Sensor Fault Detection with a Weighted Ensemble of ARMAX Forecasting Models

Mario Kahlhofer | Team Data Preparator

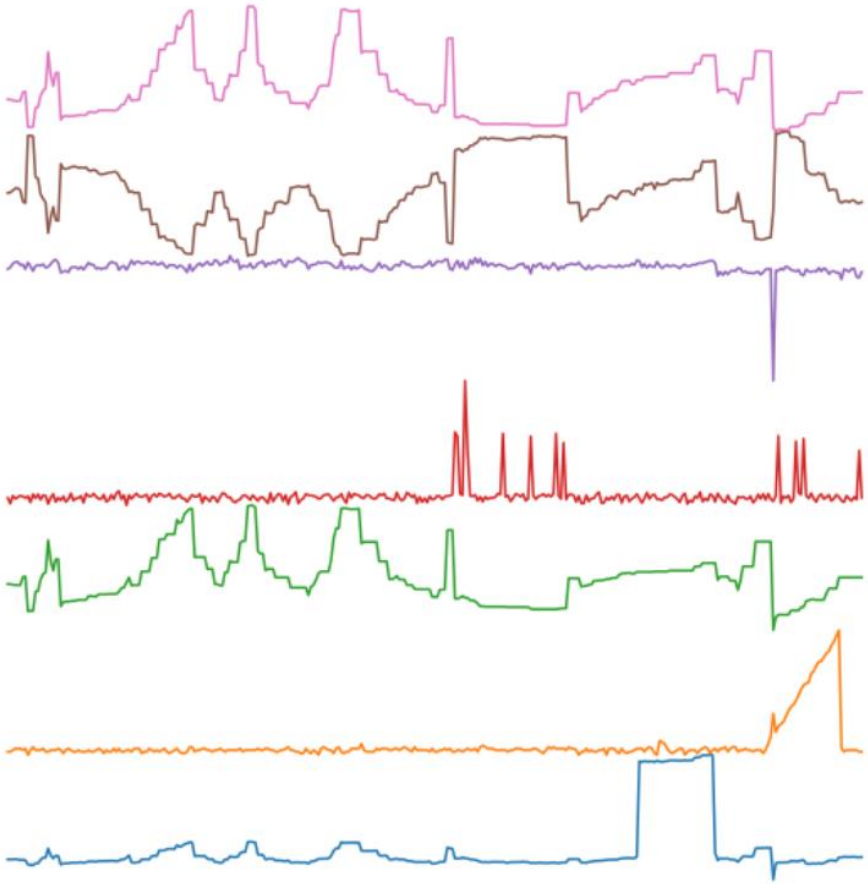
LEC Data Challenge
26 September 2019

JOHANNES KEPLER
UNIVERSITY LINZ

The Challenge

15 sensor measurements

339 points per sensor



The illustration to the right only shows the first 7 sensors.

