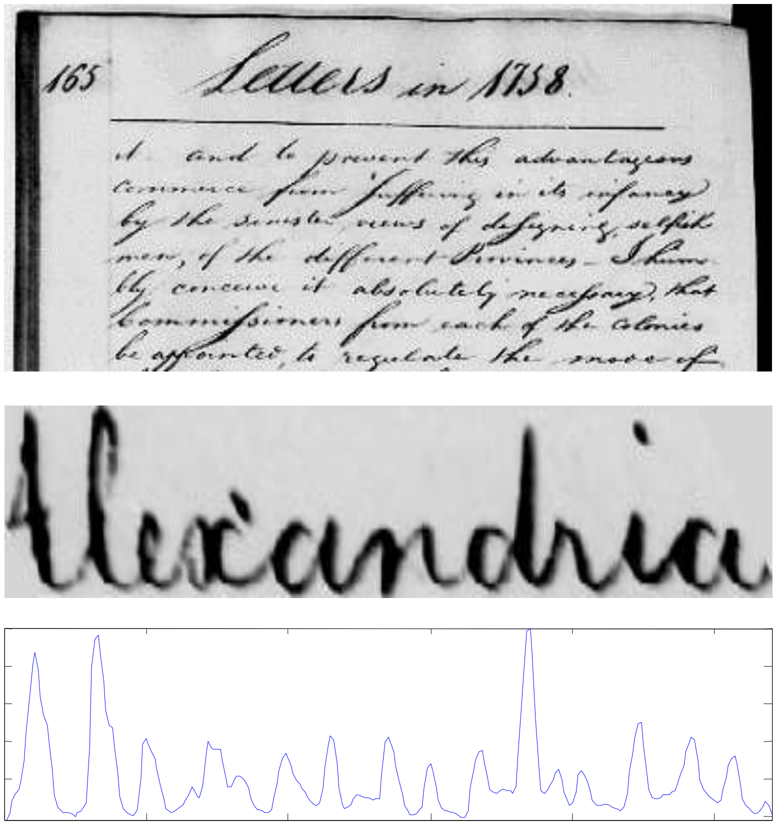


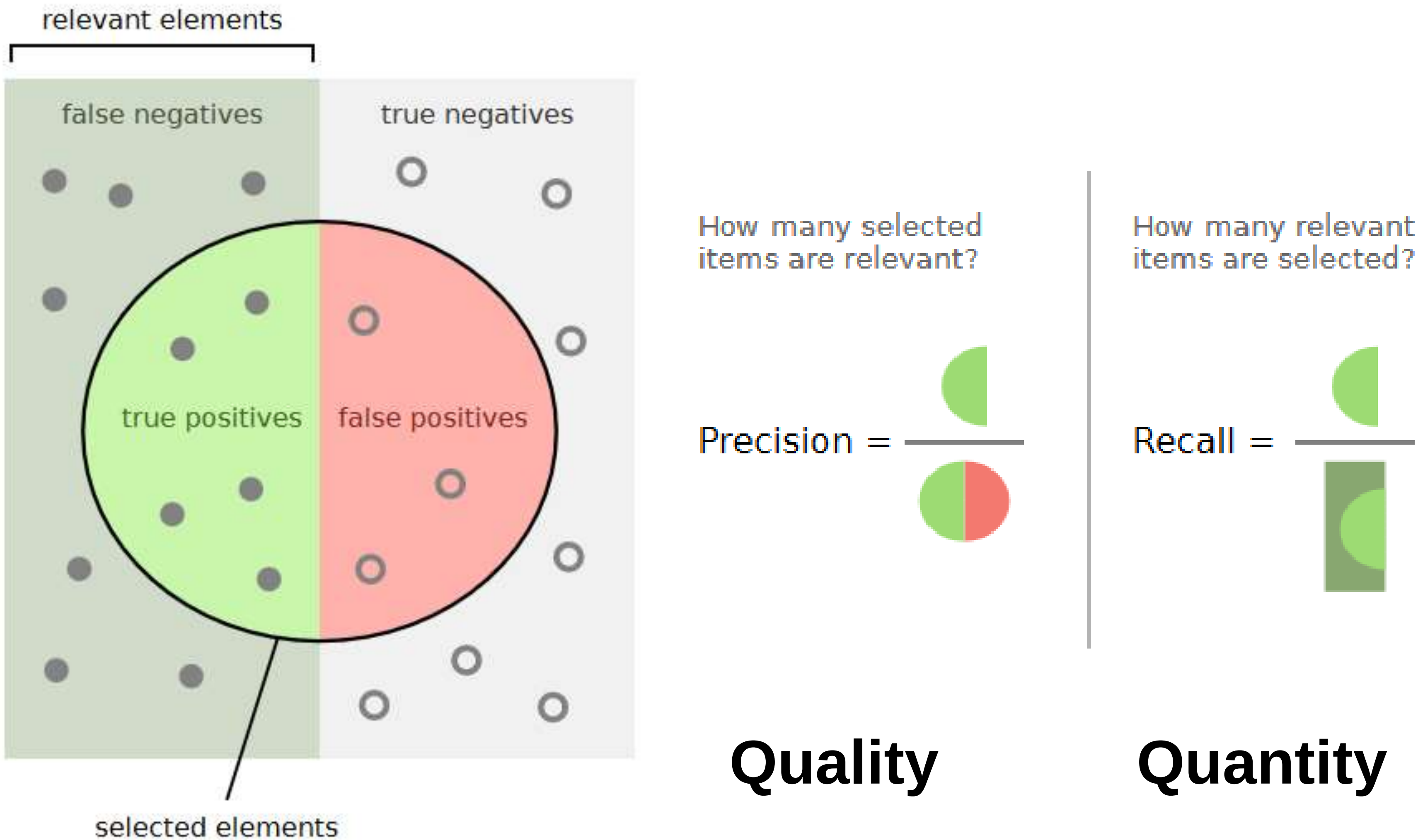
Figure 1: **A)** A sample of text written by George Washington. **B)** The word “Alexandria” after having its slant removed. **C)** A time series created by tracing the upper profile of the word (Image courtesy of Raghavan Manmatha, used with permission)

There has been a recent explosion of interest in indexing handwritten documents. Note that simply treating the words as “time series” (see Figure 1) is an extremely competitive approach for classifying (and thus indexing) handwritten documents.

