GT S3 : Sûreté, Surveillance, Supervision Meeting
GdR Modélisation, Analyse et Conduite des Systèmes Dynamiques (MACS)
January 28th, 2014

A Data-driven Approach for Remaining Useful Life Prediction of Critical Components

Ahmed Mosallam
ahmed.mosallam@femto-st.fr

Supervisors:
Prof. Noureddine ZERHOUNI
Dr. Kamal MEDJAHER

Outlines

1. Introduction
2. The method
3. Applications and results
4. Conclusion & future work

Introduction

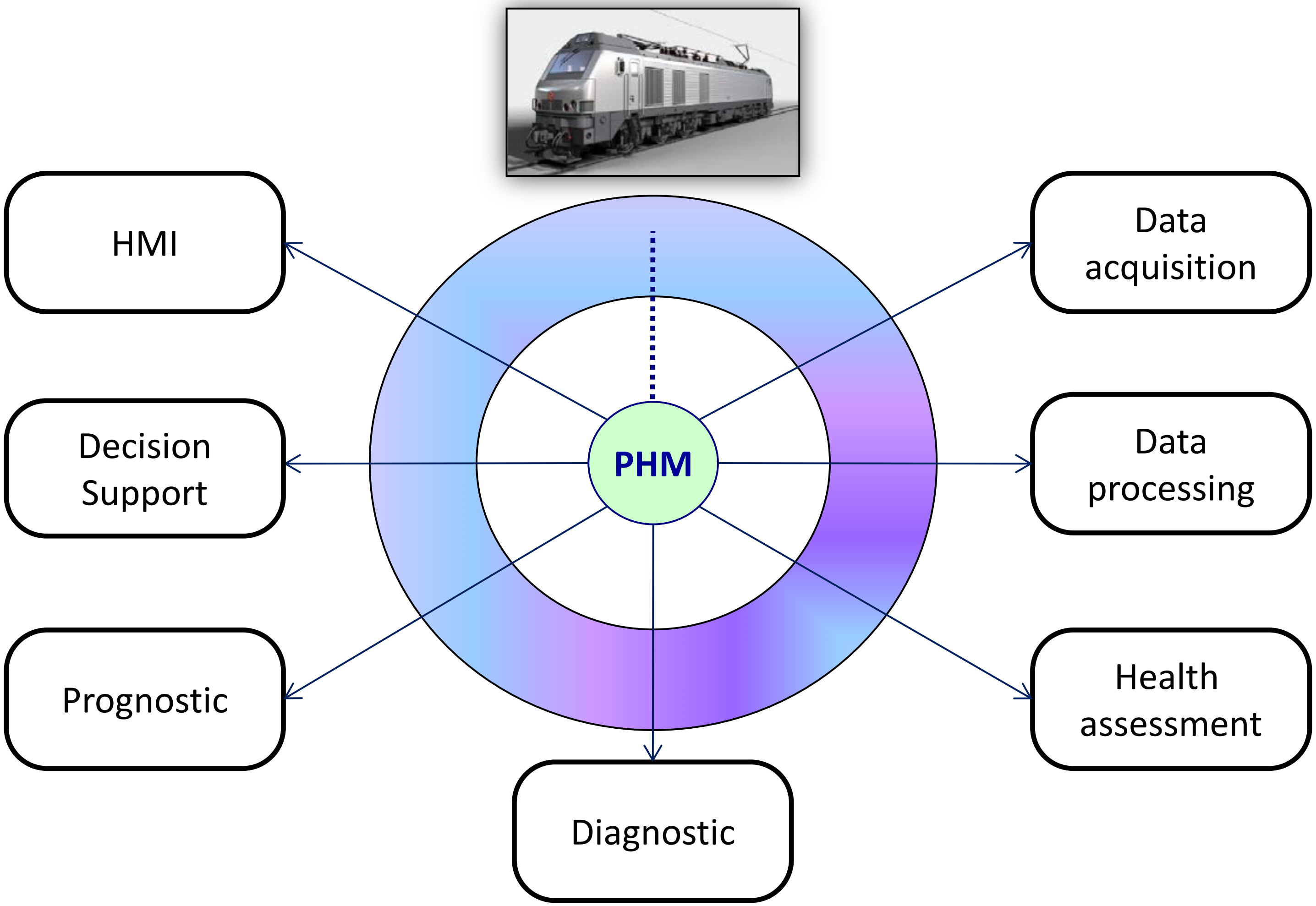


Monitor, Detect, Predict, Anticipate = Prevent and avoid such catastrophes

↗ Reliability ↗ Availability ↗ Security ↘ Costs

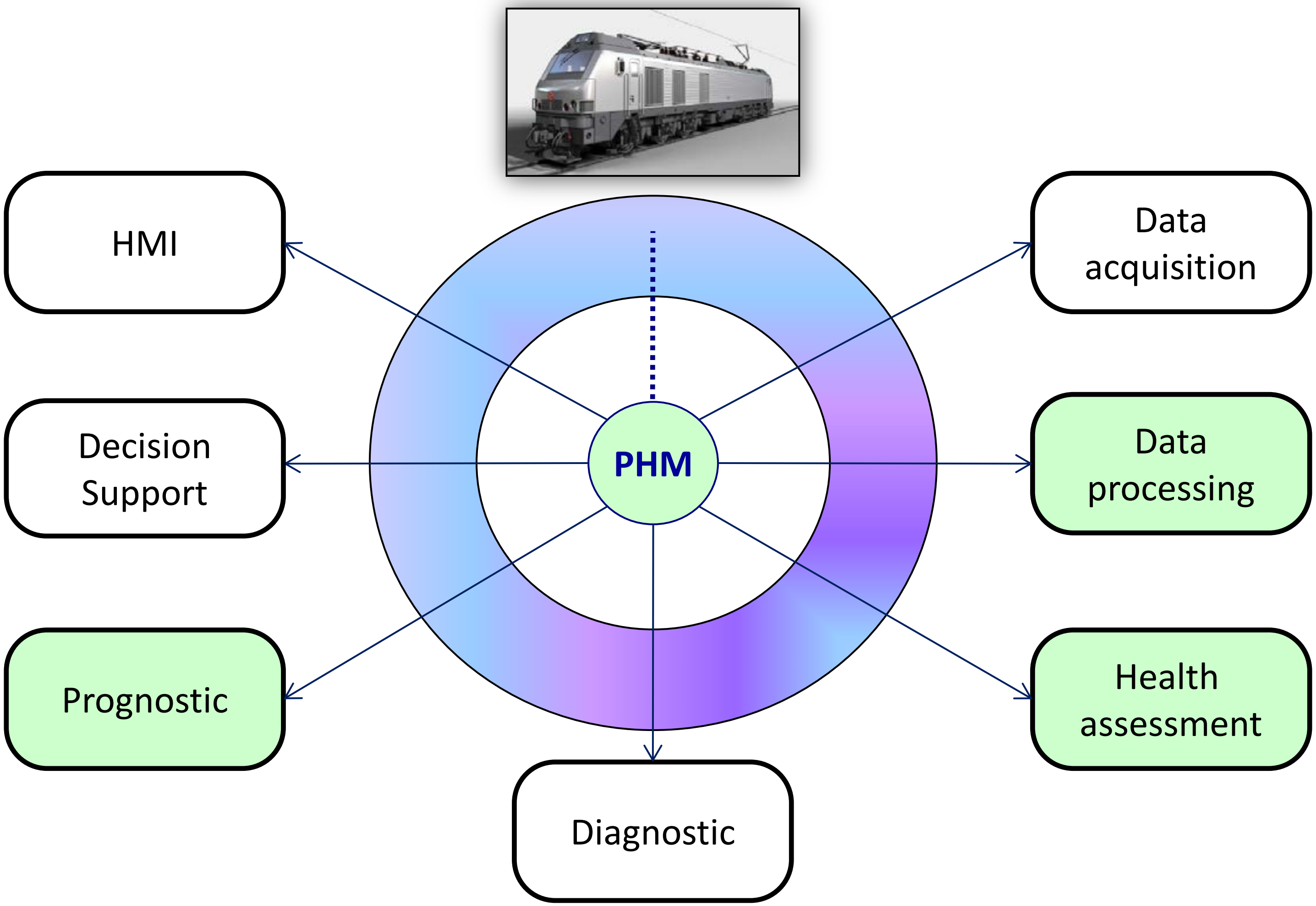
Introduction

Prognostics and health management (PHM)



Introduction

Prognostics and health management (PHM)