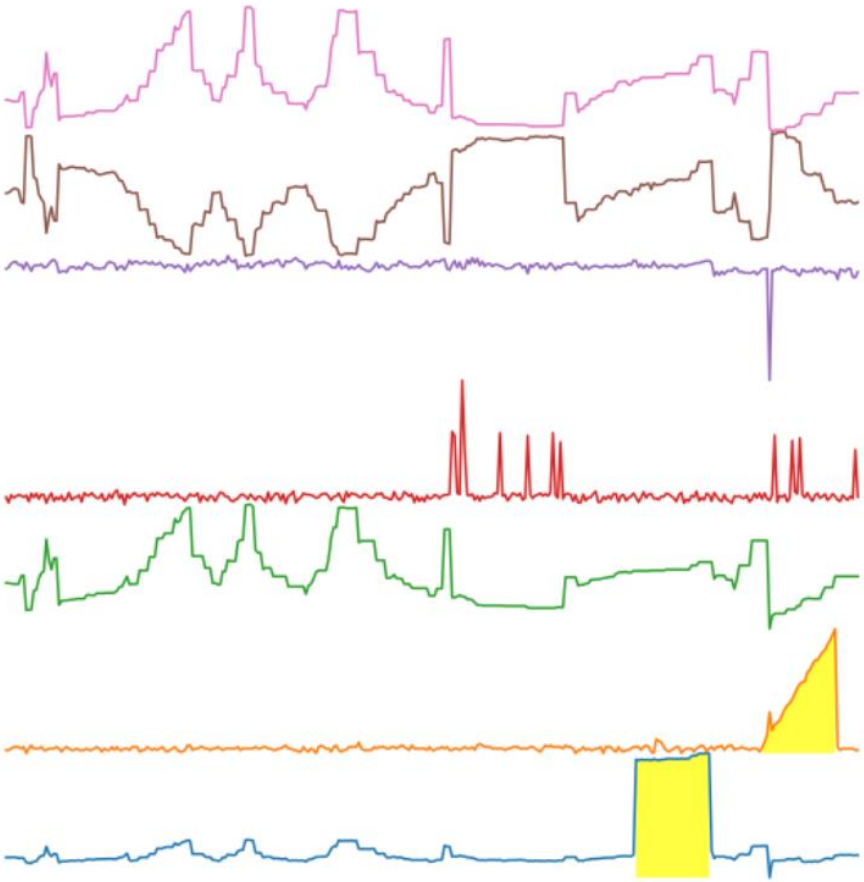
JOHANNES KEPLER UNIVERSITY LINZ

The Challenge (cont.)

15 sensor measurements

339 points per sensor

Only two regions are annotated as faulty!

