Mapper can produce a topological model for fibres of failure modes.

**Topology in the furnace:**

TDA as a diagnostics tool for process control systems.

Mikael Vejdemo-Johansson
*in transit* KTH → CUNY

Furnace picture with permission © Victor Mácha www.facebook.com/VIKTORMACHA.INDUSTRIAL www.viktormacha.com
Early work in progress

Collaboration with
Ayasdi Inc.
Outukumpu Stainless AB
KTH ITM (Industrial Engineering and Management)

Plan: use TDA to understand, diagnose and improve industrial process control and engineering.

First approach: diagnostics on existing control systems with Mapper.
What is mapper?
The Mapper algorithm

- Gurjeet Singh (2012)
- Built on a topological basic idea
- Creates intrinsic simplicial complex model of arbitrary data
Topological background

- Consider:
  - Spaces X, Y
  - Continuous map f: X → Y
  - Cover Y = ⋃ Y_i

- The cover pulls back to a cover X = ⋃ f^{-1}Y_i

- Refine cover to connected components X = ⋃ X_j; X_j ∈ π_0 f^{-1}Y_i

- If each X_j is contractible, Nerve lemma → nerve complex ≃ X.
Topological background
Topological background
Topological background
Topological background
Topological background
Topological background
From topology to data: a dictionary

- Topological space
- Continuous map $X \to Y$
- Cover
- $\pi_0$
- Nerve complex

- Point cloud
- *Filter function or lens* $X \to \mathbb{R}^d$
- Partition with overlap
- Clustering wrt *metric.*
- Nerve complex

Mapper is parametrized by a choice of *lens*(es), of *metric*, of (parameters for) *partition* and of *clustering* method.
Implementations

- mapper.m
- Ayasdi Core
- PyMapper
- TDAMapper (R)
- mirkoklukas/tda-mapper-py
- MLWave/kepler-mapper
Electric Arc Furnaces
Electric Arc Furnace

- Works by producing electric arcs from electrodes to scrap metal, producing heat that melts the metal.
- Standard 3-phase 220V 50Hz electricity.
- Consumes ~0.4 kWh/kg; theoretical minimum is ~300kWh.
Electric Arc Furnace

- Furnace in Avesta run by Outukumpu Stainless.
- Stainless: expensive scrap, high price output.
- Single charge produces 100 tonnes stainless steel.
- Approximately 5000 charges per year — 15-20 per day.
Temperature constraints

- Optimal temperature ~1600º
- Too low: not fully smelted
- Too high: entire batch spoiled

- Reference measurement possible: single use probe expensive and leaks heat.
- Metallurgical models available. Error spans ±400º. IQR -120º — +18º
Model parameters

- Known factors at any time point in production:
  - Element composition of the scrap
  - Energy used
  - Temperature of added scrap
  - Injected additives: amount & temperature
  - Metallurgical model prediction

- **Question:** Can we classify model failure modes, and dynamically recognize them?
  Can we dynamically compensate?
Process diagnostics
Basic idea: Mapper on fibres
Basic idea: Mapper on fibres

- Process model is a function [input data] → [model output]
- Given input data, we can find both measurement [T], prediction [T+ΔT] and prediction error [ΔT].
- Idea: study fibres of the map [input data] → ΔT × T
Basic idea: Mapper on fibres

- Process model is a function $[\text{input data}] \rightarrow [\text{model output}]$
- Given input data, we can find both measurement $[T]$, prediction $[T+\Delta T]$ and prediction error $[\Delta T]$.
- Idea: study fibres of the map $[\text{input data}] \rightarrow \Delta T \times T$
- Esp.: large values of $|\Delta T|$.
Master plan

- Flares and features in Mapper → classification of fibre shape
- Look for shape of input data over extreme values
- Find failure modes that can be recognized in production
- Test on future data!
The shape of steel
The shape of steel

- Mapper shape from Ayasdi Core
  Metric: Variance Normalized Euclidean
  Lenses:
  PCA1, PCA2, ΔT, T
The shape of steel

- Mapper shape from Ayasdi Core
  Metric: Variance Normalized Euclidean
  Lenses: PCA1, PCA2, $\Delta T$, $T$
- Core-generated auto-groups.
Find recognizable extremes

- Drop singleton auto groups
- Drop auto groups with any error less than 100°
- Compute global PCA, eyeball distribution of PCA1 among remaining auto groups
Find recognizable extremes

- Drop singleton auto groups
- Drop auto groups with any error less than 100º
- Compute global PCA, eyeball distribution of PCA1 among remaining auto groups
Find recognizable extremes

- Drop singleton auto groups
- Drop auto groups with any error less than 100º
- Compute global PCA, eyeball distribution of PCA1 among remaining auto groups
Candidate failure modes

- Auto-generated groups 17, 19, 21, 23, 32, 37.
- Most are very small. Group 23 bigger.
- **Question:** Can we detect membership in Group 23?
Candidate failure modes

- Auto-generated groups 17, 19, 21, 23, 32, 37.
- Most are very small. Group 23 bigger.
- **Question:** Can we detect membership in Group 23?
Membership detection, 23

- Logistic regression, class-balanced sampling
- A lot of false positives.
- Probably improves with better classifiers.
- Certainly improves with more data.
- Already somewhat useful results.
Why not just do naive regression?

- Far fewer false positives
Where do we go from here?

- Collect more data
- Validate these classifiers
- Test other possibly better classifiers

- Analyze the sound of the furnace: frequency spectra correspond to smelting stages
  use Mapper to find recognizable smelting modes?
Thank you for listening