Handwriting classifiers must be trained on each individual’s particular handwriting. However the cost of obtaining labeled data for each word, for every individual is very expensive as measured in human time. A semi-supervised approach where a user annotates just a few training examples would have great utility.

Eamonn Keogh, eamonn@cs.ucr.edu

Precision & Recall



Quality **Quantity**

More data sources

www.cs.ucr.edu/~eamonn/time_series_data/

UCR Time Series Classification Archive

Last major update, Summer 2015: Early work on this data resource was funded by an NSF Career Award 0237918, and it continues to be funded through NSF IIS-1161997 II and NSF IIS 1510741

We suggest you begin by reading the briefing document in PDF or PowerPoint, which also contains the password. Then you can download the entire archive (about 350mb in zipped format).

Please reference as: Yanping Chen, Eamonn Keogh, Bing Hu, Nurjahan Begum, Anthony Bagnall, Abdullah Mueen and Gustavo Batista (2015). *The UCR Time Series Classification Archive.* URL www.cs.ucr.edu/~eamonn/time_series_data/

```
@misc{UCRArchive,
title={The UCR Time Series Classification Archive},
author={ Chen, Yanping and Keogh, Eamonn and Hu, Bing and Begum, Nurjahan and Bagnall, Anthony and Mueen, Abdullah and Batista, Gustavo},
year={2015},
month={July},
note = {\url{www.cs.ucr.edu/~eamonn/time_series_data/}}
}
```

Name	First paper or data creator	Number of classes	Size of training set	Size of testing set	Time series Length	1-NN Euclidean Distance	1-NN Best Warping Window DTW (*r*) Note that *r* is the percentage of time series length	1-NN DTW, no Warping Window
Synthetic Control	Pham	6	300	300	60	0.120	0.017 (6)	0.007
Gun-Point	Ratanamahatana	2	50	150	150	0.087	0.087 (0)	0.093
CBF		3	30	900	128	0.148	0.004 (11)	0.003
Face (all)	Xi	14	560	1690	131	0.286	0.192 (3)	0.192
OSU Leaf	Gandhi	6	200	242	427	0.479	0.388 (7)	0.409
Swedish Leaf	Soderkvist	15	500	625	128	0.211	0.154 (2)	0.208
50Words	Rath	50	450	455	270	0.369	0.242 (6)	0.310
Trace	Roverso	4	100	100	275	0.240	0.010 (3)	0.000
Two Patterns	Geurts	4	1000	4000	128	0.090	0.002 (4)	0.000
Wafer	Olszewski	2	1000	6174	152	0.005	0.005 (1)	0.020
Face (four)	Ratanamahatana	4	24	88	350	0.216	0.114 (2)	0.170
Lightning-2	Eads	2	60	61	637	0.246	0.131 (6)	0.131
Lightning-7	Eads	7	70	73	319	0.425	0.288 (5)	0.274
ECG	Olszewski	2	100	100	96	0.120	0.120 (0)	0.230
Adiac	Jalba	37	390	391	176	0.389	0.391 (3)	0.396
Yoga	Xi	2	300	3000	426	0.170	0.155 (2)	0.164
Fish	Lee	7	175	175	463	0.217	0.154(4)	0.177
Plane		7	105	105	144	0.038	0.000 (6)	0.000
Car		4	60	60	577	0.267	0.233 (1)	0.267
Beef	Tony Bagnall	5	30	30	470	0.333	0.333 (0)	0.367

Black sheep
Detecting vehicles on the road that should not be there

Jan Skácel
December 2017

Roadside Particulate Matter Emission Measurement – Prague October 2017

Quick facts:

- 10 days of measurement

~ 12 500 vehicles

- Particulate data - count, size & mass -
- Gaseous emissions CO_2 CO, NO_x

Emission Measurement Goals

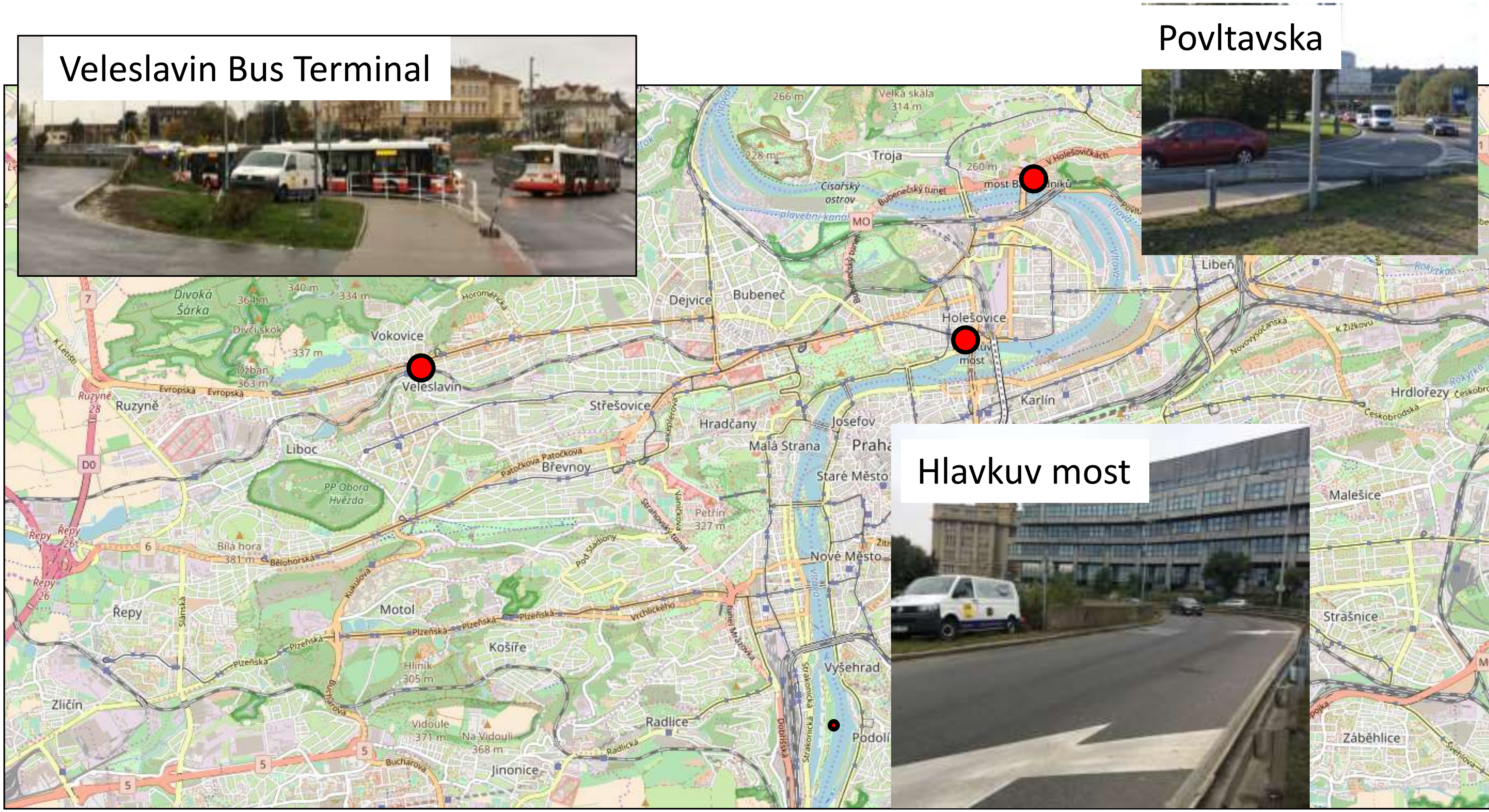
- Vehicle fleet Emission Distribution
- High particulate mass emitters identification
- Comparison with technical data from vehicle registry based on vehicle registration plate
- Judgement if vehicle OK / NOK

