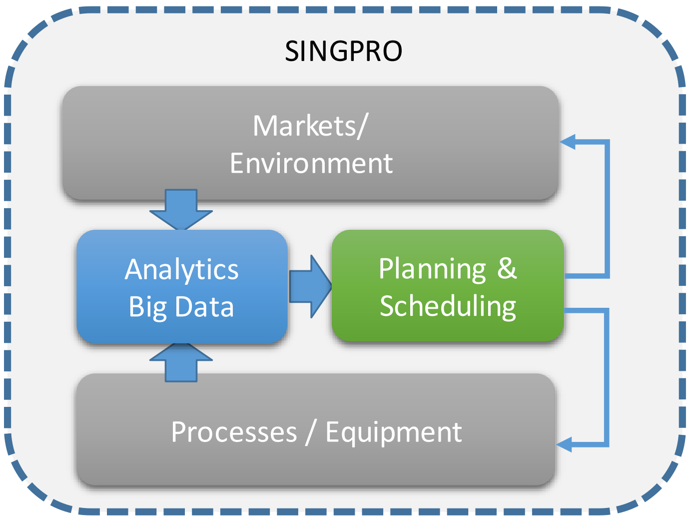
Aalto University
School of Chemical Engineering

SINGPRO

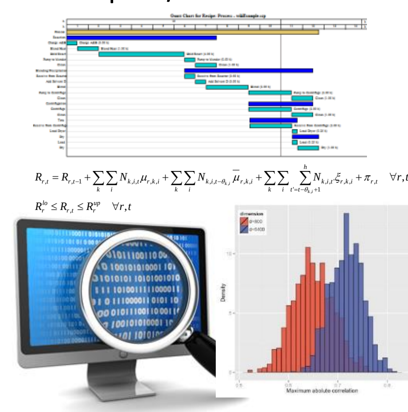
Iiro Harjunkoski (School of Chemical Engineering)
Keijo Heljanko (School of Science)
Seminar, 15.10.2018

SINGPRO Project (2018-2019)
Synergistic and intelligent process optimization

Academy of Finland project: Adj. Prof. Harjunkoski (Aalto CHEM) & Assoc. Prof. Heljanko (Aalto SCI)



Robust & agile schedules
Adapted/correlated models



Sustainable & safe operations

- ✓ Energy efficient
- ✓ Optimal throughput
- ✓ Well maintained in time
- ✓ Safe operating conditions
- ✓ On-time production
- ✓ Knowledge-based models
- ✓ Agile and adaptive decisions

SINGPRO Targets
Create and prove novel concepts in real life

- Show that big data technologies can be deployed together with optimization strategies, to close the decision loop in automation
  - The results can help defining future research needs within systems-level integration of process control systems and data-driven decision making
- Collaborate with Finnish industry on piloting the methodology
  - Get access to real data, process information and the opportunity to discuss, test and demonstrate the solution approaches in practice
  - Create concepts that are re-usable across various industries

SINGPRO Project Team

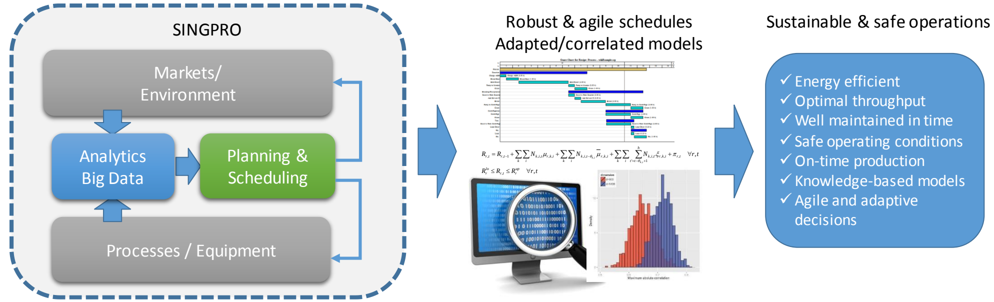
Adj. Prof. Harjunkoski (Aalto CHEM)

Prof. Heljanko (University of Helsinki)

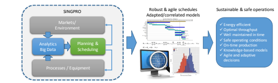
Dr. Tewodros Deneke

Dr. Teemu Ikonen

Dr. Hossein Mostafaei

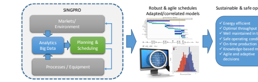


Questions to be Answered (1/2)



- **Often a production plan is already ”old” soon after being rolled out to the plant floor**
  - Could I do better planning by knowing more about the process, i.e. utilizing the real-time data?
- **Schedules are usually based on average durations (tables)**
  - Dynamically generating accurate statistics on process behaviour every time I want to schedule?
- **Disturbances and breakdowns often come as a surprise**
  - How many incidents can actually be predicted and avoided?

Questions to be Answered (2/2)