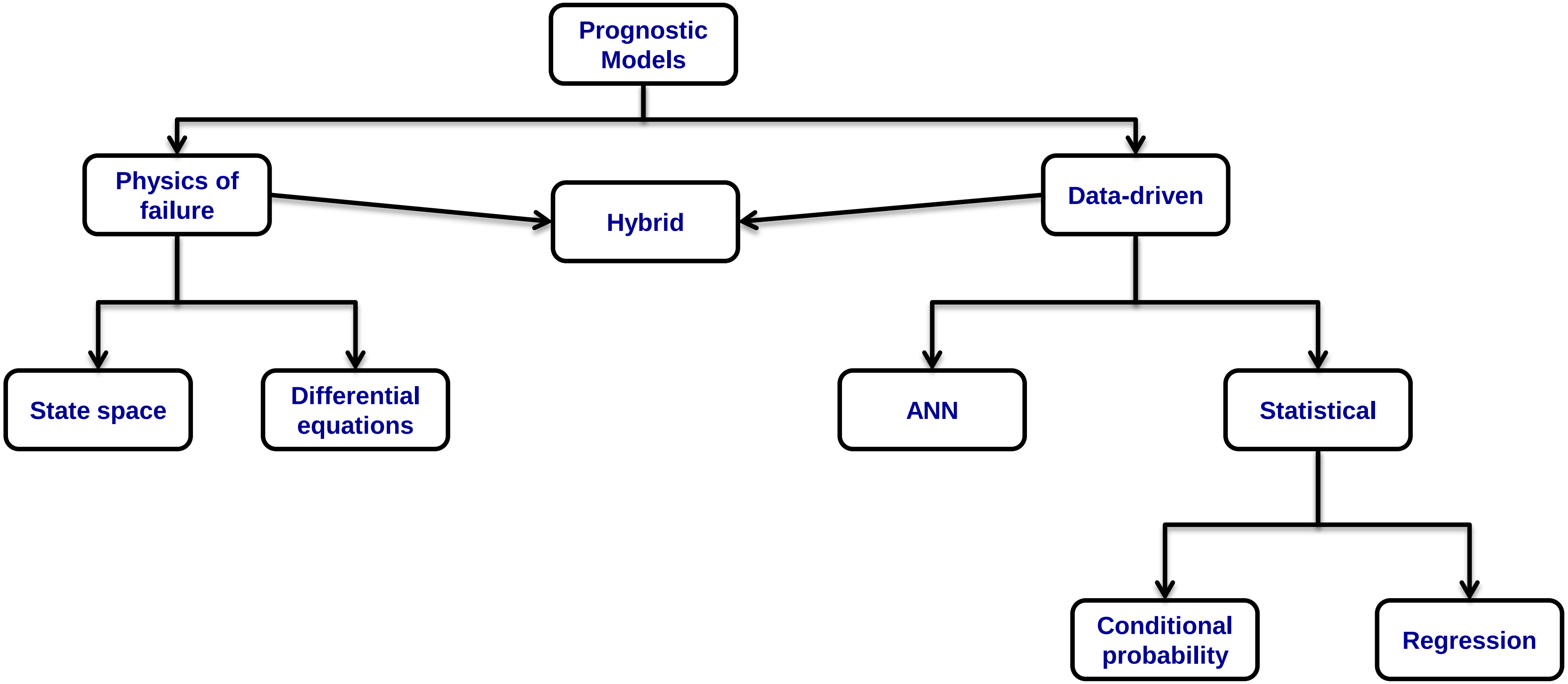
Introduction

Prognostic definitions

- An advance indication of a future event. [Oxford dictionary]
- Estimation of time to failure and risk for one or more existing and future failure modes. [ISO 13381-1, 2004]
- Estimation of the time before failure, or the remaining useful life, and the associated confidence value. [Tobon-Mejia, 2012]
- Indicates whether the structure, system or component of interest can perform its function throughout its lifetime with reasonable assurance and, in case it cannot, to estimate the remaining useful life. [Zio, 2010]
- Predicts how much time is left before a failure (or more) occurs given the current machine condition and past operation profile. [K.S. Jardine, 2006]

