Problem: Plant-wide disturbances in chemical plants act on many process variables

The disturbances affect cost, quality and safety of a process.

Research question: Find the root cause of the disturbance through cause & effect

The knowledge of all cause and effect relationships gives the fault propagation path.

Problem has been tackled by finding qualitative models of the direction of flow

The models are retrieved from expert knowledge or differential equations which are not always available.

Several effects can change the process dynamics of a disturbance along the flow

The causality measure should exploit these changes.
**Agenda**

Motivation

Method

Case Study

Conclusion

---

**Methodology:** Find nearest neighbours of past values of y to estimate a future value of x

(Feldmann and Bhattacharya, 2004)

![Diagram](image)

Does X predict Y better than Y predicts X?

---

**Methodology:** Construct embedded vectors and their images

![Diagram](image)

Parameters are embedding dimension m, time lag τ and prediction horizon h.

---

**Nearest neighbours predictability improvement (PredI) algorithm (Step 1)**

1. Finds K nearest neighbours \(x_{nn,i}\) of \(x_i\) through:
   \[ \|x_i - x_k\| \]

2. Embedd vector: \(X\)
   Nearest neighbour: \(Y\)

---

**Nearest neighbours predictability improvement (PredI) algorithm (Step 2)**

2. Compares the corresponding future value of \(y_{i+h}\) with future value of nearest neighbours \(y_{nn,i+h}\):
   \[ \rho(y|x) = \frac{1}{K} \sum_{k=1}^{K} |y_{i+h} - y_{nn,i+h}| \]

3. Embedd vector: \(X\)
   Nearest neighbour: \(Y\)

---

**Nearest neighbours predictability improvement (PredI) algorithm (Step 3, 4)**

3. Repeats for all \(i\) embedded vectors and computes average statistic:
   \[ \rho(y|x) = \frac{1}{L} \sum_{i=1}^{L} \rho(y|x) \]

4. Computes same measure with exchanged x and y and compares the two measures:
   \[ \rho(x|y) = \rho(y|x) - \rho(x|y) \]
Nearest neighbours predictability improvement (PredI) algorithm

Step 1: Finds $K$ nearest neighbours $x_{nn,i}$ of $x_i$ through minimising: $||x_i - x_k||$

Step 2: Compares the corresponding future value of $y_{i+h}$ with future value of nearest neighbours $y_{nn,i+h}$

Step 3: Repeats for all embedded vectors and computes average statistic

Step 4: Computes same measure with exchanged $x$ and $y$ and compares the two measures

Case study: Reaction process at Eastman Chemical Company, Tennessee

Most time trends of the process variables show a disturbance with an oscillatory nature.

Case study: Does the level upset the temperatures or the temperatures the level?

None of the control loop pv/op plots point towards a valve problem.

Case study: Digraph construction from causality measure is carried out in an automated way

The results indicate that the disturbance is caused upstream and not by the level control loop.
Agenda

Motivation
Method
Case Study
Conclusion

Benefits of cause & effect from historical process data

- Non invasive
- Easy to implement
- No expert knowledge or differential equations required
- Process data as from data historians such as PI system can be used
- Drill down tool

Process data can be evaluated with little additional cost from data historian.

Conclusions: Nearest neighbours PredI algorithm

- The PredI algorithm identifies causality and hence directions of fault propagation;
- A causal map can be constructed from the causality measure in an automated way;
- The causal map gives conclusions towards the root cause.