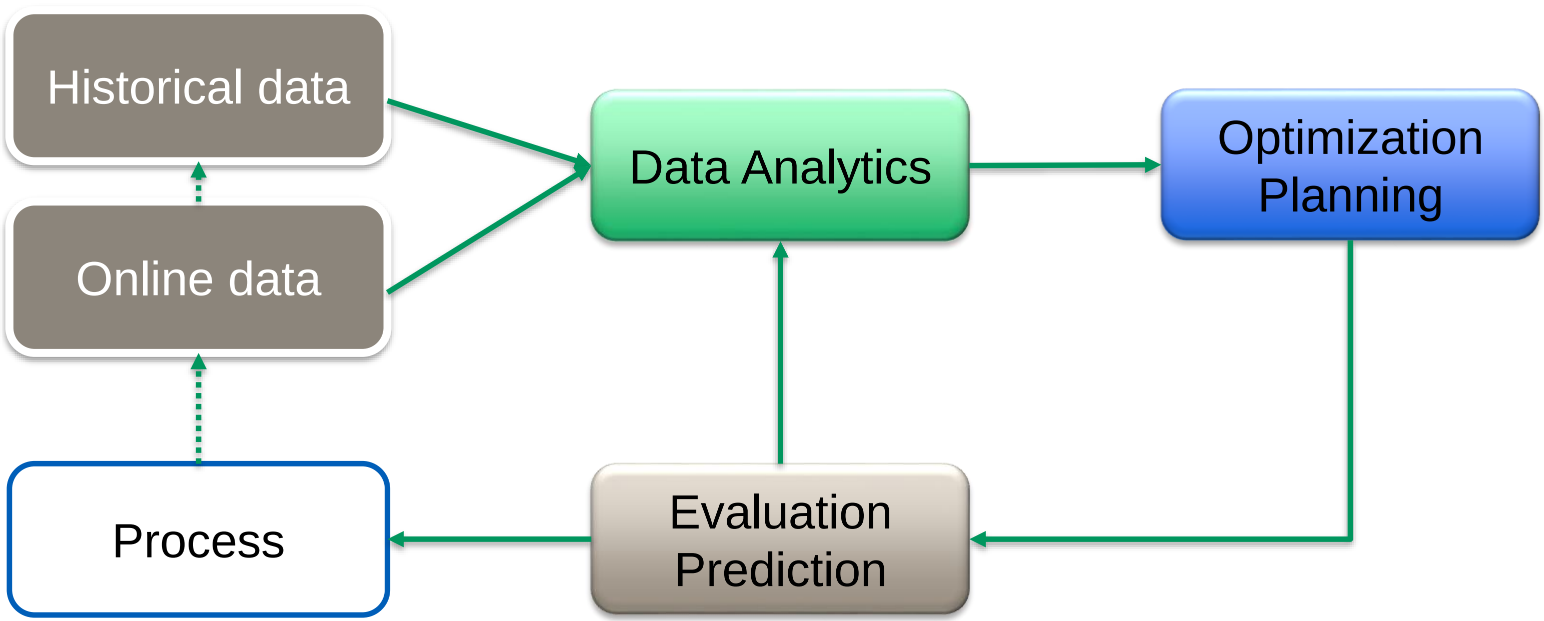


- **Often we focus on the most obvious data assuming simple causality**
  - What information actually is relevant for root-cause analysis? Are there hidden relationships?
- **Many decisions in optimization add to the complexity**
  - Are there decisions that can be excluded from the optimization scope?
- **Data is mostly collected and stored only for troubleshooting**
  - What is the actual value of this data?

SINGPRO Highlights
Combine big data analytics with optimization

- Online, reactive and anticipative tools for sustainable and efficient operations
- Collaboration interfaces between scheduling optimization and big data analytics / machine learning resulting in more agile, self-aware and flexible decisions
- Combine first-principle models with machine learning in an efficient way to reduce the modeling complexity and efforts
- Create - in a fully data driven fashion - models of normal process behaviour and predictive models of process disruptions

SINGPRO Methodology
Loop: Process → Analytics → Optimization



SINGPRO Research Activities
Combine big data analytics with optimization

- Analyze process data across multiple domains using clustering, pattern matching, identification of causalities
- Create open and adjustable production scheduling models (discrete and continuous-time) and solution concepts for large-scale problems
- Run selected pilot case studies using industrial data on both production and supply chain level
- Build a cloud-based demonstrator built on an industrial platform following generic standards validated on multiple test cases

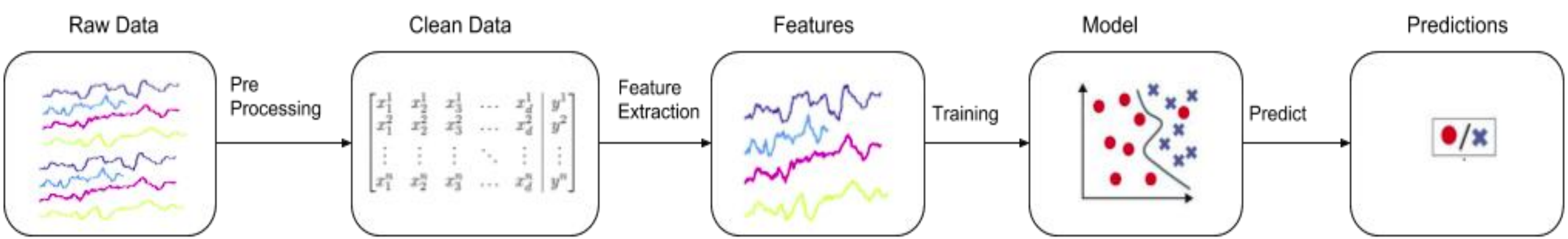
Summary of Ongoing Research

Machine Learning
Dr. Tewodros Deneke

- Starts with a raw data
- Data preprocessing
- Feature extraction
- Model training
- Prediction

Potential applications:

- Predictive maintenance
- Anomaly detection
- Parameter prediction
- AI planning
- Etc …



Parameter Predictions in Scheduling
Dr. Teemu Ikonen

Research aims

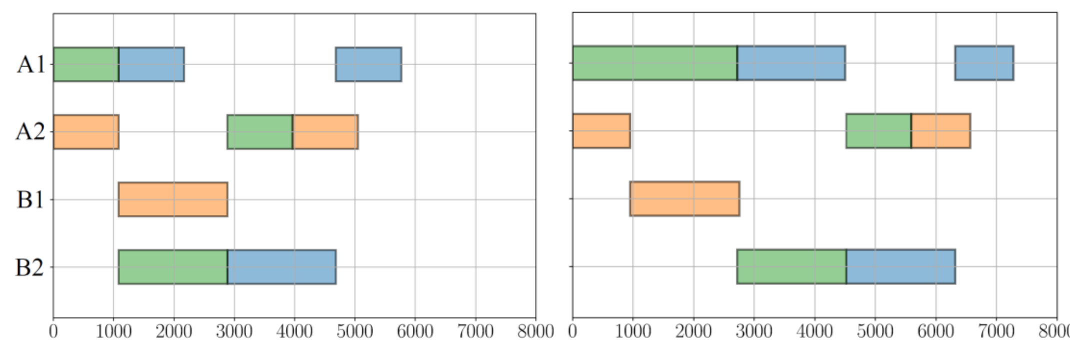
- Improve the quality of scheduling solutions via machine learning based (scheduling) parameter predictions
- Investigate the relationship between scheduled and realized schedules on real datasets