Introduction

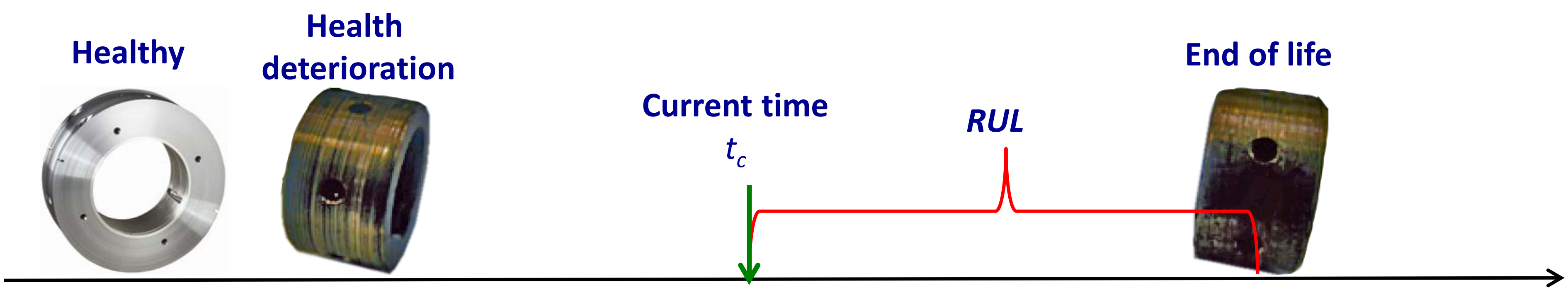
Prognostic models



Introduction

Objectives

- Assess the current status of critical component.
- Predict the remaining useful life (RUL) time at which the critical component will no longer perform its intended function.



Introduction

Challenges

- Domain knowledge

- Measurements
  - Noisy
  - Imprecise
  - Incomplete

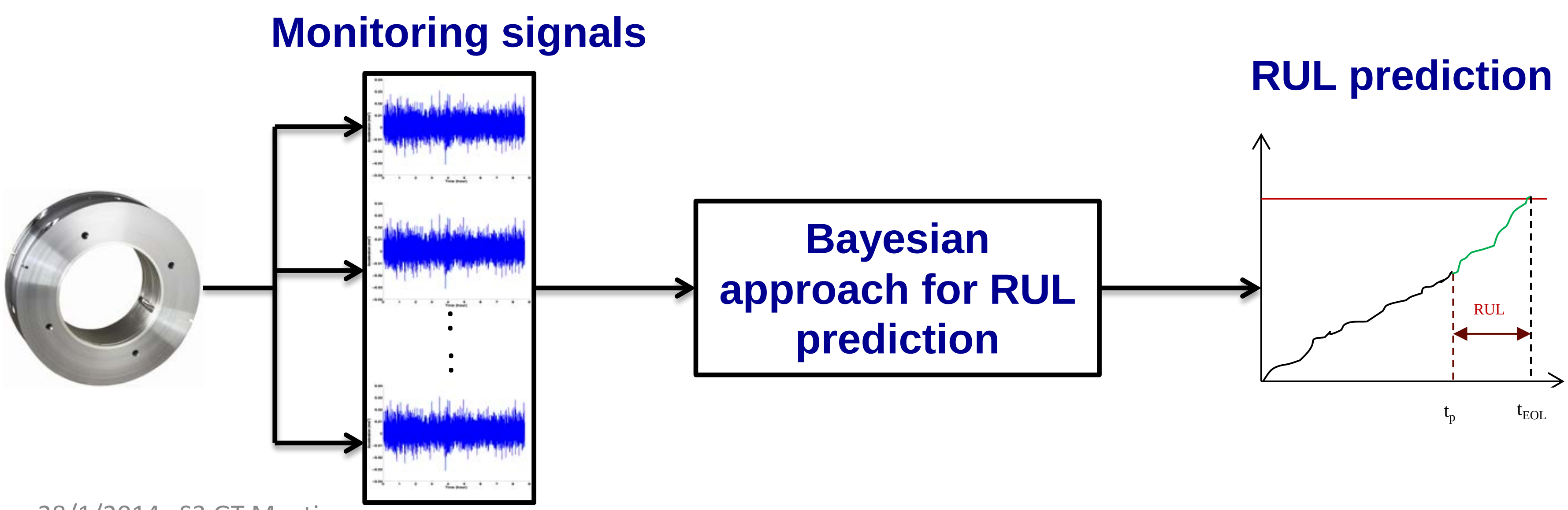
- Model
  - Un-modeled phenomena
  - Approximation and simplification

- Process
  - Unforeseen future loads and environmental conditions
  - Stochastic

The method

Overview

- Variable selection based on mining relationships between signals.
- Extract monotonic trends to represent evolution of the system.
- Using discrete Bayesian filter for online estimation.
- RUL prediction using k-NN and Gaussian process regression.