Scope
10 days, 3 locations

Recorded ~ 18.000 vehicles, ~12.500 vehicles identified by plate

Measured values:
- Particulate matter – count, size distribution počet and mass
- Gaseous emissions CO_2 CO, NO_x
- Vehicle plate, speed time

3 Emission Measurement Sites in Prague



Data - Visual

Camera 1

Online data processing
- Picture
- List of vehicles – time & plate

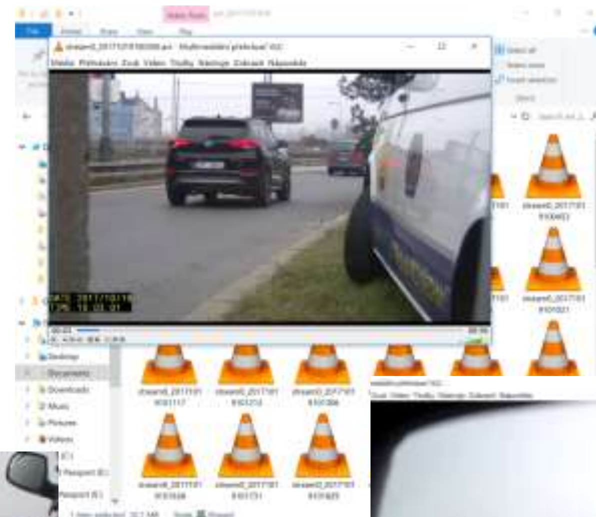
Camera 2

Continuous video
Post-processing
- plate recognition

Camera 3

Overall location video
“True vehicle count”

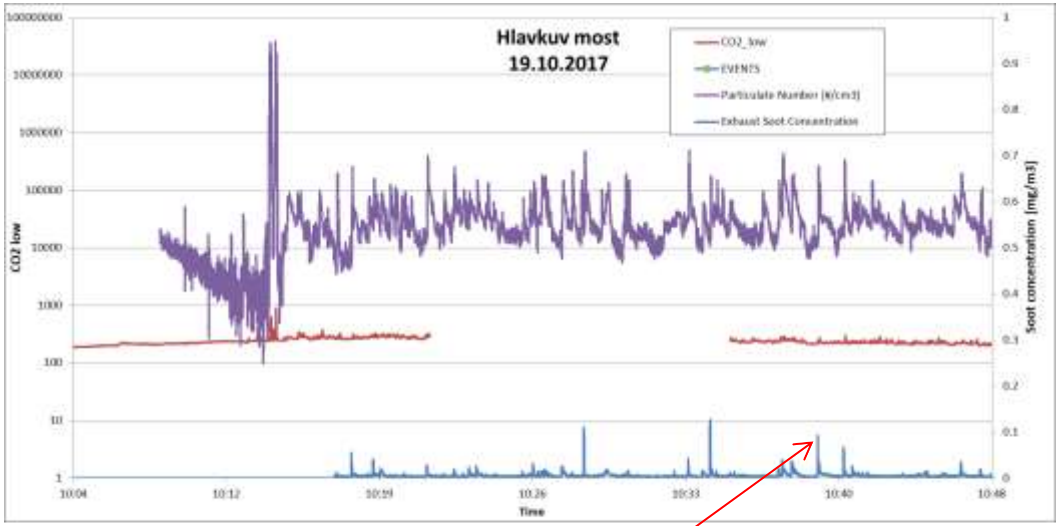
Plate No.	Capture Time	Country	Validity
2SL1722	19.10.2017 09:47:42	Czech Republic(CZE)	87
2Z68863	19.10.2017 09:47:46	Czech Republic(CZE)	78
4E78140	19.10.2017 09:47:56	Czech Republic(CZE)	84
3AY2703	19.10.2017 09:47:58	Czech Republic(CZE)	85
7A84858	19.10.2017 09:47:59	Czech Republic(CZE)	85
2SP4295	19.10.2017 09:48:02	Czech Republic(CZE)	82
5AS3754	19.10.2017 09:48:04	Czech Republic(CZE)	82
1BL8155	19.10.2017 09:48:05	Czech Republic(CZE)	82
4SC4081	19.10.2017 09:48:11	Czech Republic(CZE)	82
5AB3109	19.10.2017 09:48:38	Czech Republic(CZE)	78
3SC3641	19.10.2017 09:48:40	Czech Republic(CZE)	83
[illegible]	[illegible]	[illegible]ZE)	78
[illegible]	[illegible]	[illegible]ZE)	86
[illegible]	[illegible]	[illegible]ZE)	82
[illegible]	[illegible]	[illegible]ZE)	79
[illegible]	[illegible]	[illegible]ZE)	86
[illegible]	[illegible]	[illegible]ZE)	83
[illegible]	[illegible]	[illegible]ZE)	82
[illegible]	[illegible]	[illegible]ZE)	80
[illegible]	[illegible]	[illegible]ZE)	71
[illegible]	[illegible]	[illegible]ZE)	84
[illegible]	[illegible]	[illegible]ZE)	87
5A97[illegible]	[illegible]	[illegible]	[illegible]
4L94[illegible]	[illegible]	[illegible]	[illegible]
9T81[illegible]	[illegible]	[illegible]	[illegible]
5AN[illegible]	[illegible]	[illegible]	[illegible]
4AX[illegible]	[illegible]	[illegible]	[illegible]
2AL5[illegible]	[illegible]	[illegible]	[illegible]
5AD[illegible]	[illegible]	[illegible]	[illegible]
3AE[illegible]	[illegible]	[illegible]	[illegible]
9A58[illegible]	[illegible]	[illegible]	[illegible]
9A81[illegible]	[illegible]	[illegible]	[illegible]
5H56[illegible]	[illegible]	[illegible]	[illegible]
5M20388	19.10.2017 09:52:26	Czech Republic(CZE)	83
5AE6680	19.10.2017 09:52:36	Czech Republic(CZE)	83
2SK9114	19.10.2017 09:52:37	Czech Republic(CZE)	80
6AK4443	19.10.2017 09:52:39	Czech Republic(CZE)	86
4AR6740	19.10.2017 09:52:41	Czech Republic(CZE)	81
5AX7696	19.10.2017 09:52:42	Czech Republic(CZE)	85
MVJ832	19.10.2017 09:52:47	Hungary(HUN)	85
99H9999	19.10.2017 09:52:51	Europe Region(NON)	82



