Fault detection and diagnosis on large scale

implementation in the PULSE Core living lab

The project received funding from Dutch Ministry of Economic Affairs and Climate Policy and Ministry of the Interior and Kingdom Relations under the MOOI program

1

Introduction

- Coen Hoogervorst
  - SPIE Building Solutions
  - Consultant at PULSE department

- Olav Vijlbrief
  - TNO Building Physics and Systems
  - Scientist Integrator & Data Scientist

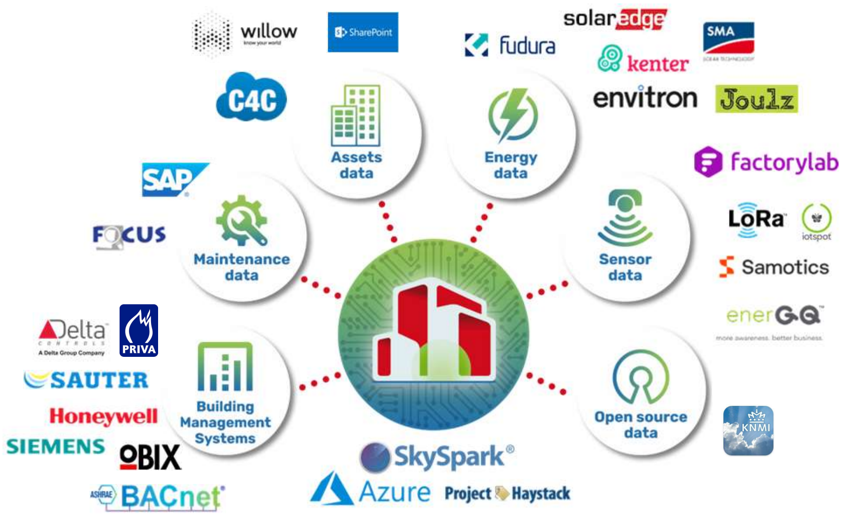
2

Contents

Introduction on Living Lab: PULSE Core

- Mission
  - Creating the most sustainable, efficient & healthy working and living environments for today and future generations using data.
- Vision
  - We are the heart of future-proof buildings and their users with data as fuel for the sustainability and efficiency of healthy buildings. In doing so, we add value for the owner and the users
- - BUILDING VALUE -

Introduction on Living Lab: PULSE Core



Introduction on Living Lab: PULSE Core

PULSE CORE as the basis for our asset management landscape:

- Data driven asset management
- Continuous commissioning
- Condition based maintenance

3. SUMMIT	Performance contract	Performance guarantee
2. SMART	Hour effort contract	Performance advice
1. START	"Software as a Service"	Performance insights

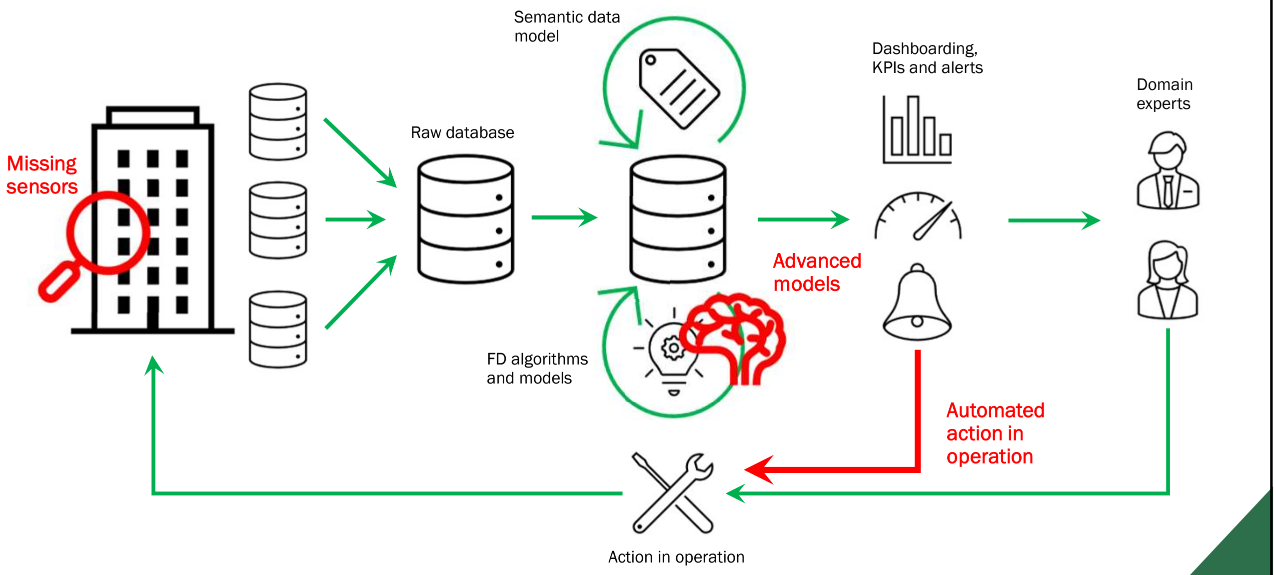
Introduction on Living Lab: PULSE Core

>900 Buildings connected	**>5 M** m^2
>250 Strategic customers	**>120 K** Sensors
280 K Ton CO_2	**360 K** BMS notifications / year
>60 B measure values	**>4500** Users

- Tilburg University
- Eindhoven airport
- Van Gogh Museum
- NIKHEF
- TNO
- Philips
- KLM (4AMS)
- Kromhout Kazerne
- Altrecht
- Carmel college
- High Tech Campus
- PGGM