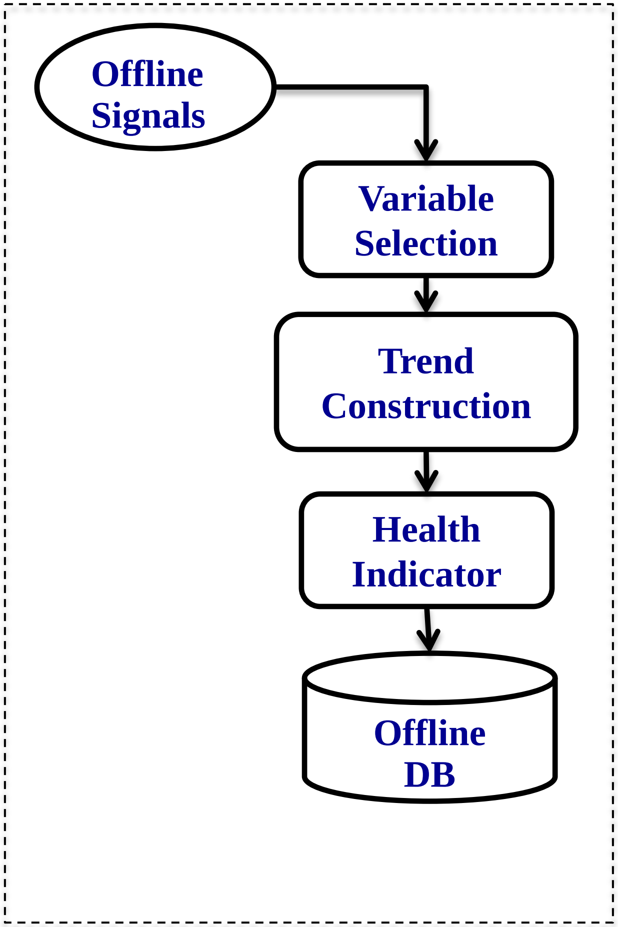
The method

General assumptions

- Domain knowledge
  - No.
- Measurements
  - Input signals: multidimensional time series matrix D_{NxM} where N is number of observations and M is number of sensors.
  - Run to failure.
  - Relations between signals are important.
- Model
  - Data set contains enough samples for training.
  - Level: component (not system).
- Process
  - Operating conditions: constant.