Primary machine learning methods

- Gaussian process regression
- Random forest regression

Scheduling models

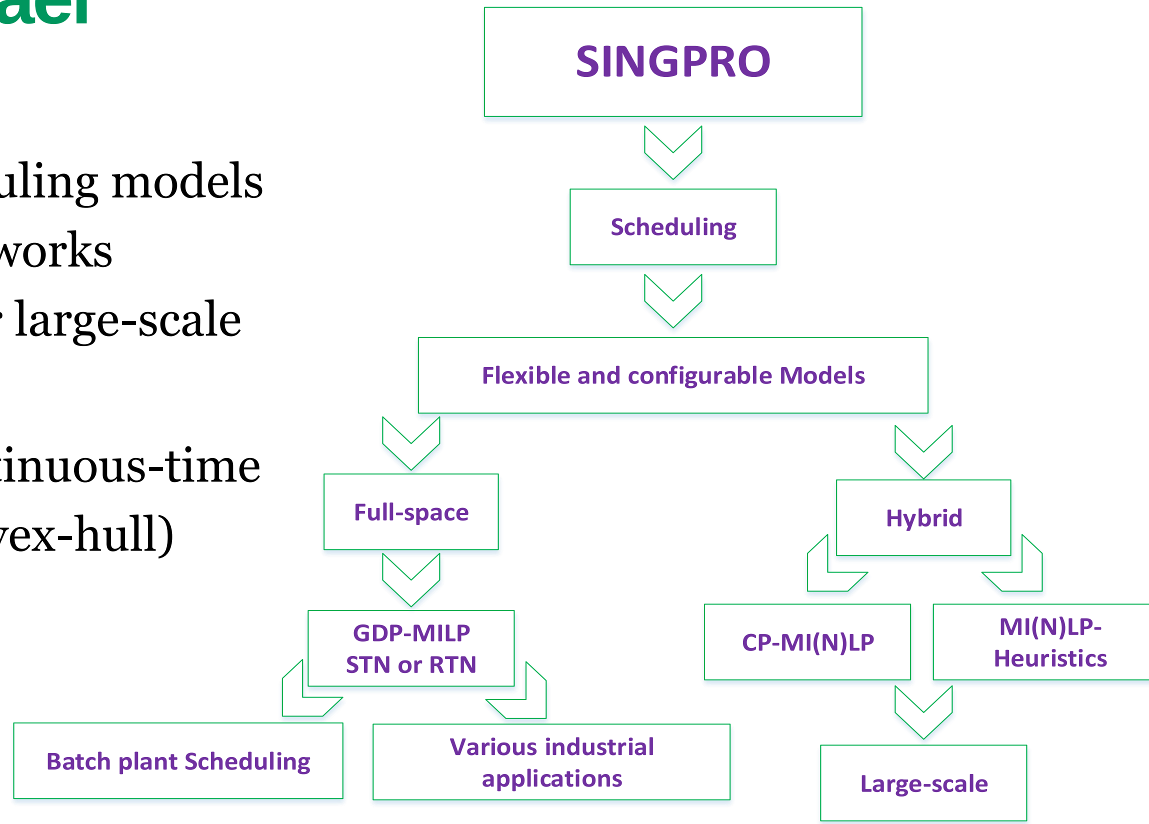
- Mainly continuous-time representations



Planning and Scheduling Optimization
Dr. Hossein Mostafaei

Aims and targets
- Modular and flexible scheduling models
- Hybrid optimization frameworks
- Decomposition schemes for large-scale

Scheduling models
- Based on discrete- and continuous-time
- Based on GDP (mainly convex-hull)
- Based on STN and RTN



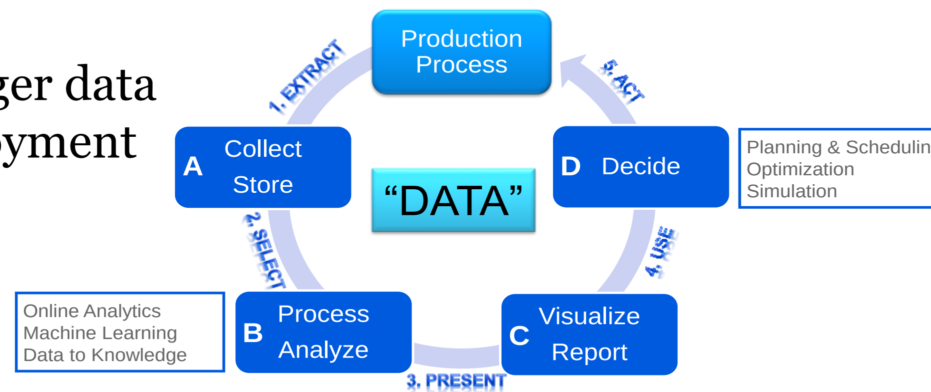
SINGPRO Collaboration

Scientific collaboration with world-leading academics: Carnegie Mellon University, University of Texas at Austin, University of Lisbon and Aalto.

- Combine cross-domains already in the research phase boosting out-of-the-box thinking and enabling a larger pool of methodologies and synergies of the existing research.

Industrial collaboration

- Identify partners that can provide larger data pools but also support technical deployment e.g. through platforms
- Define joint metrics for improvement



Machine Learning for Predictive Maintenance

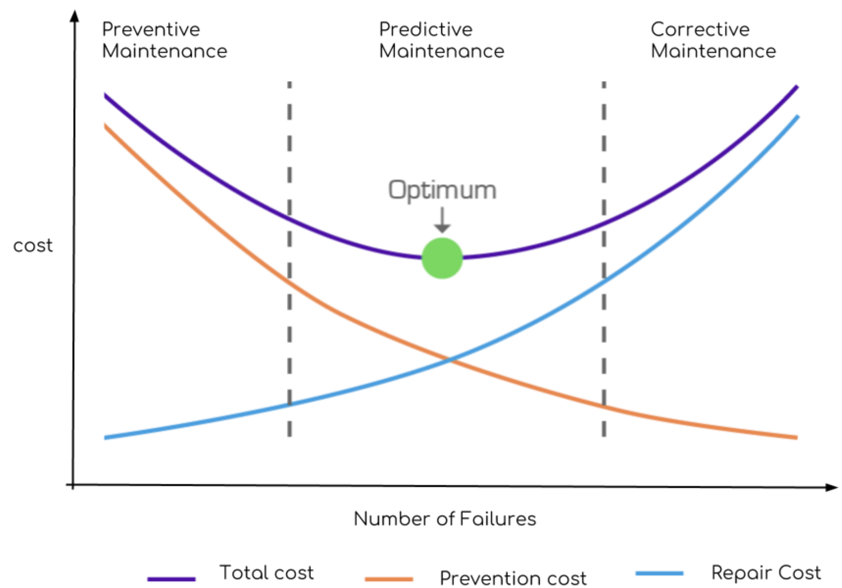
Tewodros Deneke

Outline

- Overview & Research Problems
- Data Collection & Preparation
- Modeling & Prediction
- Future Directions