7

The current situation and future ambition



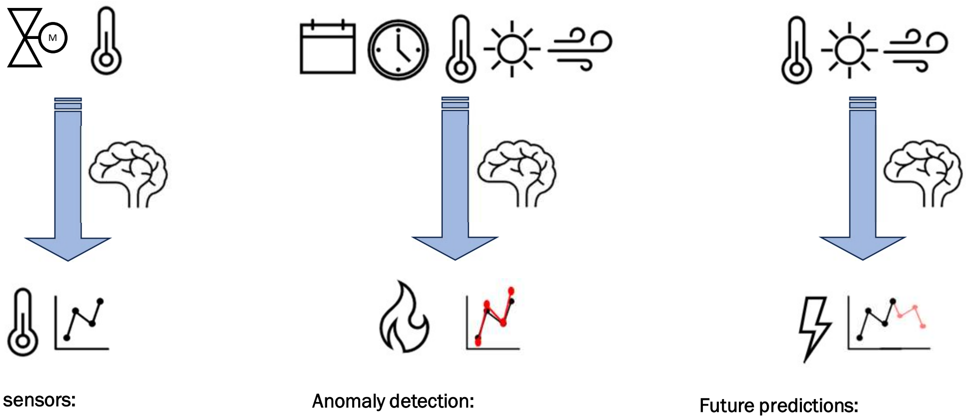
8

The current situation and future ambition

- Missing sensors:
  - Prediction missing sensor values using models
  - Enables more FD(D) analysis in sensor poor environments
  - Better fault prioritisation based on impact
- Advanced models
  - Anomaly detection on available sensors
  - Future predictions
- Automated action in operation
  - Expansion of current FD with additional symptoms to FDD
  - Enables automated integration into operational (maintenance) process
    - SPIE: Shorter time to action, lower costs and higher service level
    - Customer: Less disturbance, higher availability
  - One of the solutions to close the gap in shortage of skilled personnel

9

Virtual sensors and advanced symptoms



Missing sensors:

- Temperatures water distribution
- Energy flows
- Temperature and humidity in AHU

Anomaly detection:

- Energy consumption (profiles)
- Energy savings
- Component faults

Future predictions:

- Contract capacity exceedance
- Expected energy consumption profile

10

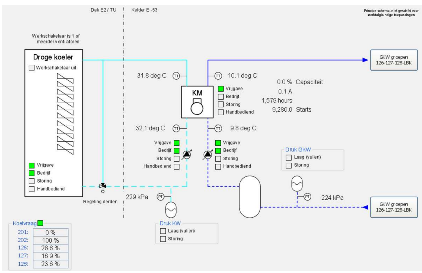
Virtual sensors and advanced symptoms

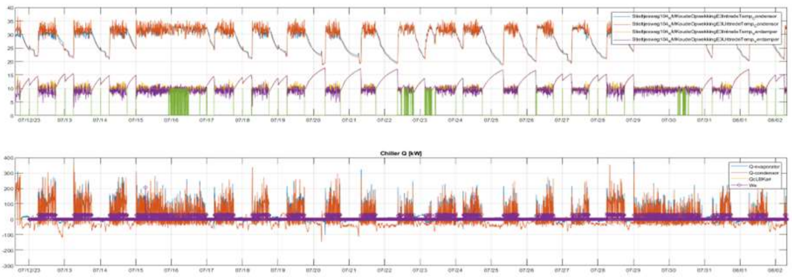
Model approaches investigated

- Missing sensors:
  - No data for prediction value
  - Physical models
- Anomaly detection:
  - Data for prediction value
  - Both physical and data driven approach
- Future predictions
  - High level predictions (on building level)
  - Data driven approach

Virtual sensors and advanced symptoms

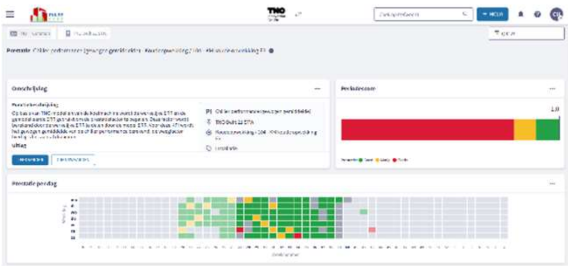
Missing sensors example: chiller cooling power