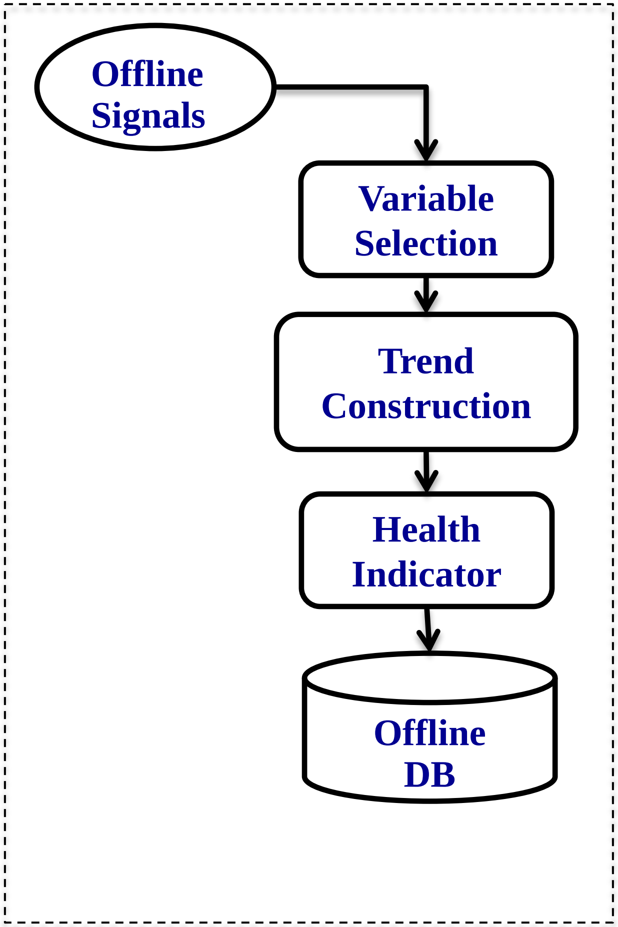
The method

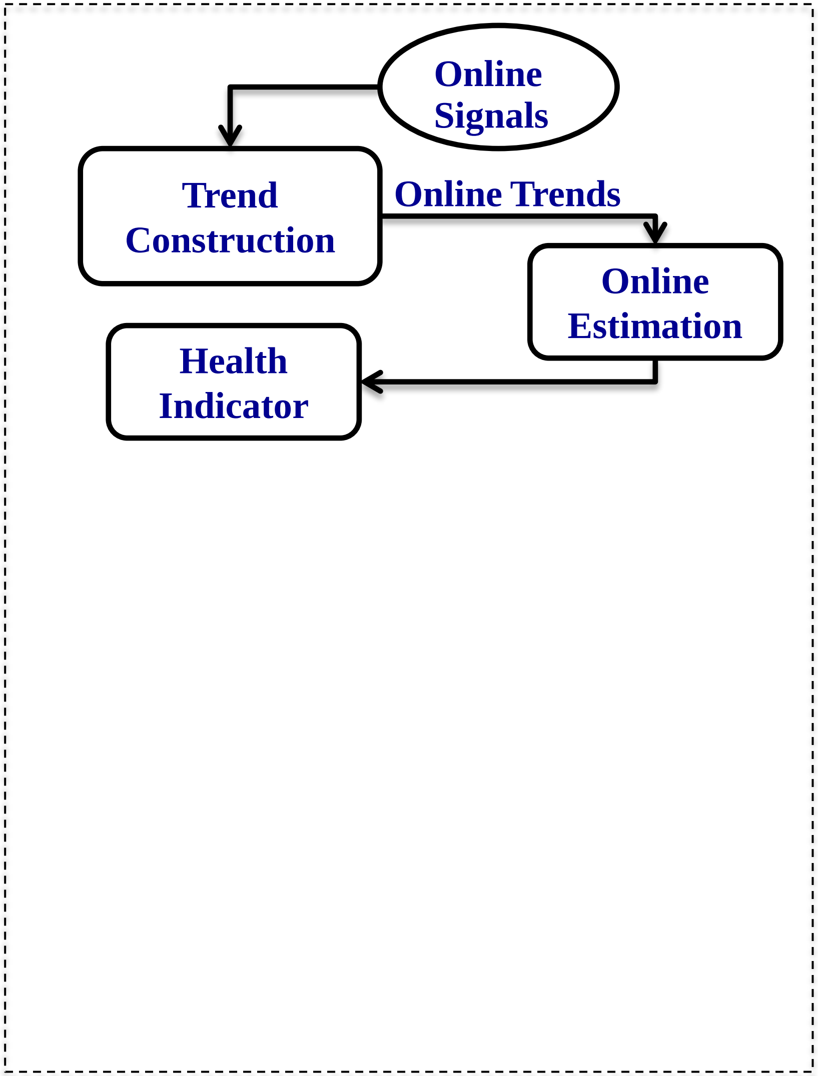
Overall scheme



The method

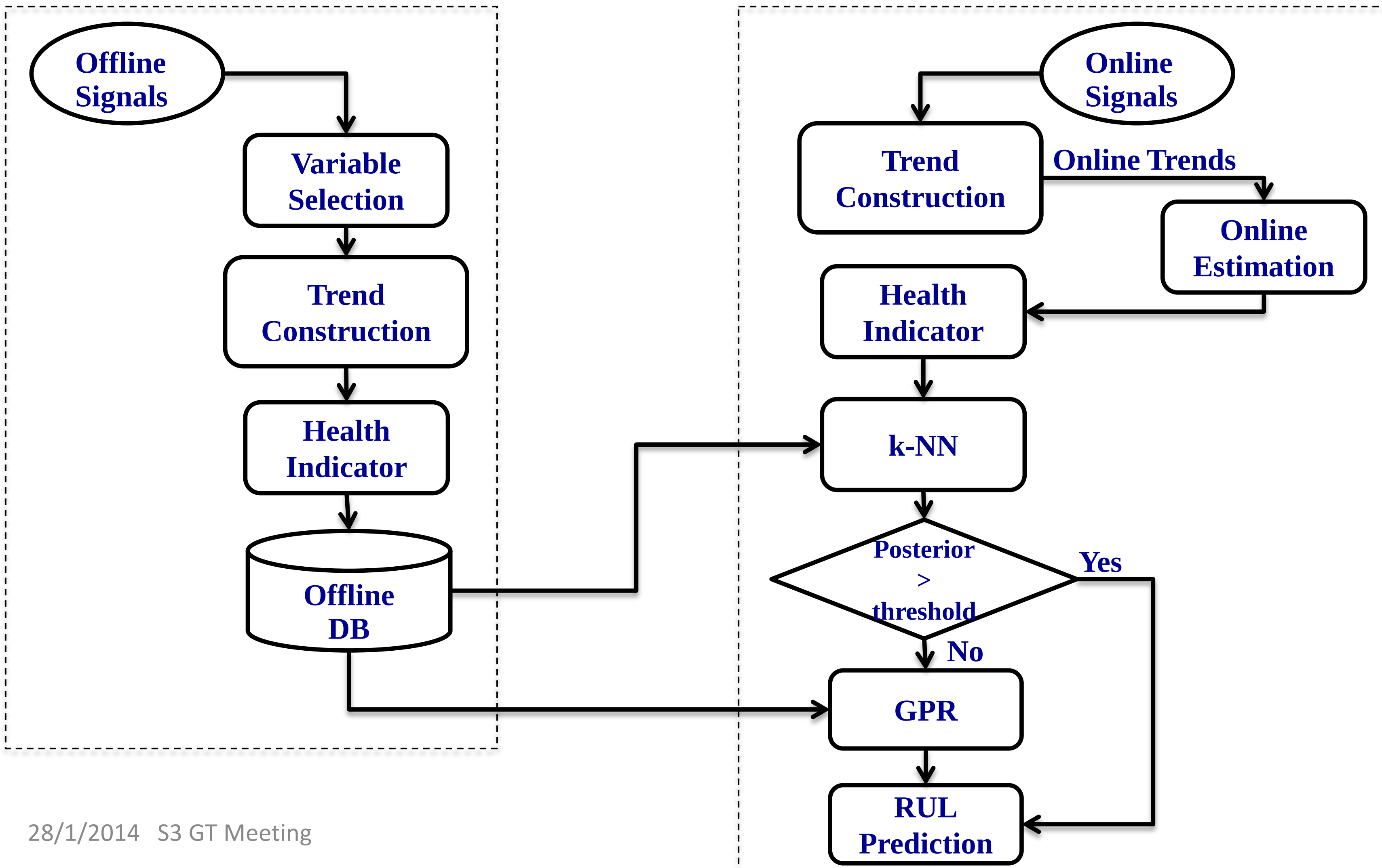
Overall scheme





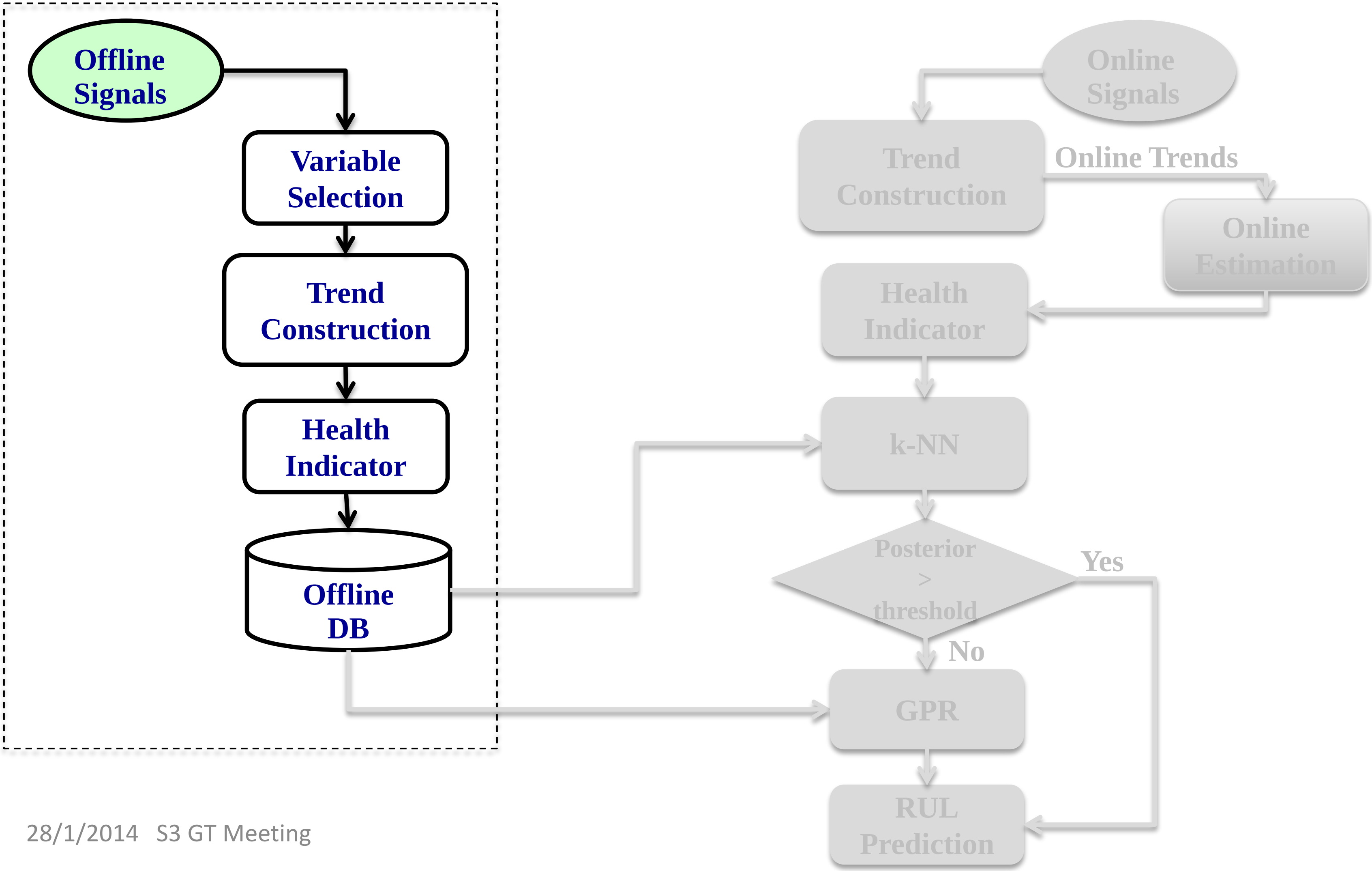
The method

Overall scheme



The method

Overall scheme



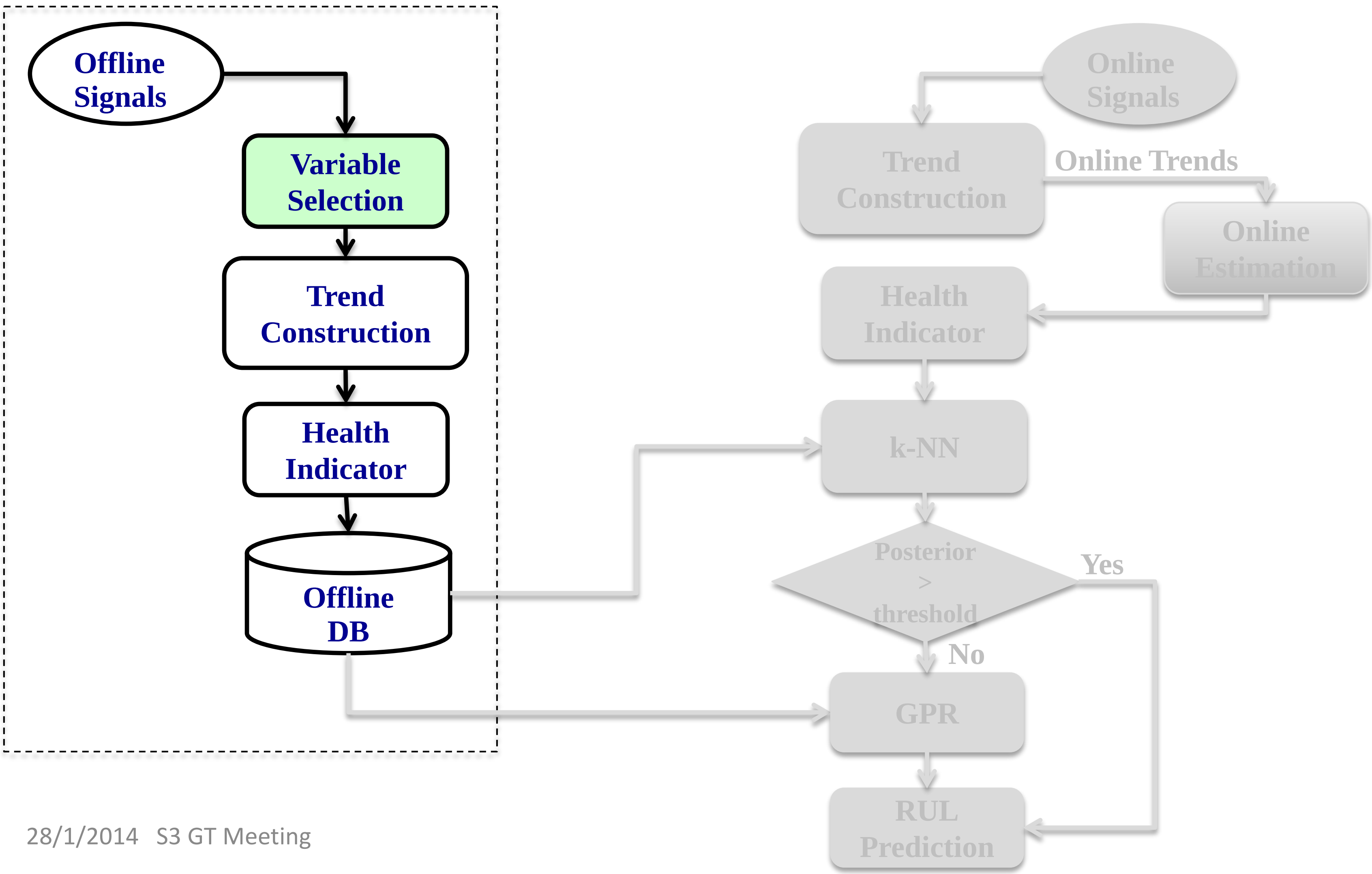
The method

Input data

- Multidimensional time series sensory data.
- Can be represented as a matrix D_{NxM}
  - Where, N is number of observations and M is number of sensors, i.e. variables.
- Signals that have non random relationships contain information about system degradation.
- The challenge is to automatically select the interesting variables.