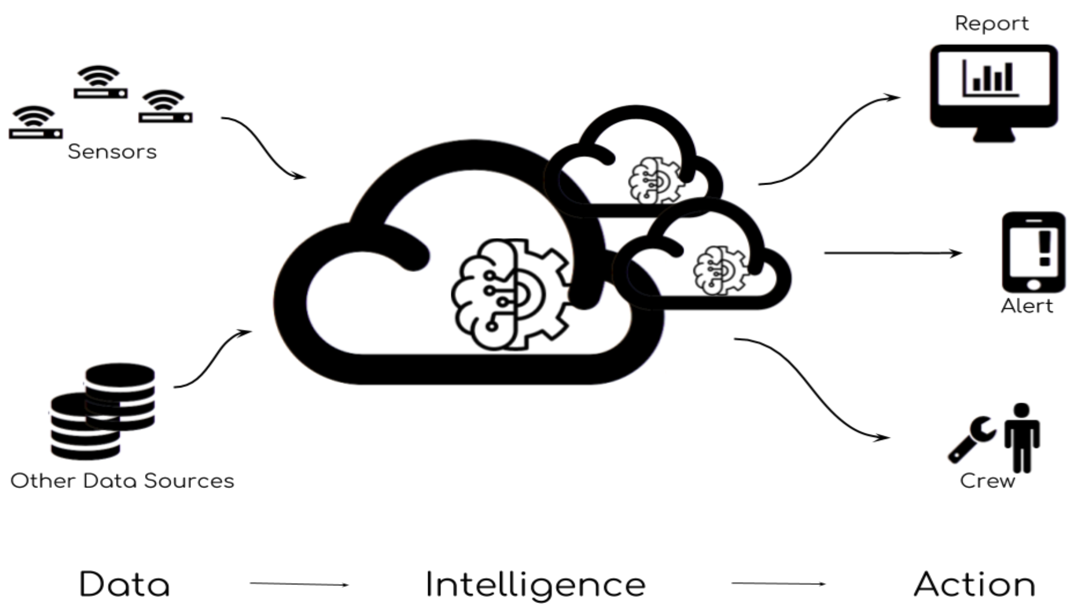
Overview & Research Problems

- Do we need maintenance?



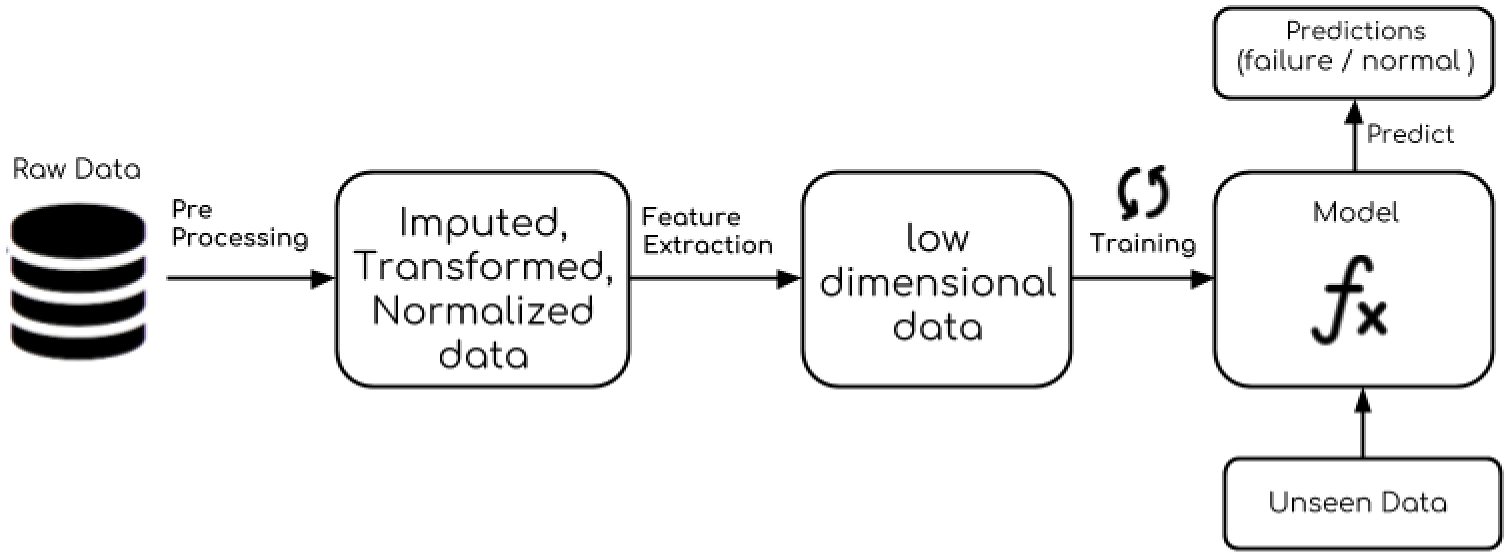
Overview & Research Problems

- Overview and promise of predictive maintenance



Overview & Research Problems

- Machine Learning: learning from data



Data Collection

- To build a failure model, we need enough historical data
- Not all data is suitable for modeling
- What failures occur and which ones to model
- How parts of the system related to failures
- What can we measure
- Ideally should be done in collaboration
- Ideally data should be labeled

ts	sensor 1	sensor 2	sensor 3	...	...	label(status)
[2018-06-30 02:20...	2.0781384267305	0.0	0.0	0.0	0.0	Normal
[2018-06-30 22:15...	8.1508338325905	0.0	0.0	0.0	0.0	Normal
[2018-06-30 22:20...	0.0	0.0	0.0	0.0	21.25	Break
[2018-06-30 22:40...	0.0	0.0	0.0	0.0	0.0	Break
[2018-07-01 00:00...	0.0	2.0781384267305074	5.851885143079255	12.802269182707134	0.0	Reduced
[2018-07-01 00:05...	0.0	2.0541220825987976	4.988659846318233	17.416142154049563	0.0	Reduced
[2018-07-01 00:10...	0.0	2.0831577652379085	6.695523312217311	13.68337741651033	0.0	Reduced
[2018-07-01 00:15...	0.0	2.1082291935765465	5.8017552841541375	14.134423278098883	0.0	Normal
[2018-07-01 00:20...	0.0	2.152975468408494	5.7489272980462935	15.922789619082497	3.0357142857142856	Normal