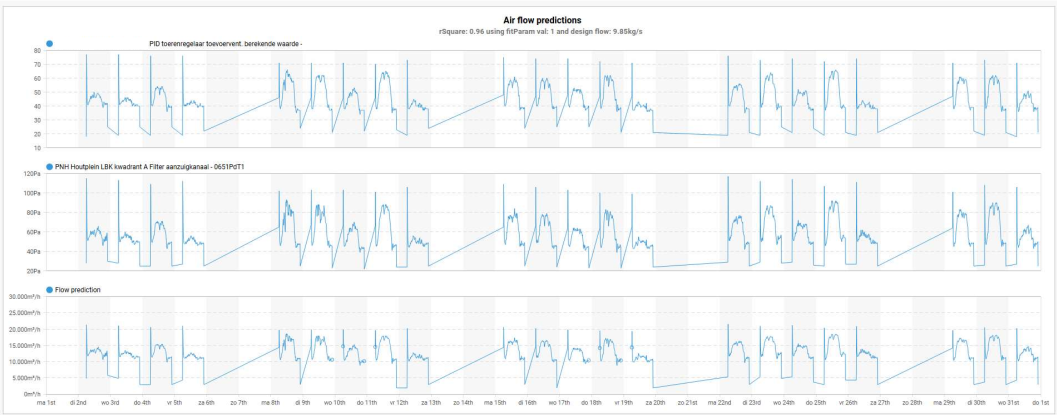
Virtual sensors and advanced symptoms

Missing sensors example: chiller cooling power



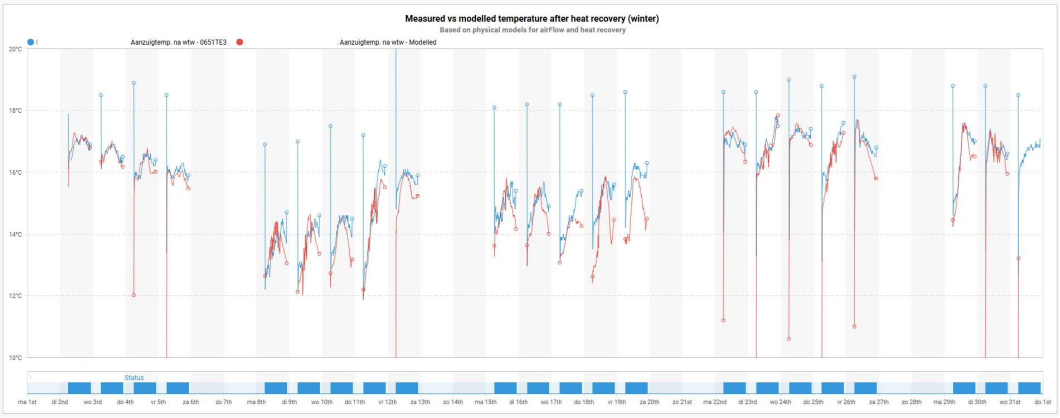
Virtual sensors and advanced symptoms

Missing sensors example: ahu



Virtual sensors and advanced symptoms

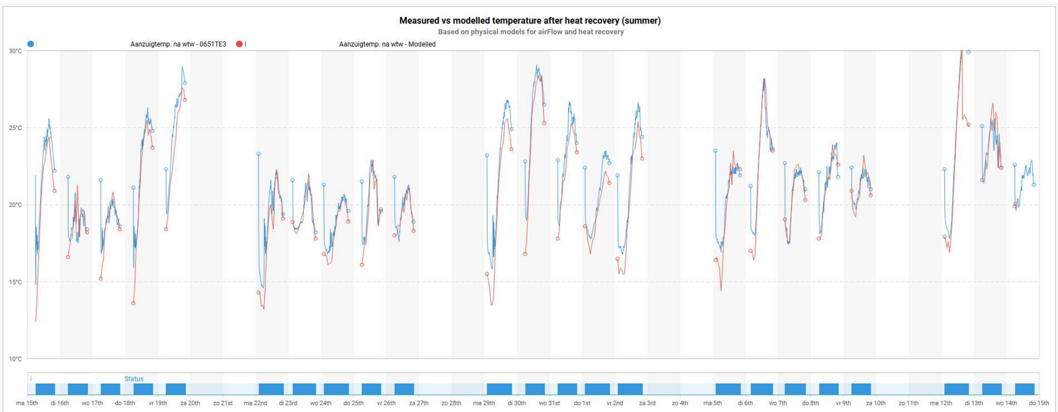
Missing sensors example: ahu



15

Virtual sensors and advanced symptoms

Missing sensors example: ahu



16

Virtual sensors and advanced symptoms

Missing sensors conclusions:

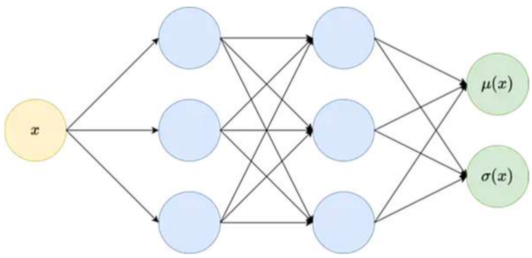
- Physical approach
- Storing design params in semantic data model makes approach replicable / scalable
- System level is hard because of different configurations and control strategies: advice to use (sub) component level
- On component level it is easy

Virtual sensors and advanced symptoms

Anomaly detection: AHU (neural network):

- Fitting uncertainty
- 2 outputs, mean, std
- Special loss function

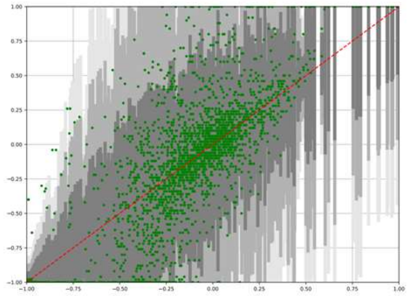
$$Loss = \log(\sigma^2) + ((y - \mu)^2) / \sigma^2$$

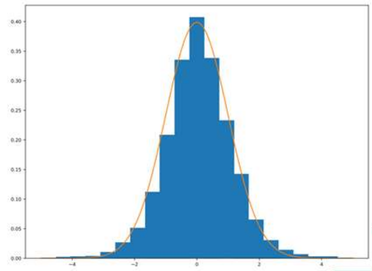


- What is really an anomaly.
- SHAP-values for better interpretation

Virtual sensors and advanced symptoms

Anomaly detection: AHU (neural network):