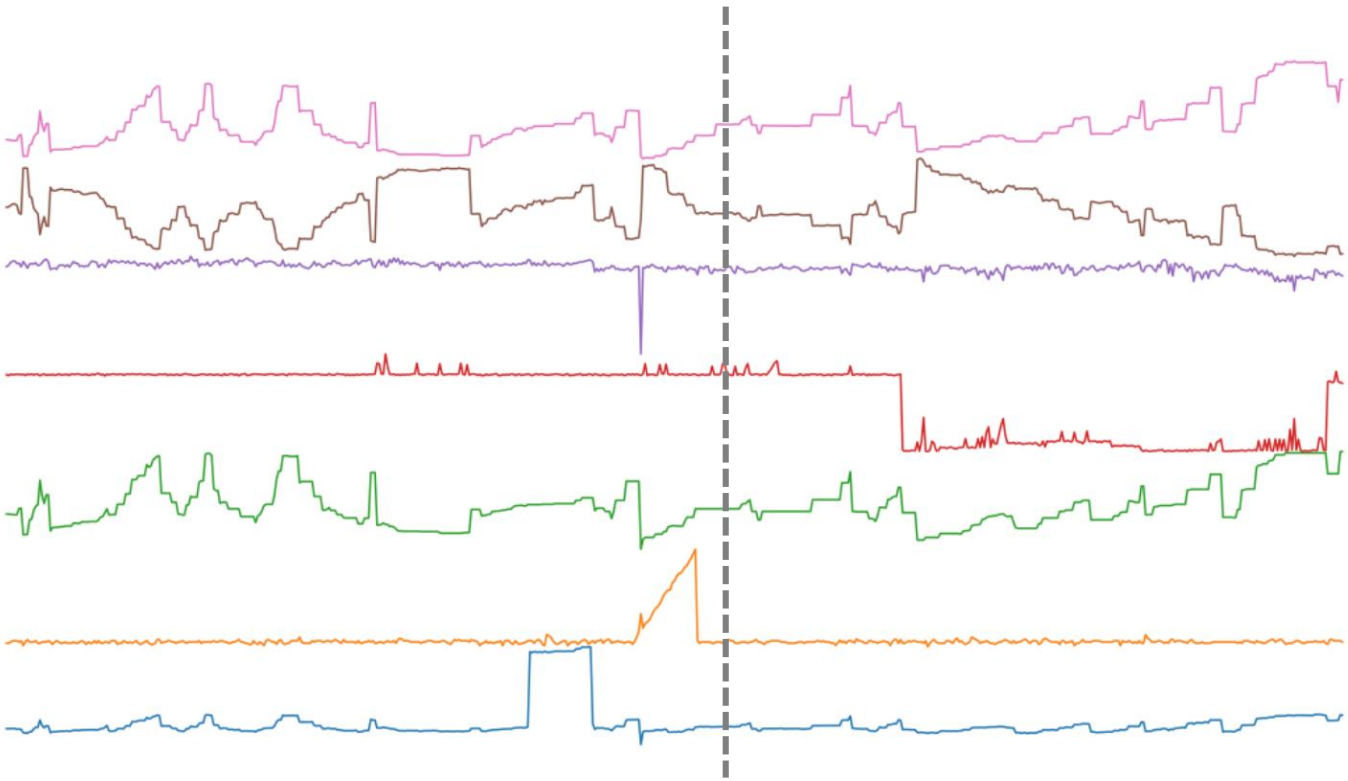


The illustration to the right only shows the first 7 sensors.

JOHANNES KEPLER UNIVERSITY LINZ

The Challenge (cont.)

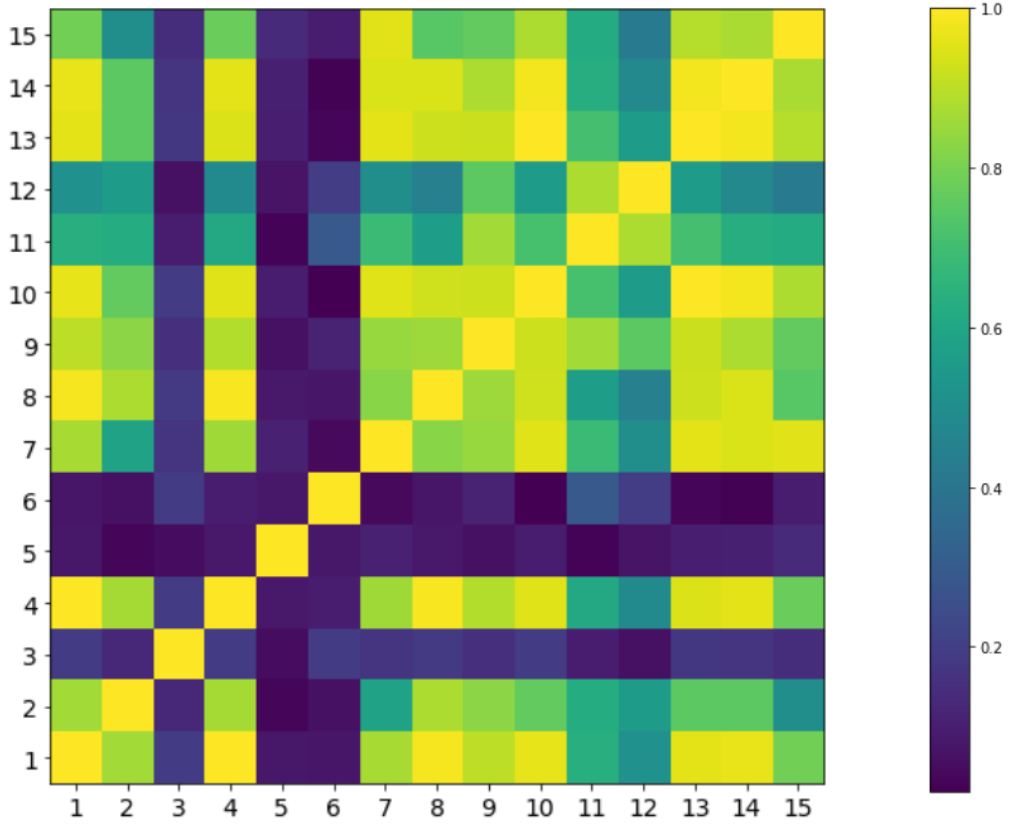
Detect the remaining faults in the test set!