The method

Overall scheme

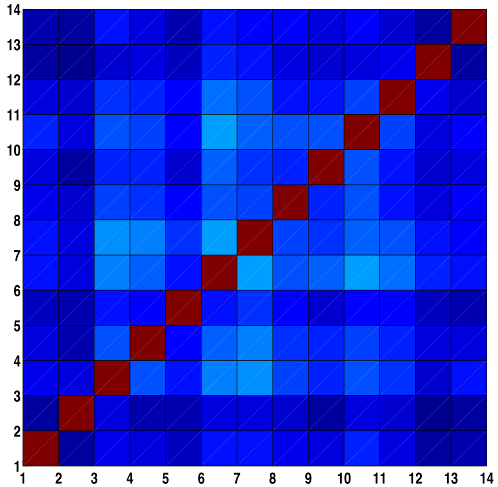


The method

Variable selection

1. Symmetrical uncertainty measure:

$$SU(X,Y) = 2 \times \frac{I(X,Y)}{H(X) + H(Y)}$$



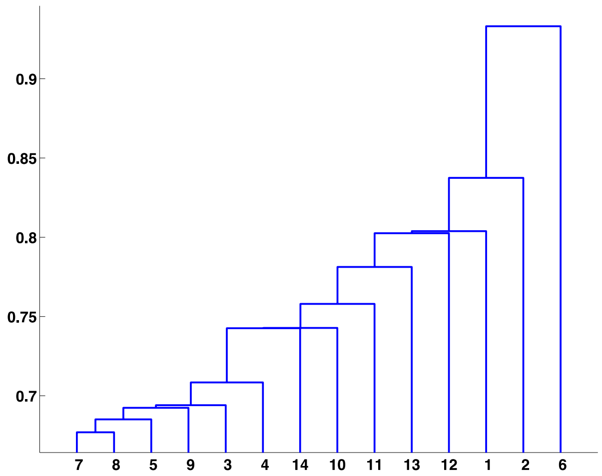
The method

Variable selection

1. Symmetrical uncertainty measure:

$$SU(X,Y) = 2 \times \frac{I(X,Y)}{H(X)+H(Y)}$$

2. Hierarchical clustering and cut off distance was selected automatically.



The method

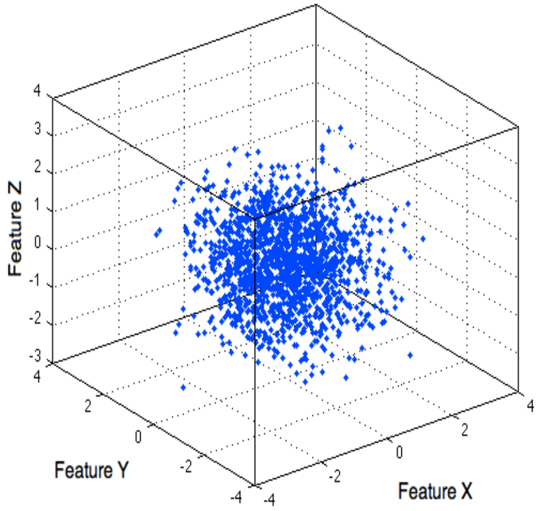
Variable selection

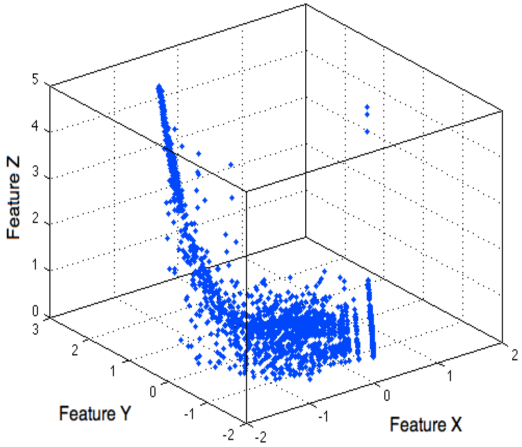
1. Symmetrical uncertainty measure:

$$SU(X,Y) = 2 \times \frac{I(X,Y)}{H(X) + H(Y)}$$

2. Hierarchical clustering and cut off distance is selected automatically.

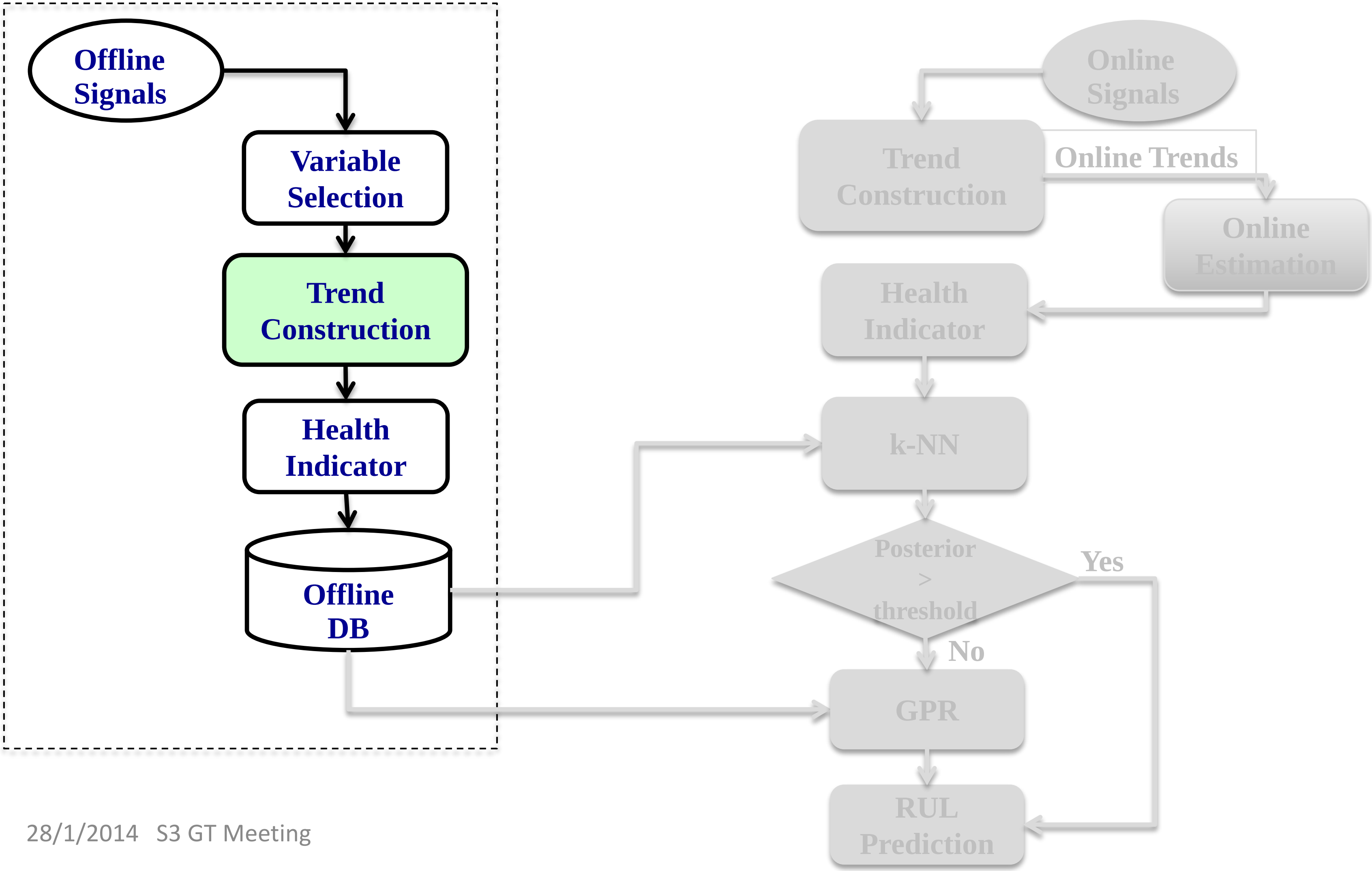
3. Clustering quality using normalized value of distortion measure of the neurons.