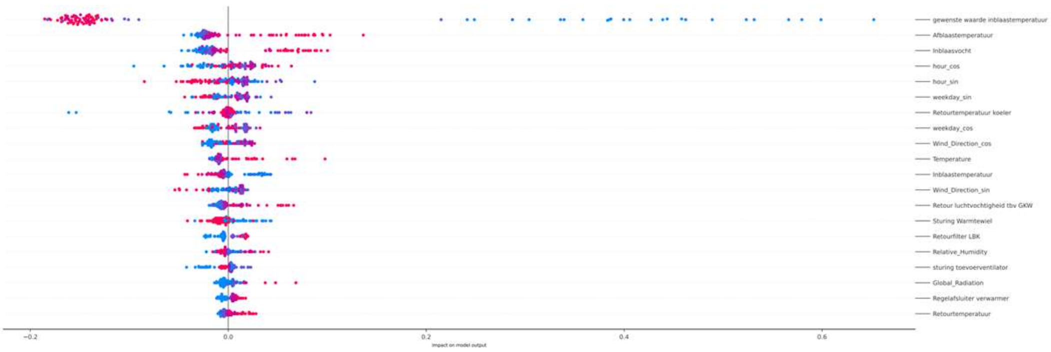


Training phase
- Step 1: selection of input features: 30 => 21
- Step 2: hypertuning neural network
- Step 3: training the network
- Step 4a: trainingsreport
- Step 4b: store selection and trained model

Virtual sensors and advanced symptoms

Anomaly detection: AHU (neural network):

Shap values
- Effect of each feature on the prediction
- Verify if those values are logical

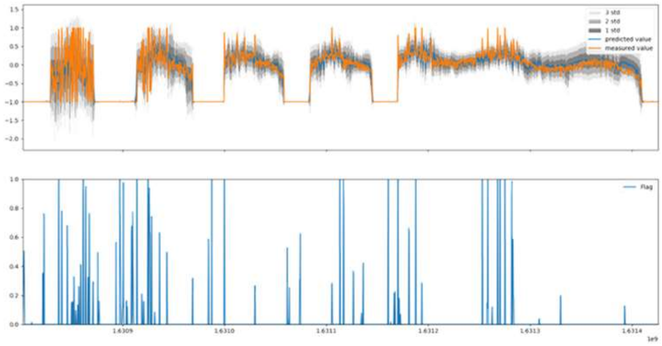


Virtual sensors and advanced symptoms

Anomaly detection: AHU (neural network):

Operational phase

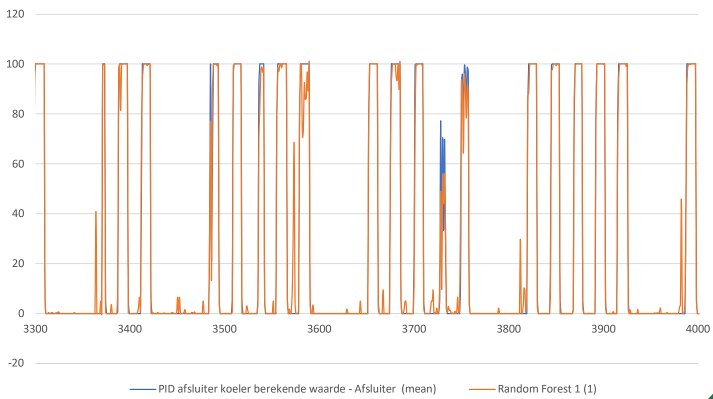
- Use Neural Network to predict mean and std.
- Measured value vs Predicted value
- Flag
  - 0 : values < 2std
  - 0-1 : values between 2 and 3std
  - 1 : values above 3std
- Averaged (running mean) => symptom



Virtual sensors and advanced symptoms

Anomaly detection: AHU cooling valve (ML)

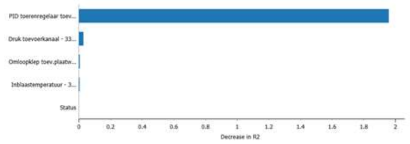
Model	Features	MSE	RMSE	MAE	R2
Model1 hourly average Random Forest	Status, office open, outdoor temperature, temperature difference between supply temperature and setpoint	101,3403	10,06679	2,326909	0,886476
Model1 hourly average kNN	Status, office open, outdoor temperature, temperature difference between supply temperature and setpoint	103,0621	10,15195	2,455018	0,885644
Model1 hourly average xgBoost	Status, office open, outdoor temperature, temperature difference between supply temperature and setpoint	111,9638	10,58129	2,406269	0,874827
Model1 hourly average AdaBoost	Status, office open, outdoor temperature, temperature difference between supply temperature and setpoint	126,9526	11,26732	2,206104	0,859297
Model1 hourly average ANN	Status, office open, outdoor temperature, temperature difference between supply temperature and setpoint	97,33293	9,865745	3,27905	0,895568

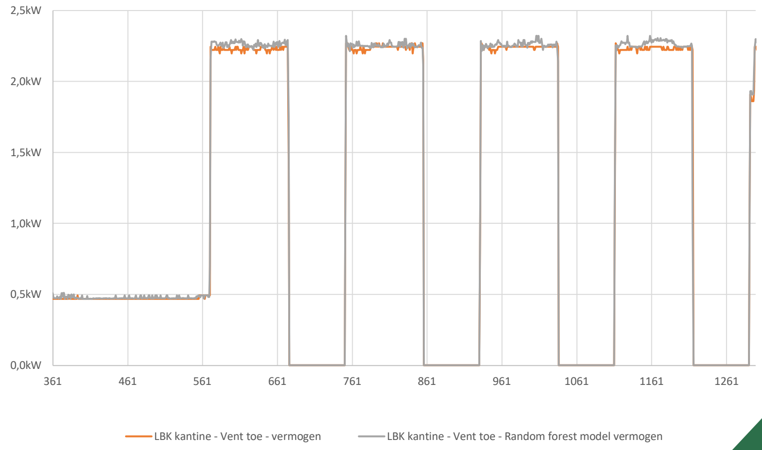


Virtual sensors and advanced symptoms

Anomaly detection: AHU fan power (ML)