Data Preparation

- Not all data is suitable for modeling
- Data might need to be transformed
- Normalization is often needed
- Feature selection / dimensionality reduction
- Data labeling in some cases

singpro

Modeling & Prediction

- The problem can be framed differently based on
  - Kind of output expected (prediction / classification)
  - Weather data is labeled or not (supervised / unsupervised)
  - The performance targets that the model should be optimized for?
- The choice of the learning algorithm is further influenced by
  - The proportion of the labels
  - What are the performance targets

Remaining Useful Life (RUL)

- How many time units are left before the system fails?
- Needs labeled data
- Labels are numeric
- Skewed scoring and cost function
- Model needs to capture temporal pattern

Failure Classification

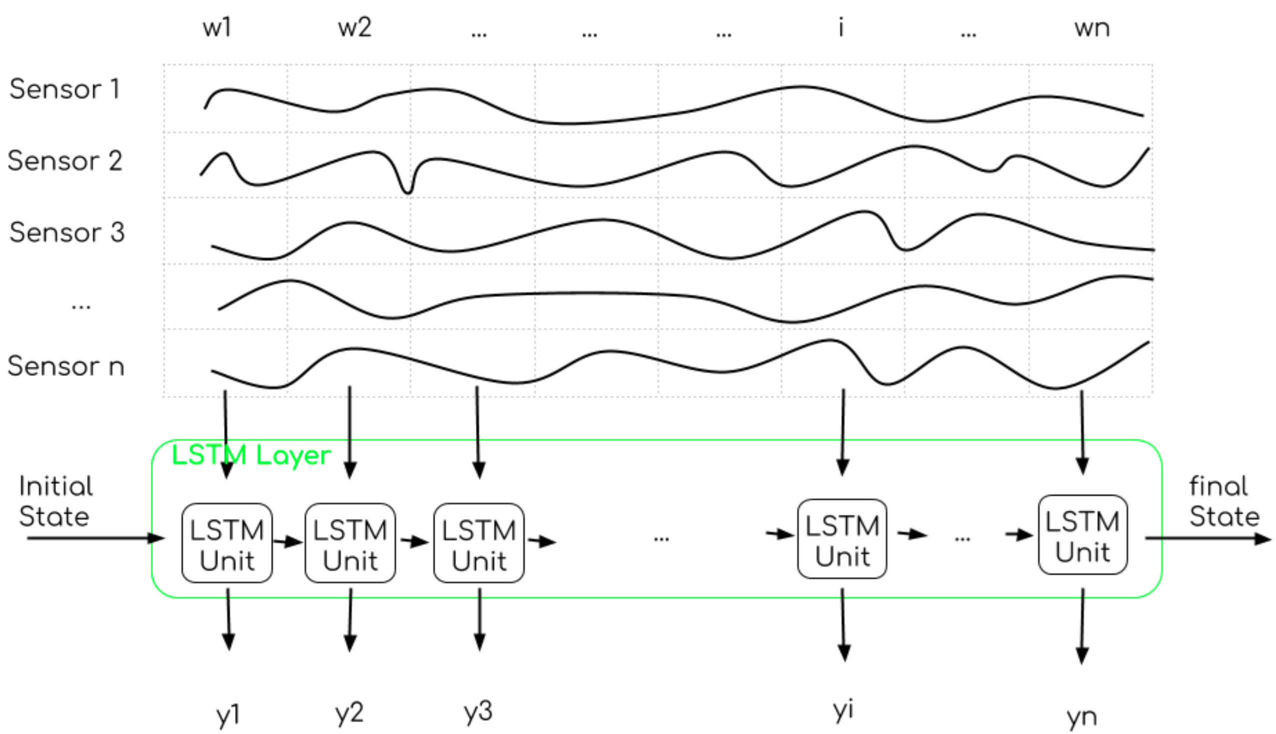
- Will a machine fail in the next N time units?
- Needs labeled data
- Labels are categorical or boolean
- Could be multi-class (Normal / Reduced / Failure)

Anomaly Detection & Labeling

- Is the behaviour shown normal?
- Label unavailable or too few
- Can be used to create some sort of labeling
- Reconstruct observations, monitor residuals
- Cluster and detect outliers

singpro

LSTM Network



Challenges & Future Directions

- Missing value imputation
- Large scale model training
- Exploring anomaly detection approaches

singpro

Aalto University
School of Chemical
Engineering

Parameter prediction and realization in job shop scheduling

A Gaussian process approach

Teemu Ikonen and Iiro Harjunkoski
15th October 2018
SINGPRO seminar
Aalto University