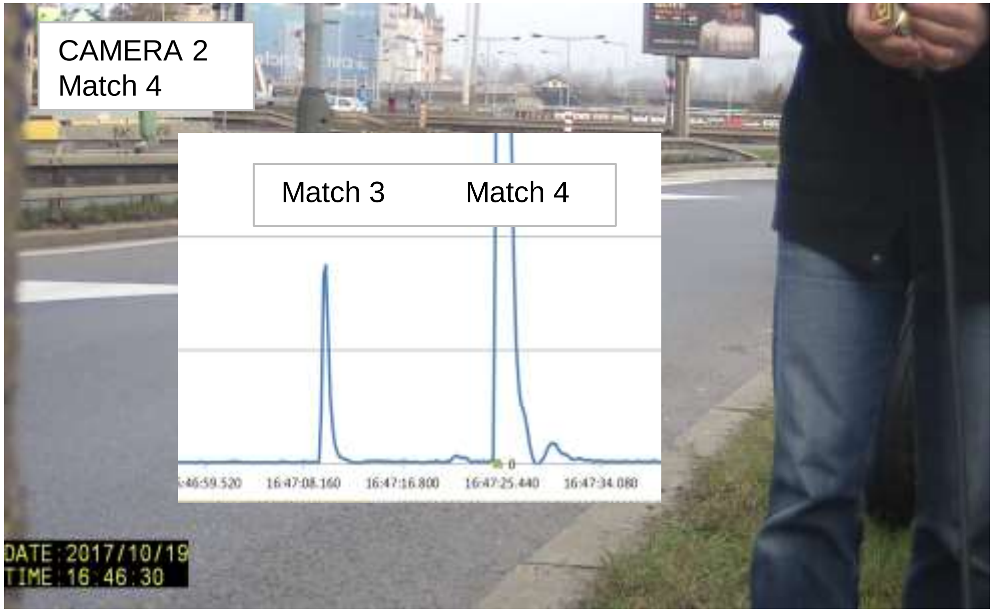
Visual Data Processing, Time synchronization

			MSS	FTIR	KAMERA 1	KAMERA 2
19.10.2017						
HLAVKUV MOST						
10:14		START				
13:29	6A0 7613	IVECO	13:29:07		13:30:13	13:28:13
13:51	2K8 1351	BILY BUS	13:51:15		13:52:21	13:50:21
14:25:30	4AU 3681	VW TRANSPORTER MODRY (KAMERA HIKVISION CAS 14:27:08)	14:26:02		14:27:08	14:25:08
15:12:30	1C9 0914	AVIA (KAMERA HIKVISION CAS 15:14:00)	15:12:54		15:14:00	15:12:00
			16:33:37		16:34:43	
		SIRKA 3	16:47:09	0:00:55		16:46:14
16:47	FINISH	SIRKA 4	16:47:24	0:00:54	16:48:30	16:46:30
DOHLEDANE		SIRKA 1	9:54:55		9:56:01	9:54:01
		SIRKA 2	10:14:05		10:15:11	10:13:11
			11:30:50		11:31:56	11:29:56
		CO	12:04:09		12:05:15	12:03:15
	1SR 440	skutr - ?	12:05:35		12:06:41	12:04:41
	2C0 0786	OPEL OMEGA	14:20:42		14:21:48	14:19:48
	2AF 4473	odtahovka zluta	15:22:49		15:23:55	15:21:55
	X 1161 BX	autobus Bulgary	15:47:36		15:48:42	15:46:42
	LOT 052	autobus Hungary	16:38:53		16:39:59	16:37:59