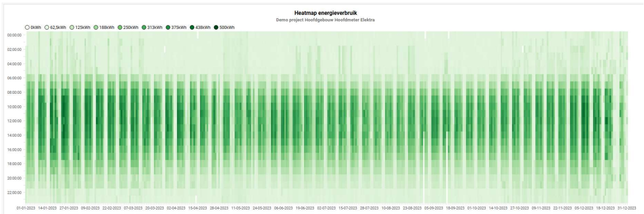
Input variables	MSE	RMSE	MAE	R2
Omloopklep toev.plaatw. (percentage) - Omloopklep PID toerenregelaar toevoervent. berekende waarde - Druk toevoerkanaal - 332PT1 Status	0,001	0,034	0,008	0,999
PID toerenregelaar toevoervent. berekende waarde -	0,005	0,069	0,013	0,996
Status	0,005	0,069	0,013	0,996
PID toerenregelaar toevoervent. berekende waarde - Druk toevoerkanaal - 332PT1 Status	0,004	0,062	0,012	0,996
PID toerenregelaar toevoervent. berekende waarde - Druk toevoerkanaal - 332PT1 Status	0,001	0,035	0,009	0,999
Inblaastemperatuur - 332TT4PID toerenregelaar toevoervent. berekende waarde -Druk toevoerkanaal - 332PT1 Status Omloopklep toev.plaatw. (percentage) - Omloopklep	0,001	0,034	0,008	0,999
Inblaastemperatuur - 332TT4PID toerenregelaar toevoervent. berekende waarde -Druk toevoerkanaal - 332PT1 Omloopklep toev.plaatw. (percentage) - Omloopklep	0,001	0,032	0,008	0,999

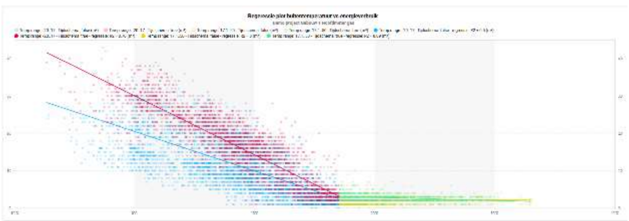




Virtual sensors and advanced symptoms

Anomaly detection: energy demand





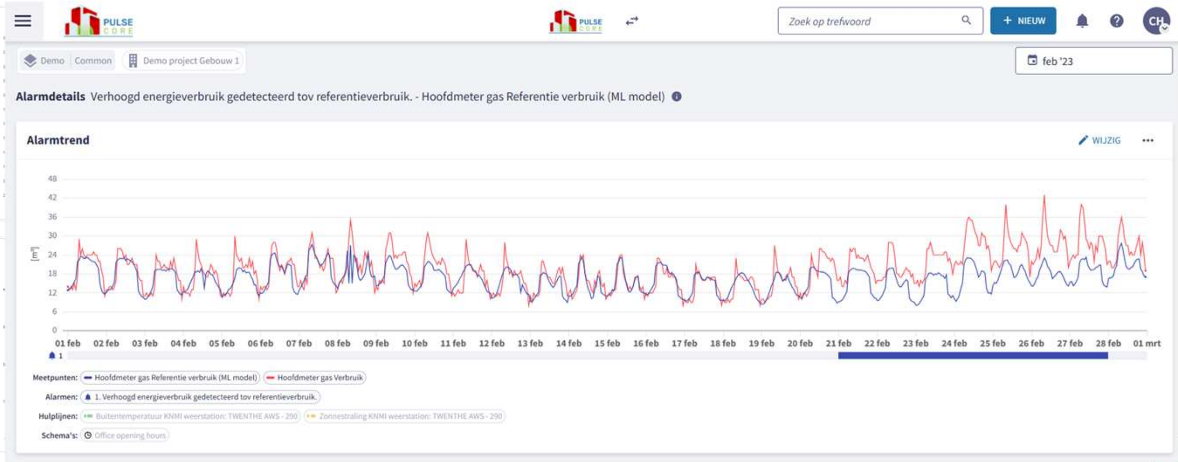
Train model

Training resultaten

	Model resultanten (geen splitsing train/test data)
Versie naam	Geen versie gedefinieerd opgeslagen als standaard.
Aantal samples	356
Coefficient of determination (R2)	0,949
Mean absolute error	31,61
Symmetric mean percentage error	14,3%
Root mean square error	43,47
	Model splitsing eigenschappen
Combinaties succesvol gefit	13
Combinaties niet gefit	1
Temperatuurranges	-20..17, 17.1..50
Dagelijkse periodes	0..23
	Model adviezen
	Model opmerkingen
Opmerking	Door trainingsinterval van 1dag word er geen tijdschema gebruikt en 1 dagperiode (0 t/m 23h) gebruikt