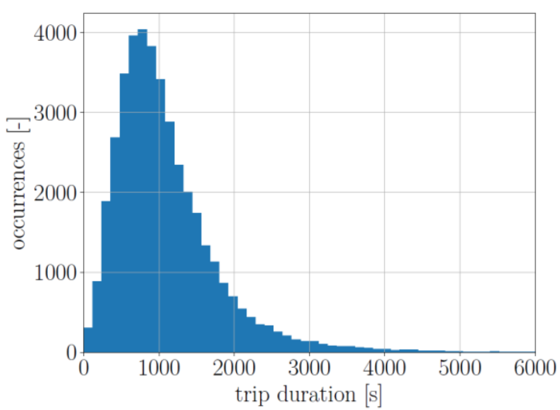
Outline*

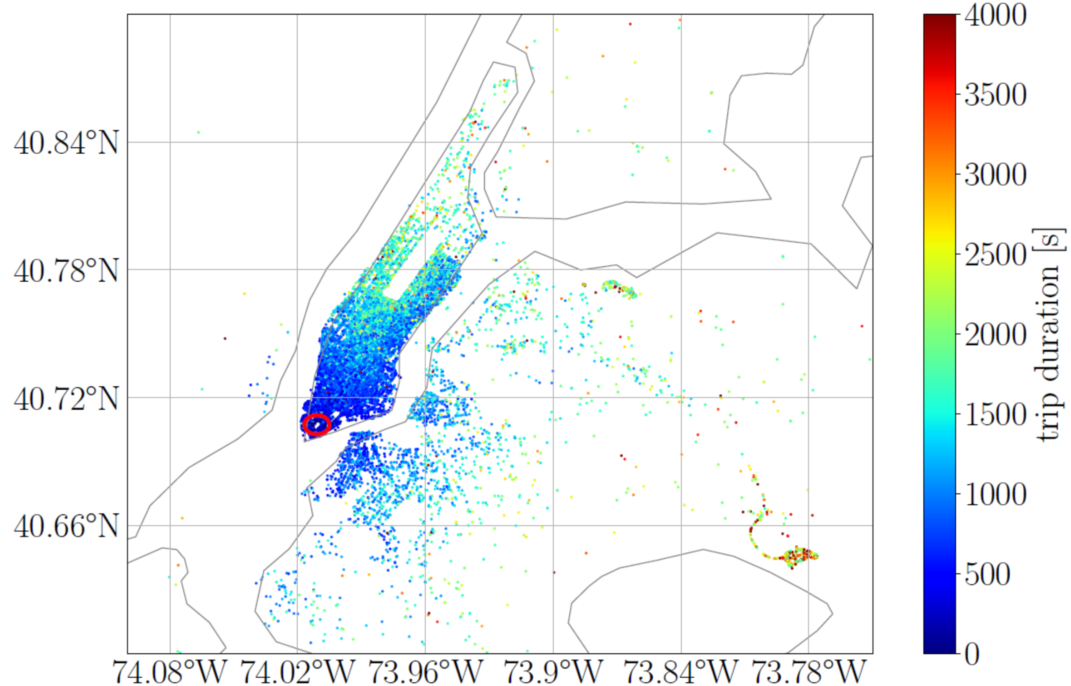
- **New York City taxi duration dataset**
- **Scheduling problem**
- **Methods**
  - Prediction models
  - Scheduling model
- **Results**
- **Conclusions and future work**

*This work will also be presented in the AIChE annual meeting 2018 later this year.

New York City (NYC) taxi duration dataset

- **Openly available dataset at https://www.kaggle.com/c/nyc-taxi-trip-duration**
- **Data of over 1.4 million taxi tips in NYC**
- **The ground truth is the trip duration**
- **Examples of features**
  - Passenger count, pick-up and drop-off date, time and coordinates





Scheduling problem

- **A company, head quartered at Wall street, performs surveys at remote sites in NYC**
- **Objective: minimize the make span of performing surveying trips at six different sites (Fig. 1)**
- **Surveying trip consists of outbound taxi trip, survey and inbound taxi trip (Fig. 2)**
- **Constraints:**
  - Only two taxi trips can be performed at the same time
  - Only two surveys can be performed at the same time
- **Durations of taxi trips are predicted, the surveys have a fixed duration of 1800 s**

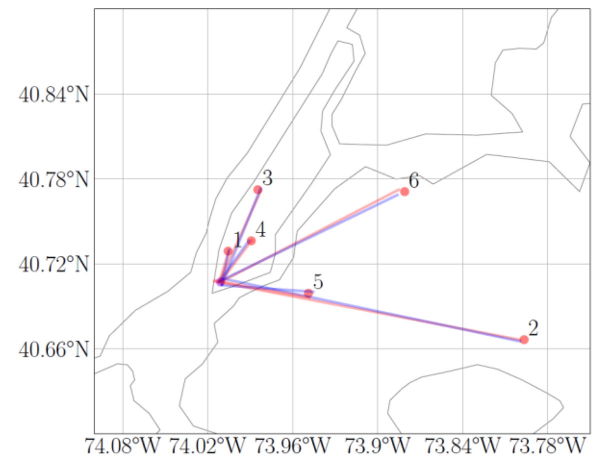


Fig. 1: An example set of surveying sites

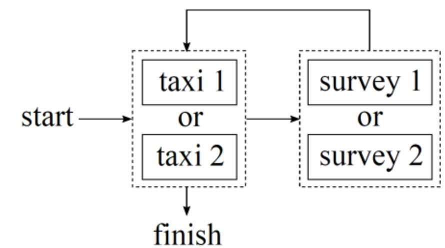


Fig. 2: Procedure of performing a surveying trip

Scheduling problem: industrial correspondence

Scheduling problem	Corresponding industrial examples
Outbound taxi trip	Preparation, heating
Inbound taxi trip	Cleaning, cooling
Survey	Chemical reaction, mechanical operation

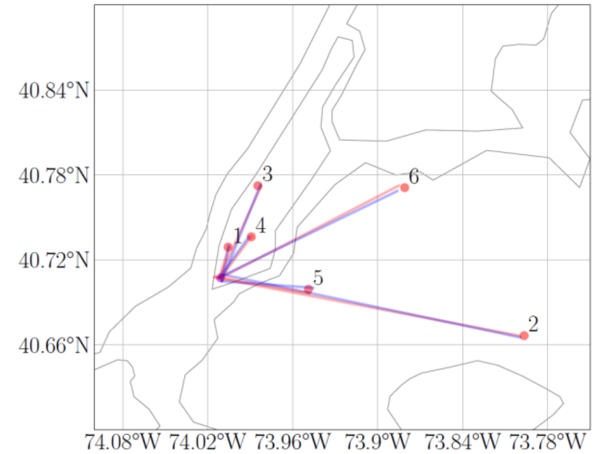


Fig. 1: An example set of surveying sites

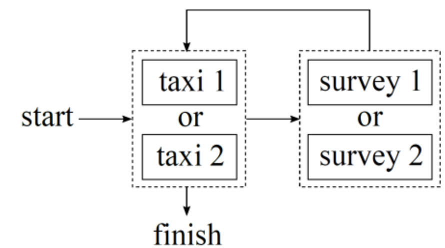
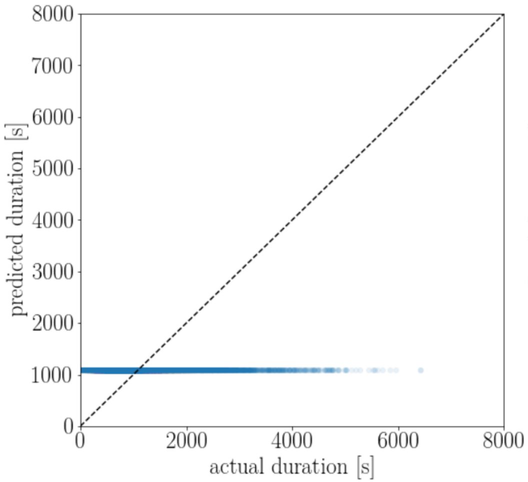