Time synchronization of visual and emission data

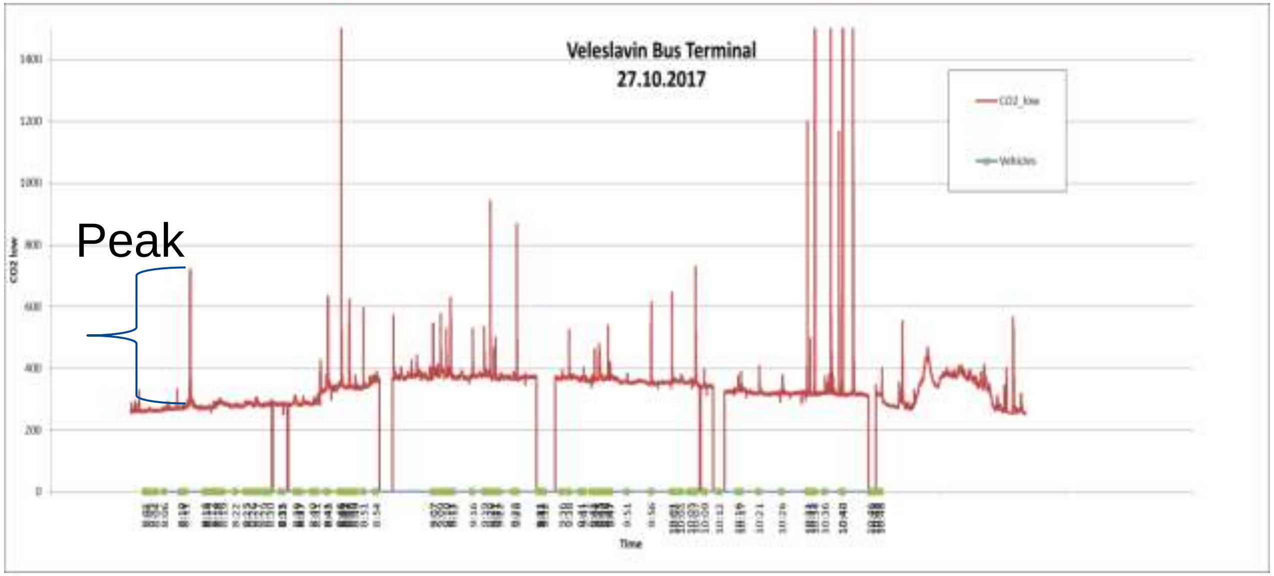


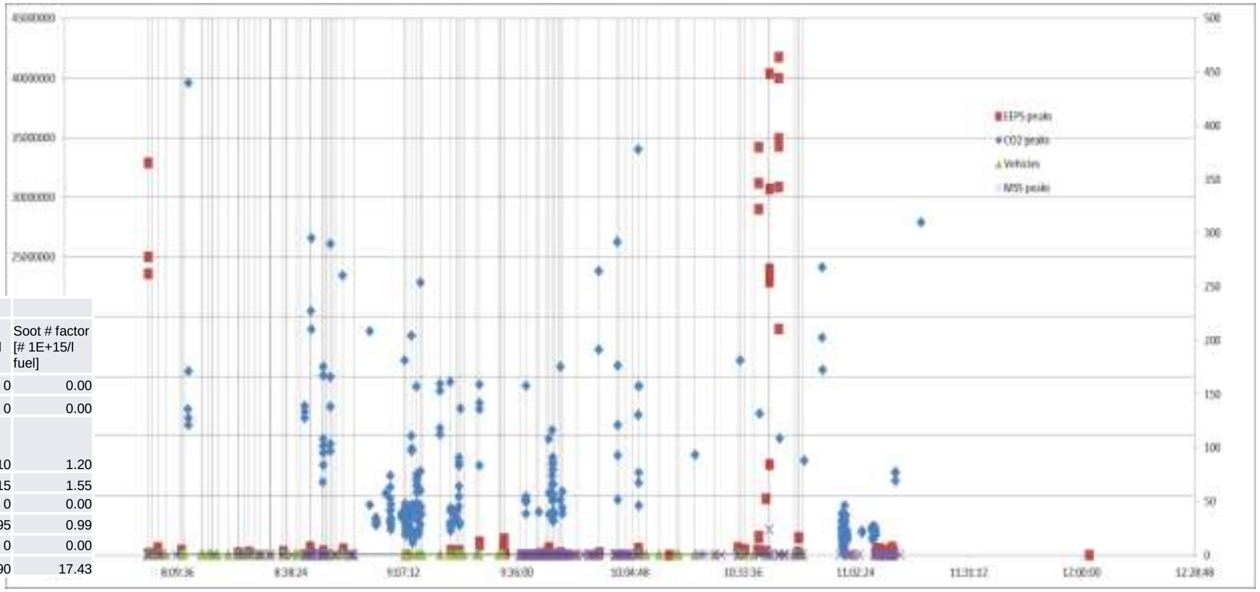
Disruption - “match test”
Significant polluter
Anomaly



Data Processing

- Time synchronization
- Offset
- Background removal
- Segmentation – peak representation
- Peaks above threshold without background
- Dynamic time warping
- Vectorization - Matching Vehicle, it's parameters and emission trace