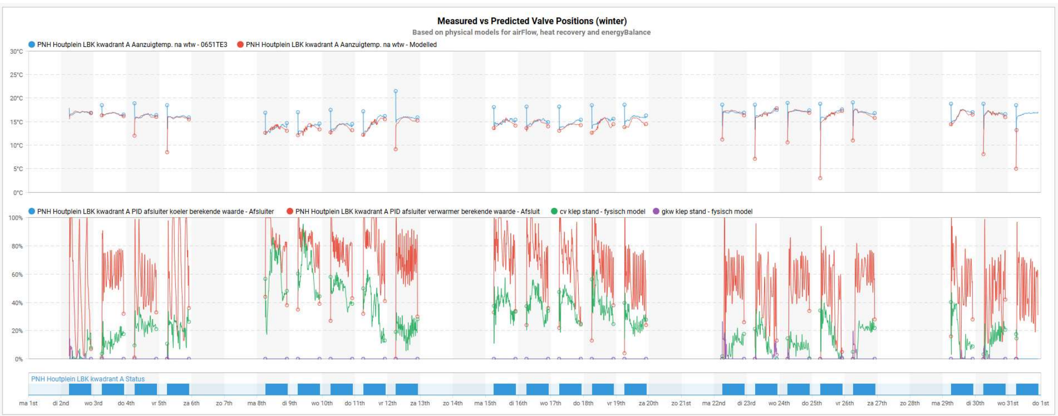
Virtual sensors and advanced symptoms

Anomaly detection: energy demand



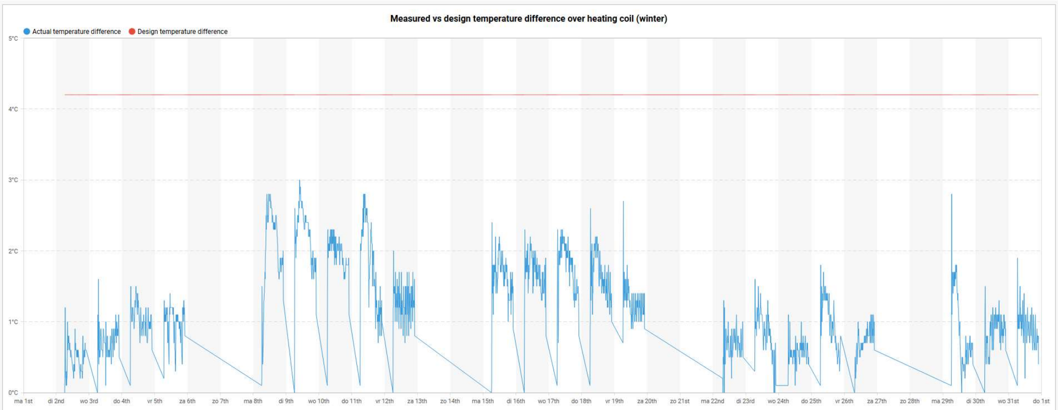
Virtual sensors and advanced symptoms

Anomaly detection: AHU valve positions (physical model)



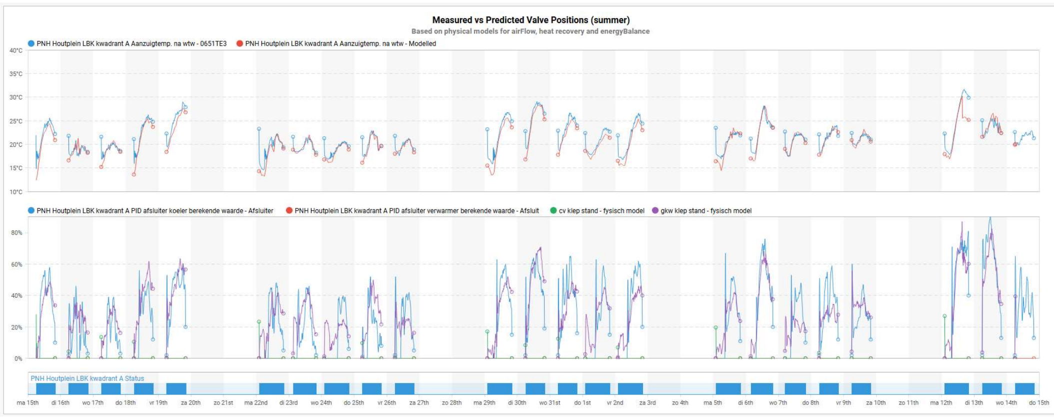
Virtual sensors and advanced symptoms

Anomaly detection: AHU valve positions (physical model)



Virtual sensors and advanced symptoms

Anomaly detection: AHU valve positions (physical model)



Virtual sensors and advanced symptoms

Anomaly detection conclusions:

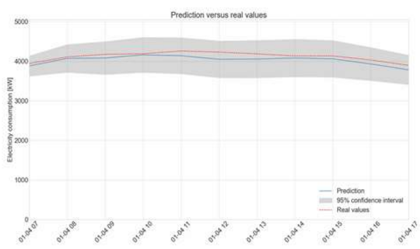
- Selecting input features based on semantic data model makes approach replicable / scalable, but not 100% accurate.
- 80% can be automated, last 20% needed for final feature selection and data quality checks. Create good supportive apps to make this easy for users.
- Physical approach on system level becomes quickly complex, but combining multiple small models is powerful.
- Boosted Decision Trees are powerful in cases where prediction values have a specific range (eg valve 0-100%).
- When prediction ranges are expected to change over time and are not present in training data other methods which derive relations might have better performance.
- Smart implementation of (simple) linear regression is in a lot of cases also performing well.
- Neural Networks can be used as 1 model for multiple sensors and detect anomalies on them. This comes with computational penalty.