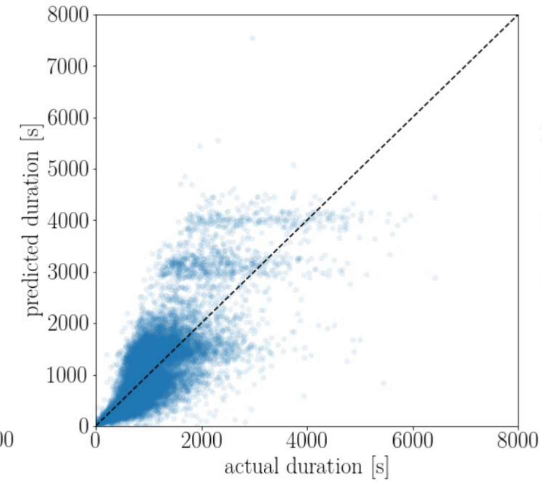
Fig. 2: Procedure of performing a surveying trip

Prediction models

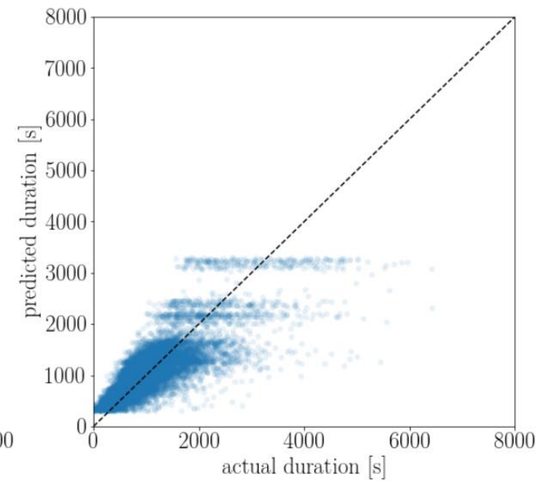
Prediction models, starting from the lowest fidelity



1) Average model



2) Rate model



3) Gaussian process (GP) model

prediction model	RMSE	r^2
average	701.8	-2.37e-4
rate	556.5	0.371
GP	438.4	0.610

RMSE = the root mean square error
r^2 = the coefficient of determination

Prediction models: industrial correspondence

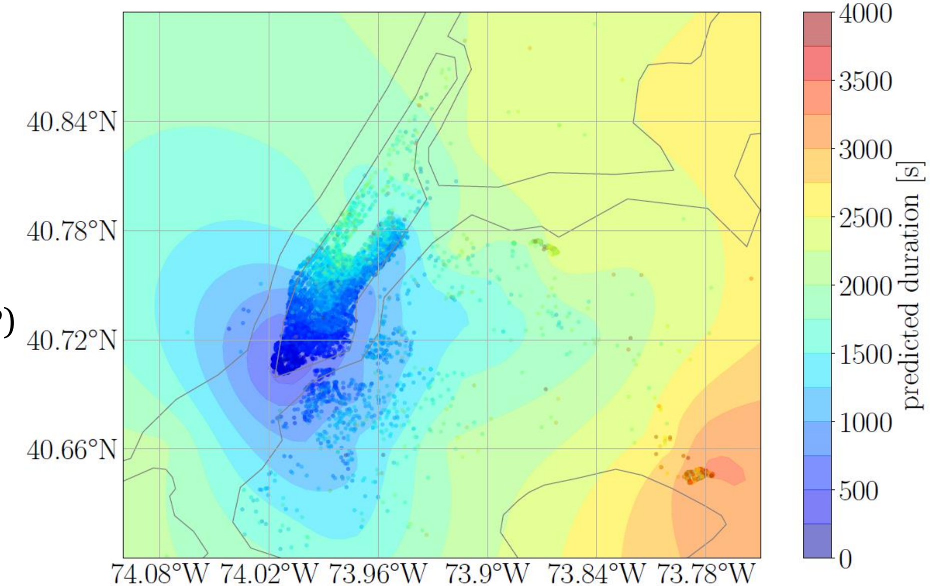
Prediction model	Industrial correspondence
Average	A static table value, which is determined as an average of historical values
Rate	A model which considers the task to have a constant processing rate (e.g. heating or purification of a volume of material)

Prediction models: Gaussian process regression

- **Kernel**
  - Exponential kernel

$$\kappa(x, x') = \sigma_f^2 \exp\left(-\frac{\|x - x'\|}{2l^2}\right)$$

  - Noise term
- **Choosing the hyperparameters**
  - Tuned using the maximum a posteriori (MAP) estimate



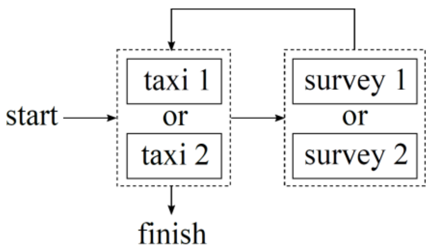
Scheduling model

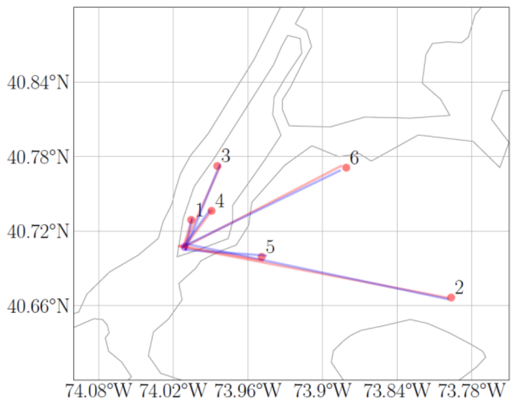
- **Unit-specific continuous-time scheduling model by Shaik and Floudas (2009)**
- **We use the value $\Delta n = 1$ (which defines the number of event points a job can span over)**
- **The number of total event points is determined by iteratively increasing n until the model has a feasible solution**