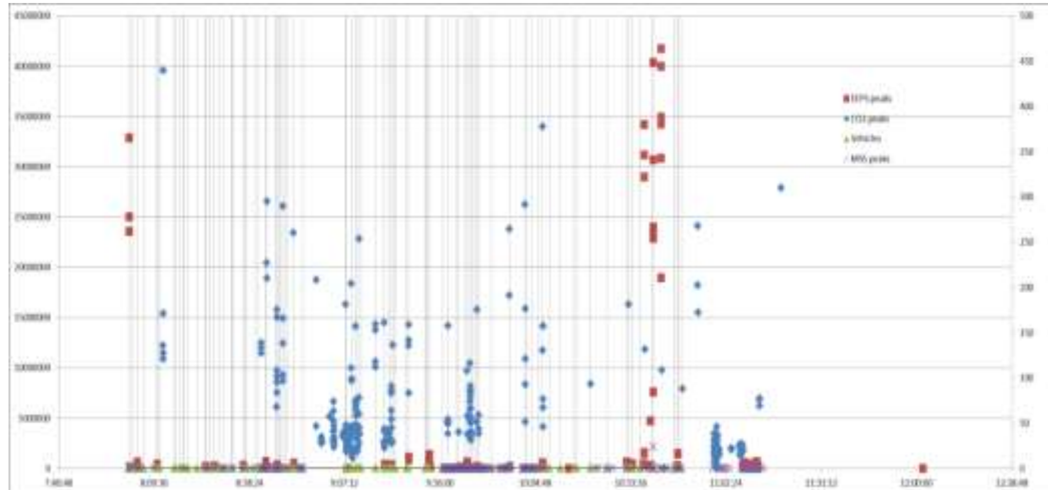


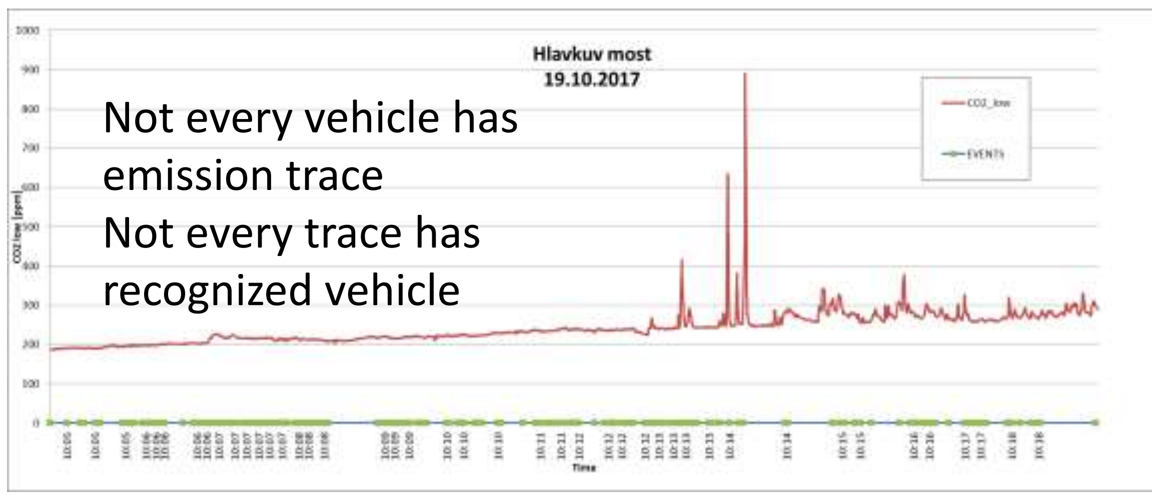


		ppm (vol)	mg/m3	#/cm3			
		PEAK CO2	PEAK MSS	PEAK EEPS	Fuel [l/m3 air]	Soot mass factor [mg/l fuel]	Soot # factor [# 1E+15/l fuel]
9:12:00	1S5 4286	0	0	0	0	0	0.00
9:12:00	1SM 4286	65.1946	0	0	4.59E-05	0	0.00
9:02:20	2AD 6284	274.204	0.04058	232330	0.000193	210	1.20
10:15:35	2AD 6291	1309.673	0.56661	1425016	0.000922	615	1.55
10:17:20	2AD 6291	0	0	0	0	0	0.00
10:32:25	2AD 7690	891.0962	0.12259	620400	0.000627	195	0.99
8:49:50	2AD 7690	0	0.00584	0	0	0	0.00
9:33:05	2AE 8019	128.9407	0.08985	1582300	9.08E-05	990	17.43

Data Processing - Dynamic time warping

– adjustments to analyzer sampling delay – vehicle distance & speed
- Nearest later vehicle within threshold
- Possibly no vehicle detected "orphan trace"





Not every vehicle has emission trace
Not every trace has recognized vehicle

Issue: Vehicle recognition

Online camera detects only about 60% of vehicles

Radar also detects only ~ 70% - closed coupled result in one long vehicle

Possible solution:

1 – additional vehicles from video post processing

2 - vehicle detection from emission trace

Hlavkuv most 18/10/2017 15:15:00 – 15:45:18

Source	Number of Vehicles	Success rate
Camera 1 – online plate list	261	61%
Camera 1 – secondary postprocessing from picture	TBD	
Camera 2 - postprocessing	TBD	
Camera 3 - Gopro manual count	426	100%
Radar – vehicle list	289	68%

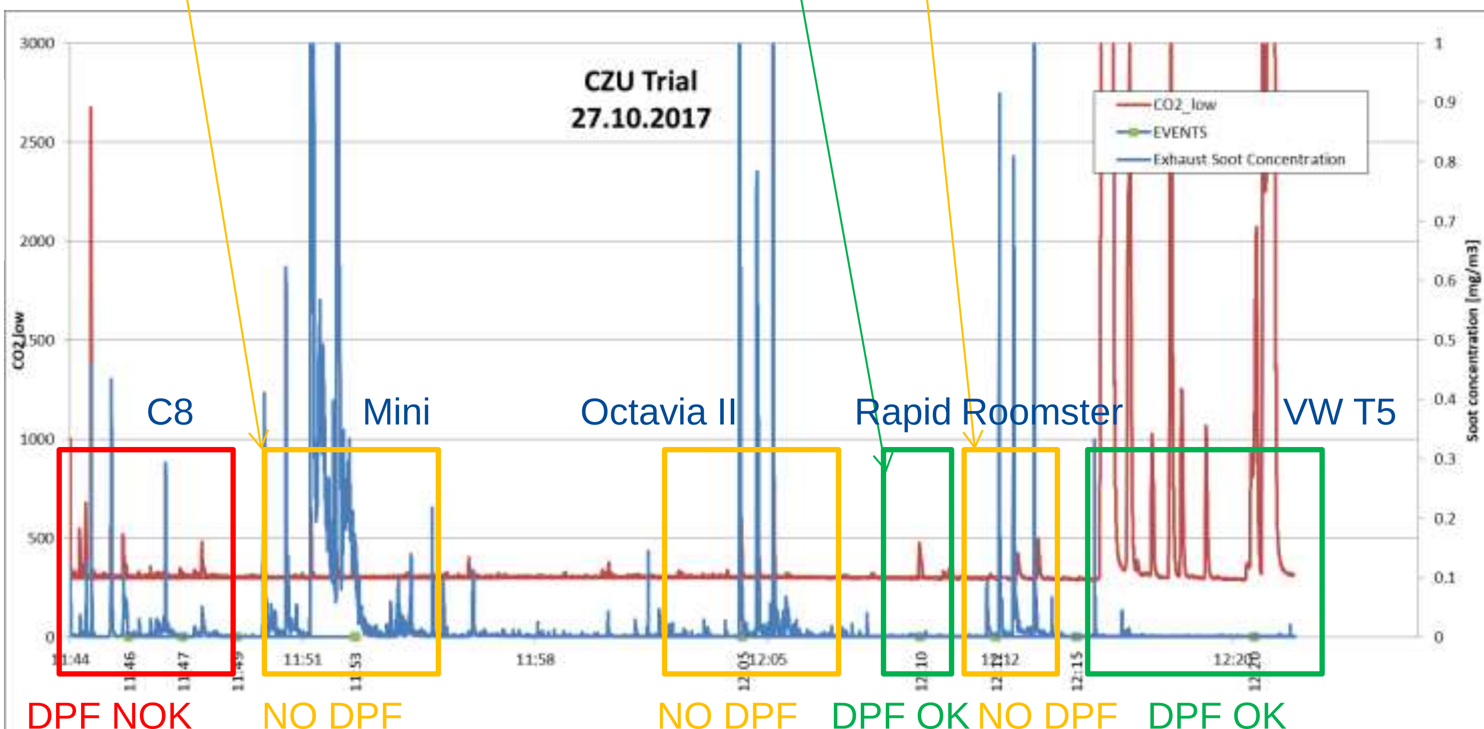
Roadside Measurement Concept Verification