Explorative Data Analysis

Spearman Correlation of sensor measurements

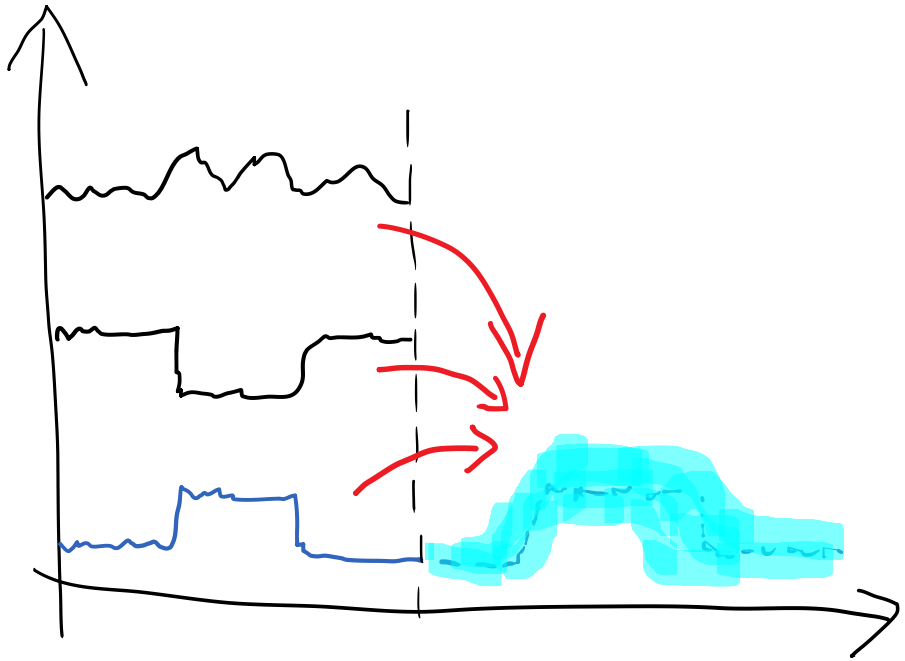
Why?
Be resistant to outliers!



JOHANNES KEPLER UNIVERSITY LINZ

Approach

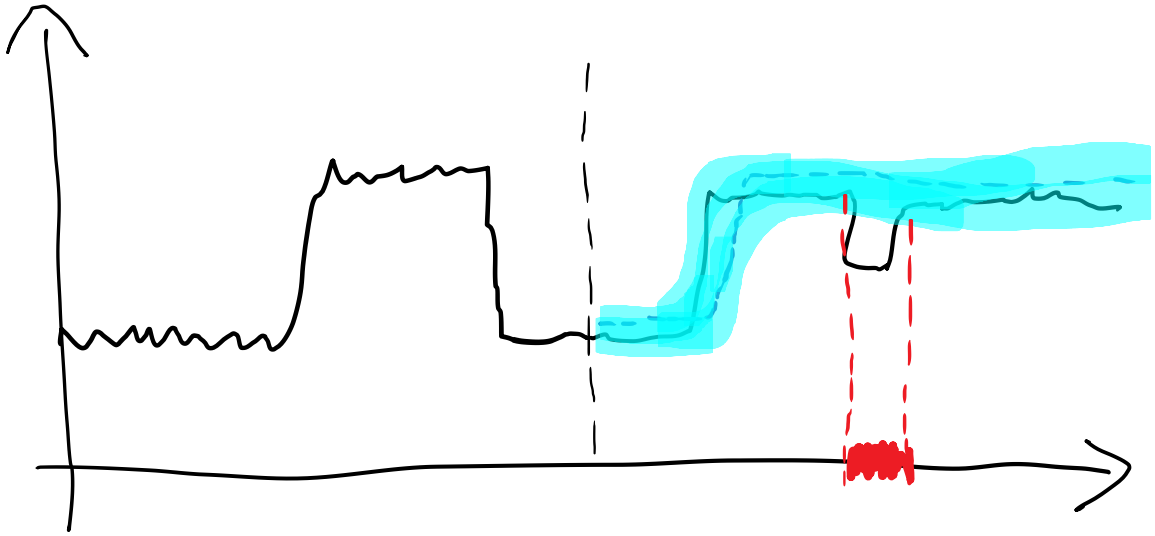
1. Per sensor, use ***entire past*** to forecast the future