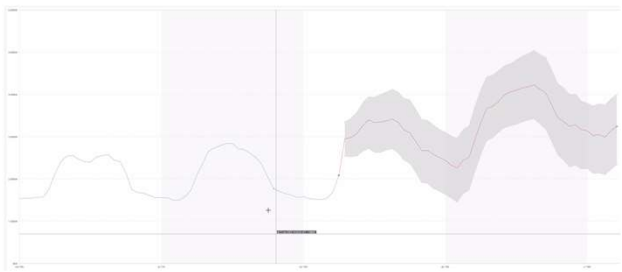
29

Virtual sensors and advanced symptoms

Future prediction: electrical power of a campus

- Challenges:
  - Big campus with multiple building types (offices, production)
  - No insight in occupants, production numbers
- High accuracy on complete profile is easy, high accuracy on peaks is harder
- Alternative approach than timeseries forecasting; ensemble model (linear regression and boosted decision trees)
- Features:
  - Short term historical energy consumption params
  - Weather forecast





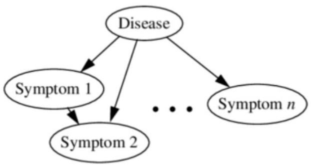
30

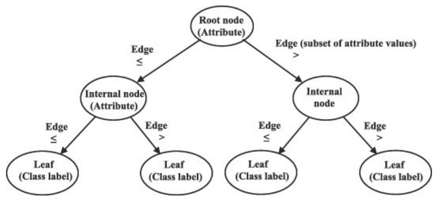
Virtual sensors and advanced symptoms

Future predictions conclusions:

- Data driven approach
- Performance is good
- Added value proven and confirmed by customers
- Pilot phase where multiple questions are open:
  - Performance of timeseries forecasting with forecasting variables as input
  - Scalability and replicability

Fault detection and diagnosis





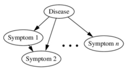
Diagnostic Bayesian networks:

- Calculate probabilities for each diagnose
- Model definition:
  - Prior probabilities
  - Conditional probabilities (relation between symptom and diagnosis)

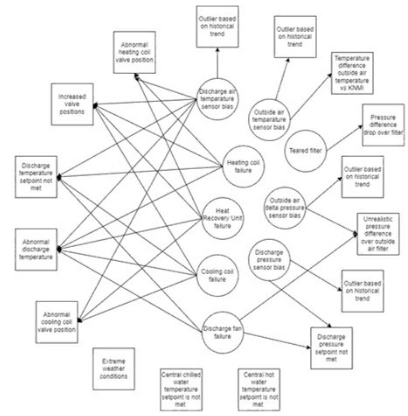
Rule based:

- Detect what will be the most logical diagnosis
- Model definition:
  - Define for each possible diagnosis what is /are the combination of symptom states
  - Ideally each diagnosis has each own unique combination of symptom states
  - Practically this is not the case, a leaf can have multiple diagnosis

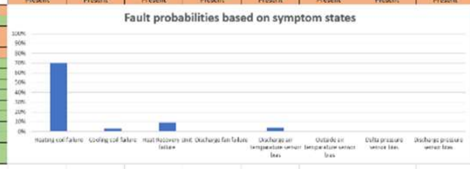
Fault detection and diagnosis



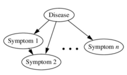
FDD using DBN: AHU



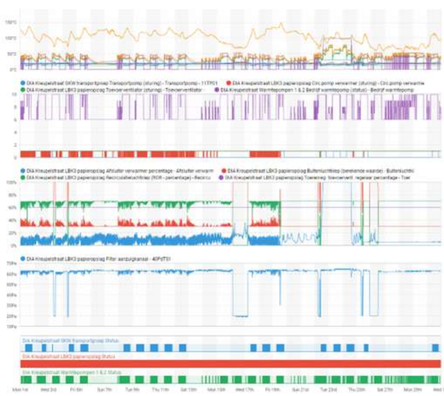
	Symptom state	Heating coil failure	Cooling coil failure	Heat Recovery Unit failure	Discharge fan failure	Discharge air temparature sensor bias	Outside air temparature sensor bias	Delta pressure sensor bias	Discharge pressure sensor bias
		Present	Present	Present	Present	Present	Present	Present	Present
Abnormal heating coil valve position during heat demand	Present								
Abnormal cooling coil valve position during cool demand	Absent								
Increased valve positions	Present								
Abnormal discharge temperature	Present								
Discharge temperature not met	Present								
Discharge pressure not met	Absent								
Central hot water temperature is not met	Absent								
Central chilled water temperature is not met	Absent								
Extreme weather conditions	Absent								
Outlier discharge temperature	Absent								
Outlier outside temperature	Absent								
Outlier delta pressure filter	Absent								
Outlier discharge pressure	Absent								
Unrealistic delta pressure filter	Absent								
Big temperature difference between outside air temperature and KNMI	Absent								
FAULT PROBABILITIES		70%	3%	9%	0%	4%	0%	0%	0%