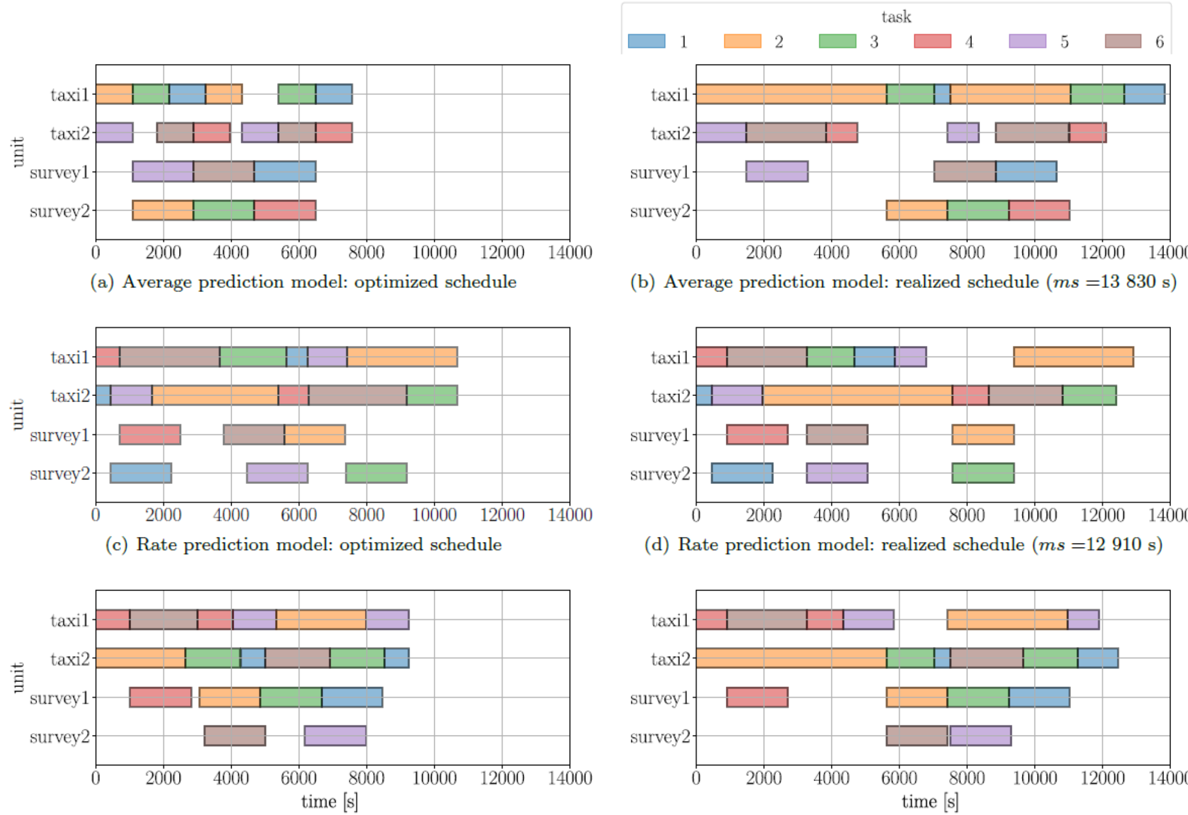
Shaik, M.A. and Floudas, C.A. (2009). Novel unified modeling approach for short-term scheduling. *Industrial & Engineering Chemistry Research*, 48(6), 2947-2964.

Results: a single scheduling problem

Optimized and realized schedules of six tasks, consisting of outbound taxi, survey, and inbound taxi







(a) Average prediction model: optimized schedule

(b) Average prediction model: realized schedule (ms =13 830 s)

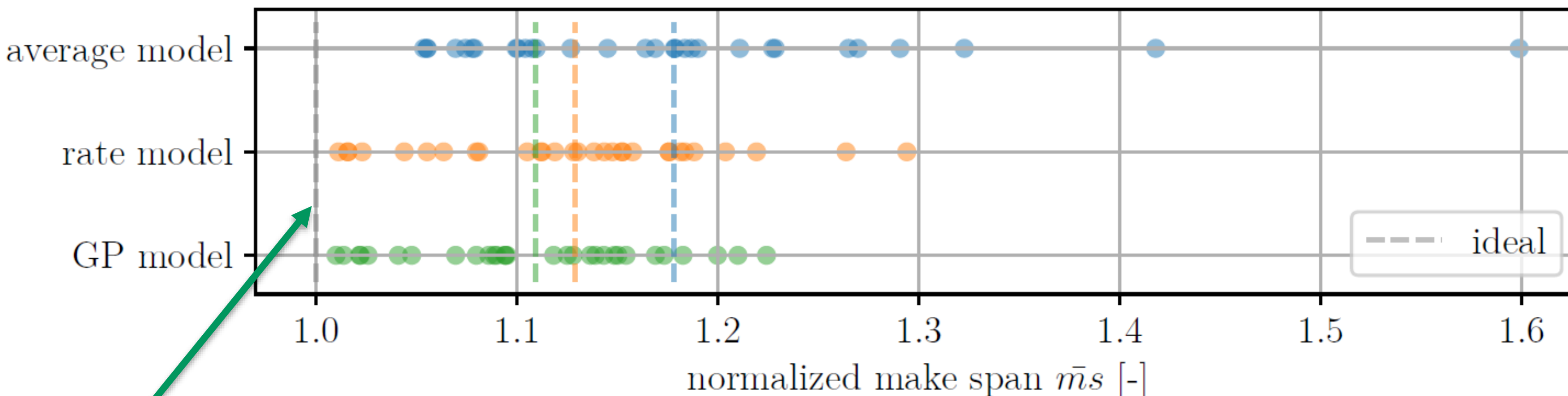
(c) Rate prediction model: optimized schedule

(d) Rate prediction model: realized schedule (ms =12 910 s)

(e) GP prediction model: optimized schedule

(f) GP prediction model: realized schedule (ms =12 450 s)

Results: 30 different scheduling problems



The reference model is the *ideal predictor*, for which $r^2 = 1$ and RMSE $= 0$.

Aalto University
School of Chemical Engineering

Results: 30 different scheduling problems

prediction model	$\bar{ms}$ [-]		solution time [s]		
	mean	sd	mean	min	max
average	1.178	0.119	4.45	4.25	5.74
rate	1.129	0.072	45.94	7.71	121.69
GP	1.109	0.061	79.66	16.44	344.82

Conclusions

- **Job shop scheduling with data-driven duration prediction of three levels of fidelity is studied**

Conclusions

- **Job shop scheduling with data-driven duration prediction of three levels of fidelity is studied**
- **In the studied problem, the GP prediction model yields shorter make spans than average and rate prediction models by the margins of 5.8% and 1.8%, respectively**

Conclusions

- Job shop scheduling with data-driven duration prediction of three levels of fidelity is studied
- In the studied problem, the GP prediction model yields shorter make spans than average and rate prediction models by the margins of 5.8% and 1.8%, respectively
- However, the computational cost of GP and rate prediction models are an order of magnitude higher than that of the average prediction model

Future work

- **Proactive scheduling (e.g. adjustable robust optimization) with the uncertainty predictions from Gaussian process regression**
- **More complex scheduling problems**

Aalto University
School of Chemical
Engineering

SINGPRO

Iiro Harjunkoski (School of Chemical Engineering)
Keijo Heljanko (School of Science)