Vehicle without DPF

Vehicle with DPF OK

Vehicle without DPF

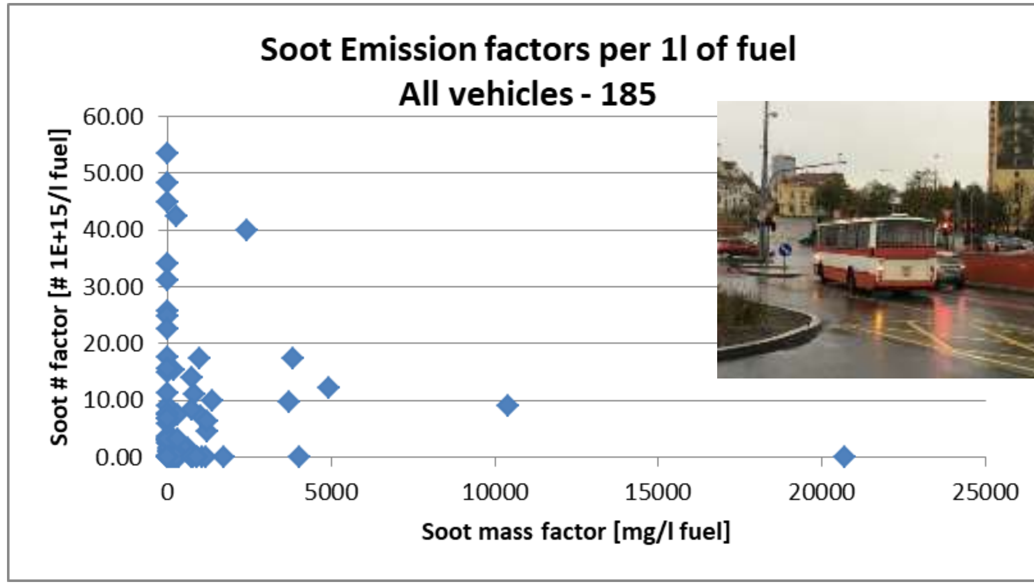
Pattern:
High CO_2 & no soot indicates good DPF