JOHANNES KEPLER
UNIVERSITY LINZ

Approach (cont.)

2. Compare the real measurements with the forecast and alert on significant deviations

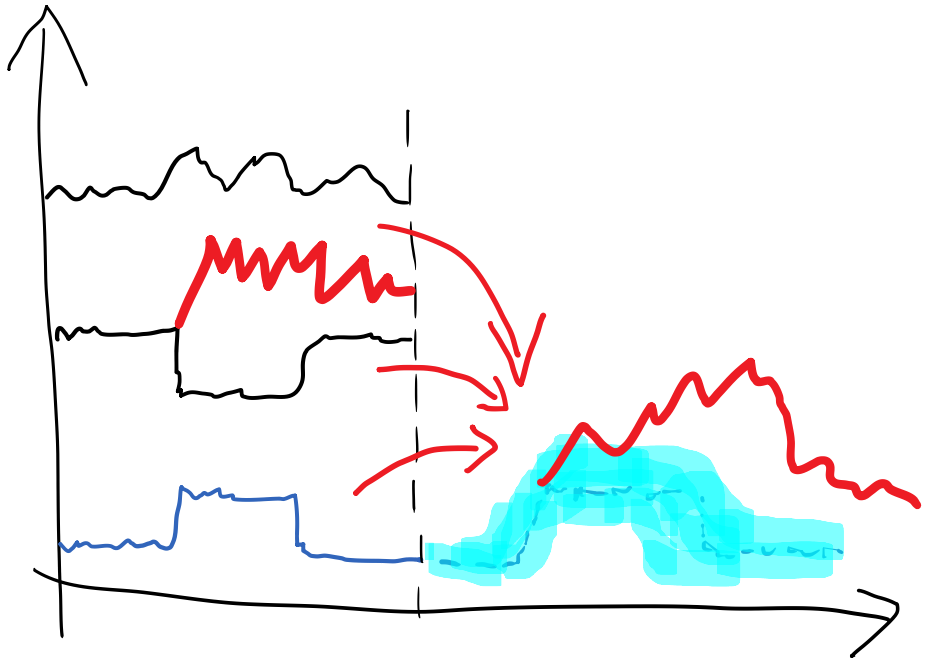


JOHANNES KEPLER UNIVERSITY LINZ

Approach (cont.)

There's a catch!