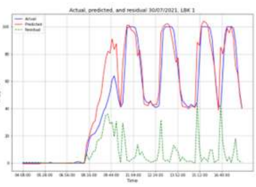


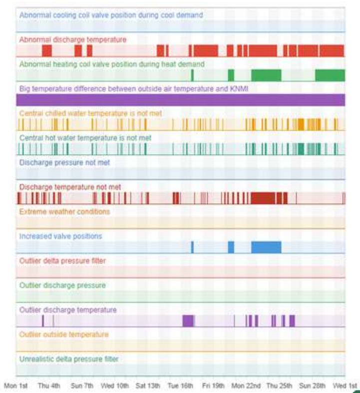
Fault detection and diagnosis



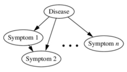
FDD using DBN: AHU







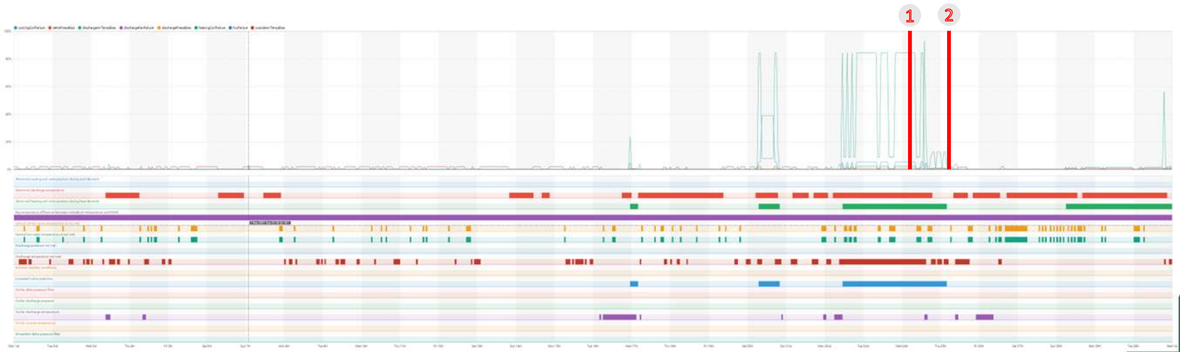
Fault detection and diagnosis



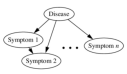
FDD using DBN: AHU

Fault probabilities and symptom states for 1 month:

1. Reported failure: stuck heating coil valve and broken valve actuator, temporary fix
2. Replaced valve actuator



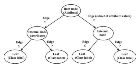
Fault detection and diagnosis



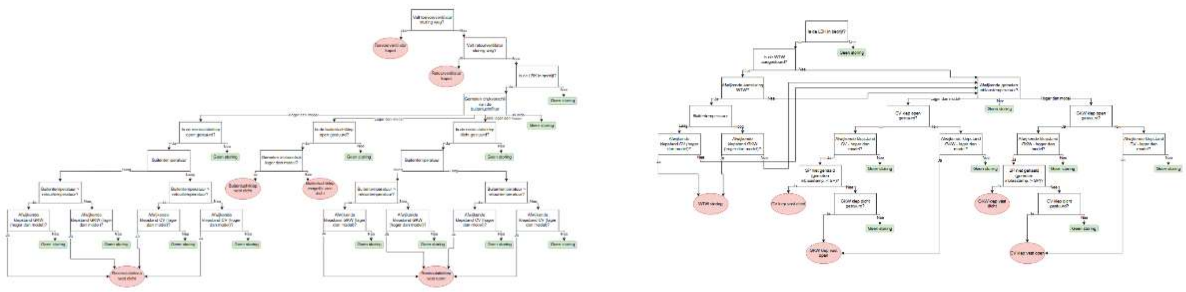
DBN conclusions:

- Failure probabilities have high added value
- No labelled data sets means that DBN can only be defined by experts; difficult process.
- Not flexible to fit different situations; our test show it is not easy to adjust a “master“ version model:
  - Symptom states that cannot be calculated
  - Different system configurations or controls
- Symptom threshold definition: when should a symptom change its state?
  - Absolute threshold and duration
  - Impact of model reliability
  - Local characteristics

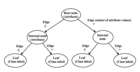
Fault detection and diagnosis



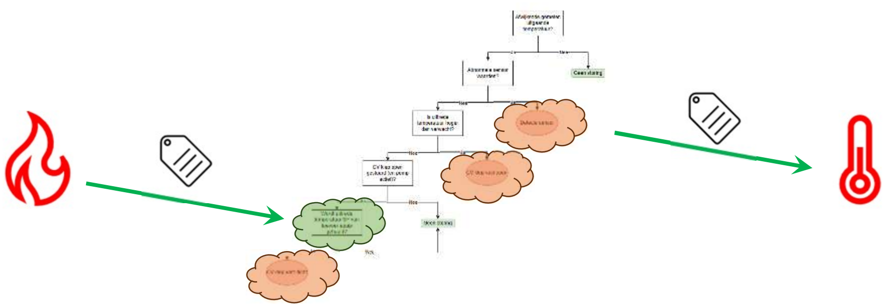
FDD using rule based: AHU



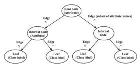
Fault detection and diagnosis



FDD using rule based: distribution group



Fault detection and diagnosis



Rule based conclusions:

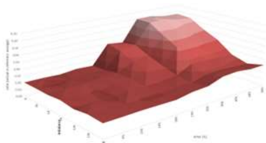
- Single diagnosis is a drawback
- Easier to define using expert knowledge but harder to define complex systems
- (Sub) component approach is promising, but translation to system level is still work in progress
- (Sub) component approach gives flexibility toward different system configuration.
- Symptom threshold definition: same challenges as in DBN approach

Challenges / opportunities

- Balance between scalability / replicability and reliability
  - Focus: on specific type building (office buildings), specific system types
  - Data model for whole life cycle
  - Remaining manual work should be as easy as possible, supported by specific apps
- Symptom tuning
  - Automate based on model performance parameters
- Data is not labelled
  - Backward engineering: created labelled datasets based by historical FDD analysis
  - Failure definition in operational process

Conclusions and further work

- End 2 end test containing:
  - Anomaly detection and missing sensors, combination of:
    - ML models
    - simple physical models
  - FDD: rule-based approach on subcomponent level (aim for 80% most common failure)
  - Focus: AHU and (mixing) valves
  - Automated integration in operational process
- Energy forecasting
  - Generic approach on office buildings
- Predictive maintenance
  - Vibration sensors on rotating equipment (chillers and fans); time domain evaluation



Questions

Thank you for your attention!

Any questions?