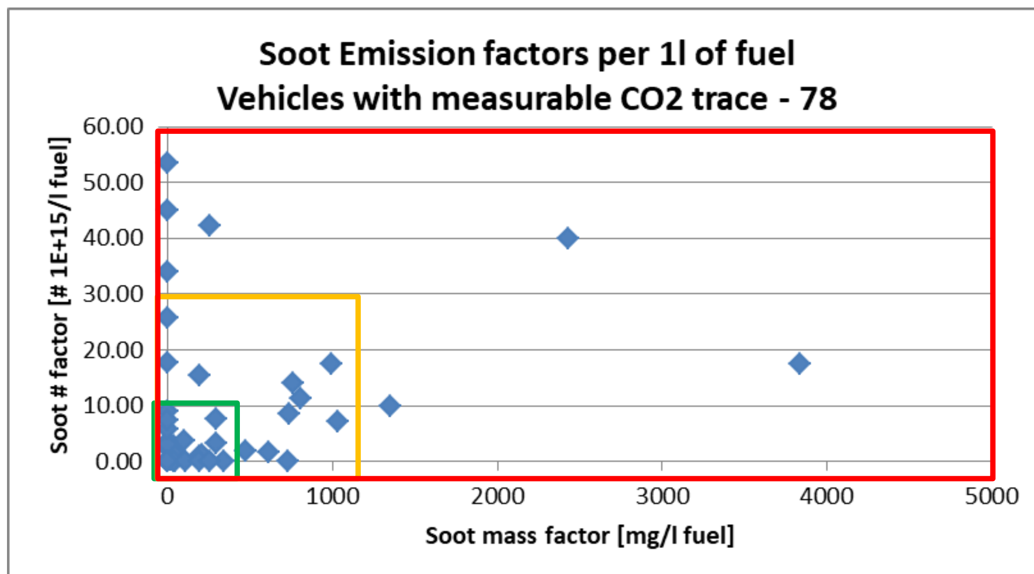


Partial Output – Emissions of Buses

180 Vehicles identified

		ppm (vol)	mg/m3	#/cm3			
		PEAK CO2	PEAK MSS	PEAK EEPS	Fuel [l/m3 air]	Soot mass factor [mg/l fuel]	Soot # factor [# 1E+15/l fuel]
9:12:00	1S5 4286	0	0	0	0	0	0.00
9:12:00	1SM 4286	65.1946	0	0	4.59E-05	0	0.00
9:02:20	2AD 6284	274.204	0.04058	232330	0.000193	210	1.20
10:15:35	2AD 6291	1309.673	0.56661	1425016	0.000922	615	1.55
10:17:20	2AD 6291	0	0	0	0	0	0.00
10:32:25	2AD 7690	891.0962	0.12259	620400	0.000627	195	0.99
8:49:50	2AD 7690	0	0.00584	0	0	0	0.00
9:33:05	2AE 8019	128.9407	0.08985	1582300	9.08E-05	990	17.43
10:50:00	2AH 1458	0	0	0	0	0	0.00
9:21:10	2S8 9377	91.0344	0	373900	6.41E-05	0	5.84
08:18:20	2ST 7861	0	0	0	0	0	0.00
10:46:40	2ST 7861	0	0	0	0	0	0.00
08:47:15	2ST 7862	0	0.02133	0	0	758	0.00
08:11:10	2SV 6964	0	0	0	0	0	0.00
10:02:10	2SV 7438	0	0.00562	0	0	0	0.00
10:02:12	2SV 7438	0	0	0	0	0	0.00
10:02:05	2SV 7439	433.616	0	0	0.000305	0	0.00
9:25:58	2SV 7449	0	0	0	0	0	0.00
9:26:29	2SV 7449	0	0	0	0	0	0.00
08:16:00	3A1 4198	0	0	0	0	0	0.00
08:33:12	3A1 4579	0	0.01205	0	0	0	0.00
08:49:00	3A2 3682	0	0	0	0	0	0.00
08:36:22	3AD 6291	0	0.13872	341100	0	4927	12.12
9:41:40	3AH 5486	39.9735	0.00717	0	2.81E-05	255	0.00
08:25:10	3AH 5486	0	0.00892	0	0	0	0.00
10:19:50	3AM 4059	0	0	0	0	0	0.00
08:43:20	3B5 4537	0	0.00695	700400	0	0	24.88
9:25:45	3J9 0188	0	0	0	0	0	0.00
08:30:30	3S1 3536	0	0.01148	0	0	0	0.00
10:48:57	3S1 3536	0	0	0	0	0	0.00





Next stetps

Improve vehicle be identified
– visual
- emission trace

Repeatability / Robustness
- repeated measurements consistent

Combine measurement with technical data from registry
- true negatives

Any patterns / typical features of a polluter – vehicle type, age, region of registration, make, model

Automated polluter recognition? Plate – emission trace – make, model

Variation between identical vehicles i.e. poor maintenance vs. responsible owner

Time/Location of last MOT (I/M) inspection

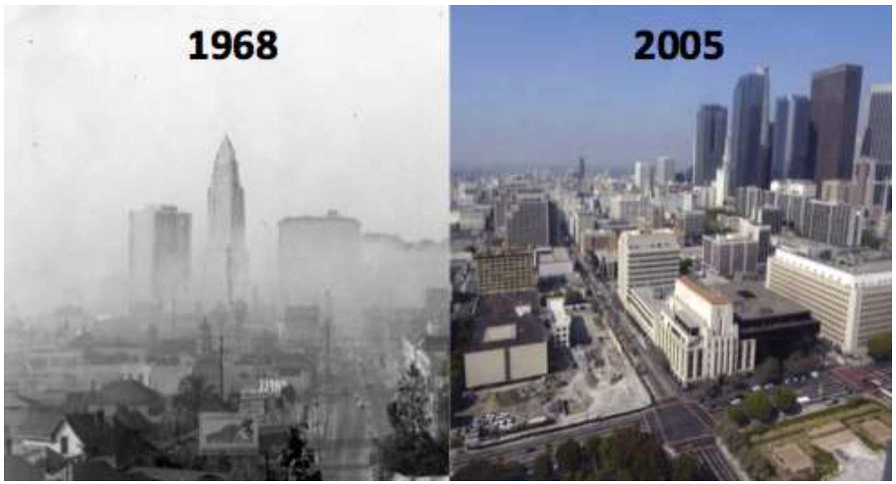
From 12.500 vehicles

Repeats	Number of Vehicles
2	429
3	49
4	52
5	23
6	6


There is Still Work to Do
Thank you!

Air pollution challenges of the past…

Los Angeles



L.A. 1979 smog siege

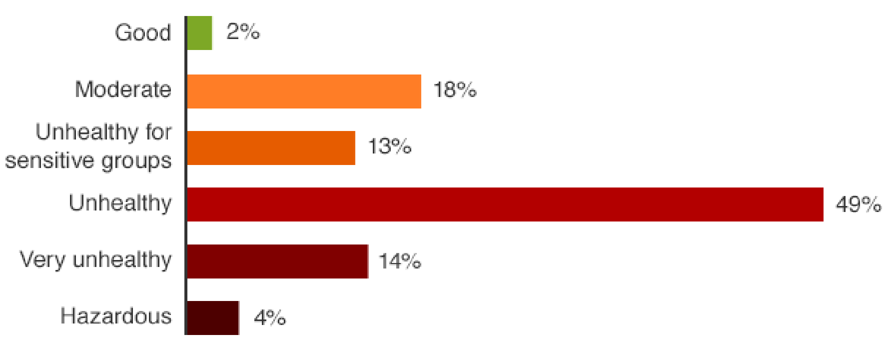
New York

Today's pollution is still a challenge… Beijing issued first Red Smog Alert ever

12/7/2015 Beijing, China issues first ever red smog alert expected to last 3 days
Schools in Beijing are to close and outdoor construction to stop
Beijing has 6 million registered cars.
Cars with odd and even number plates will be banned from driving on alternate days.

Beijing air quality 2008-2015

Daily average air quality index (AQI*) at US embassy, based on PM2.5 concentration readings



Daily average compiled from valid hourly readings Apr 2008-Jun 2015.
*AQI categories as set by the US Environmental Protection Agency

Source: US embassy, Beijing

Catalyst Testing – Early Beginning…

Engine Laboratory
Vehicle Testing – Light Duty Applications

Automated Robot Driver
(Horiba)