We can't trust the other sensors, because the could be faulty as well

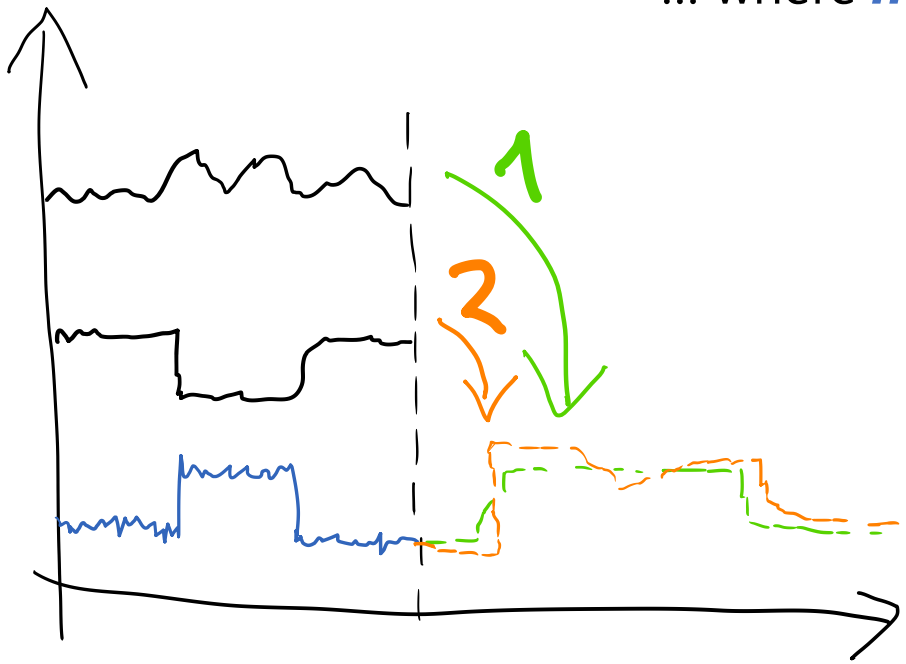


JOHANNES KEPLER UNIVERSITY LINZ

Revised Approach

Use an ensemble of models (feature bagging):
Per sensors, make ***n – 1*** forecasts, and combine them

… where ***n*** is the number of sensors



Why?
A smart combination should somehow neglect faulty predictions

JOHANNES KEPLER UNIVERSITY LINZ

Method

1. For each sensors, train ***n – 1 ARMAX models***
   - ***AR*** Autoregression *(this sensor's past)*
   - ***MA*** Moving Average *(this sensor's past)*
   - ***X*** Exogenous Factors *(another sensor's past)*

$$X_t = \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \sum_{i=1}^{b} \eta_i d_{t-i}$$

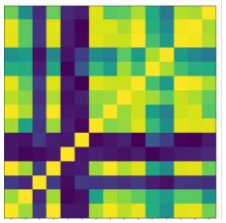
c ***AR*** ***MA*** ***X***

Why?
Because we can use the entire history to model the future

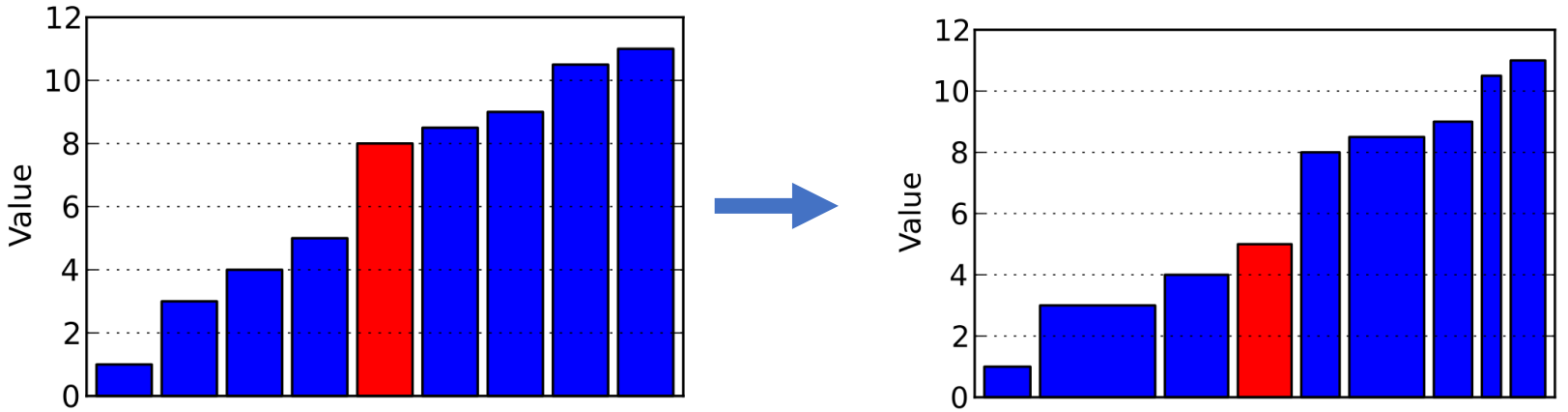
JOHANNES KEPLER UNIVERSITY LINZ

Method (cont.)

2. Combine the forecasts and error bands by taking their ***weighted median***



Remember the correlation matrix from before?
These are our weights!



User Bscan, Wikimedia Commons. 2016.
URL: https://commons.wikimedia.org/wiki/File:Weighted_median.svg

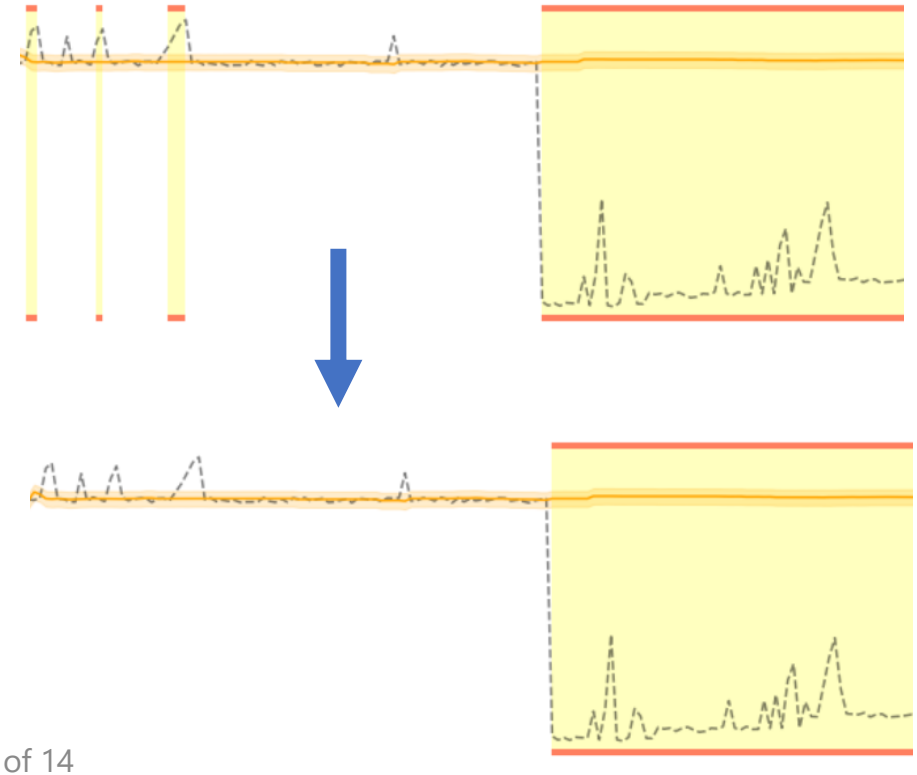
Why? Neglect predictions with weak correlation

JOHANNES KEPLER UNIVERSITY LINZ

Method (cont.)

There's one more thing: Smooth the final predictions!

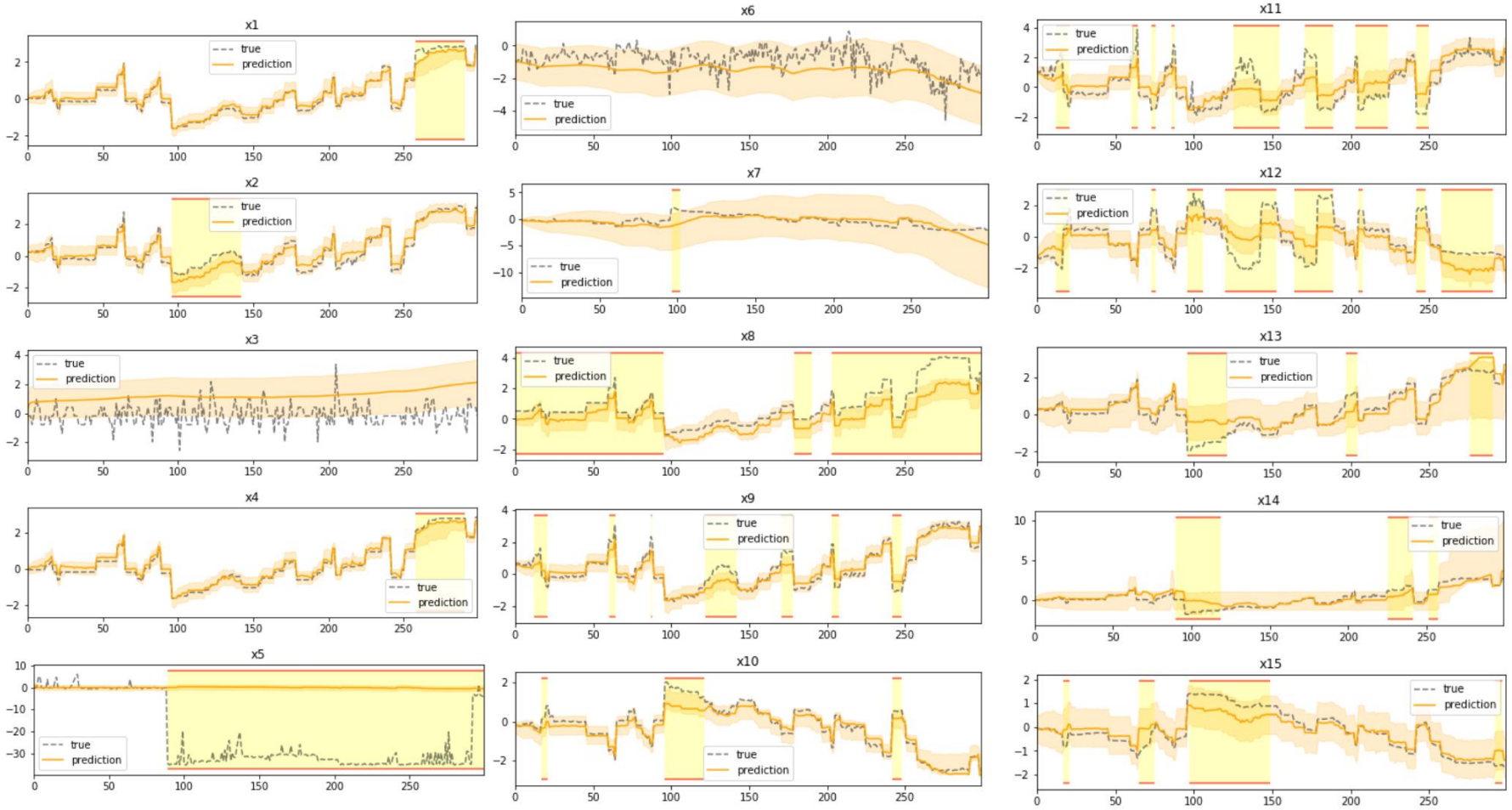
*E.g., with a **1D convolution** with window size ~ 5-10*



Why?
Remove uncommon temporal outliers

JOHANNES KEPLER UNIVERSITY LINZ

Final Results on Test Dataset



JOHANNES KEPLER UNIVERSITY LINZ

Conclusion

- Per sensor, train an ensemble of ***n – 1 ARMAX forecasting models*** and combine the predictions with a ***weighted median***, based on their ***Spearman correlation*** coefficient
- **Future Work:** Adjustable sensitivity by computing different quantiles of the error bars

- Check out the Python source code on GitHub

github.com/blu3r4y/lec-data-challenge-2019

JOHANNES KEPLER UNIVERSITY LINZ

Sensor Fault Detection with a Weighted Ensemble of ARMAX Forecasting Models

Mario Kahlhofer | Team Data Preparator

LEC Data Challenge
26 September 2019

JOHANNES KEPLER
UNIVERSITY LINZ