The method

Overall scheme



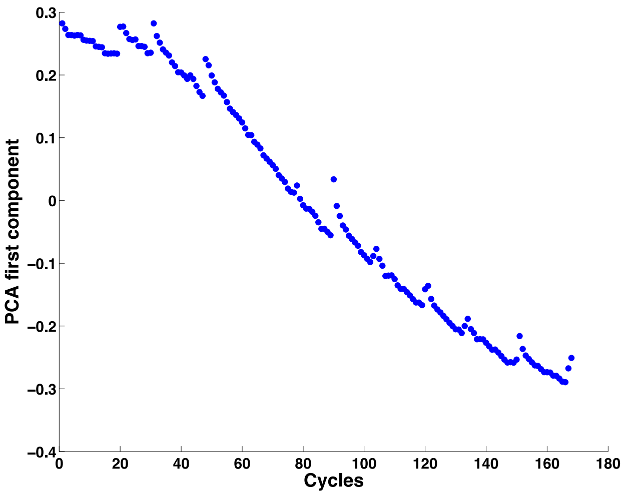
The method

Trend construction

- Using principle component analysis

$$C\lambda_i = \lambda_i v_i$$

  - Where λ_i is eigenvalue and v_i is eigenvectors for covariance matrix of C of the selected features.



- Linear projection

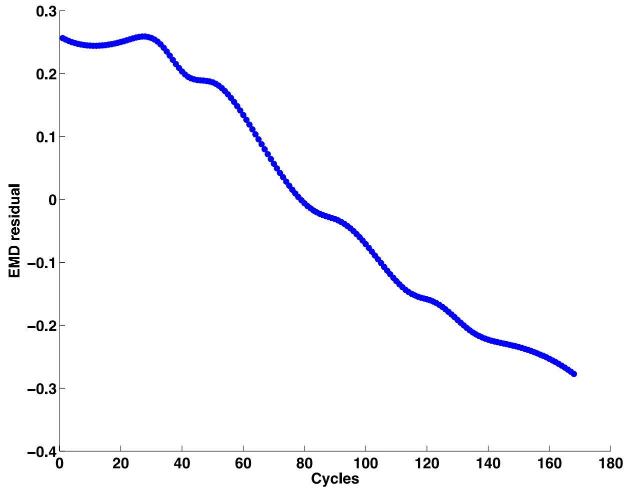
The method

Trend construction

- Using empirical mode decomposition

$$r_n(t) = X(t) - \sum_{i=1}^{n} imf_i\,(t)$$

  - Where *X(t)* is input signal, *imf* is intrinsic mode function and r(t) is residual.

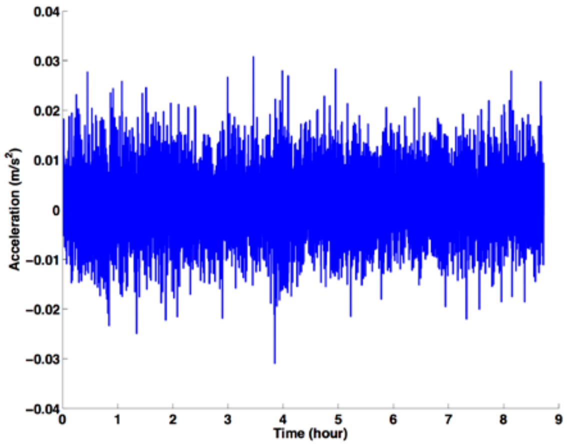


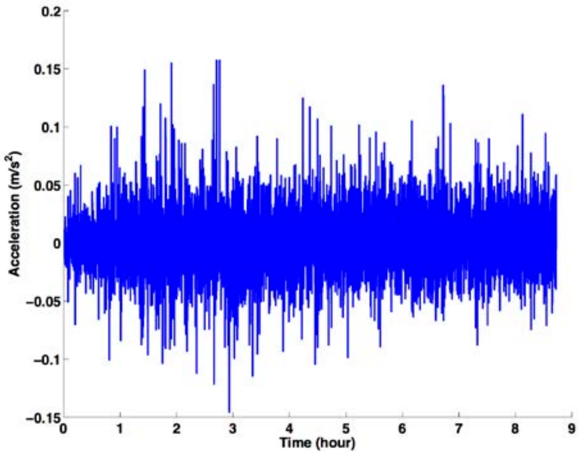
- The residual should be constant or monotonic function.

The method

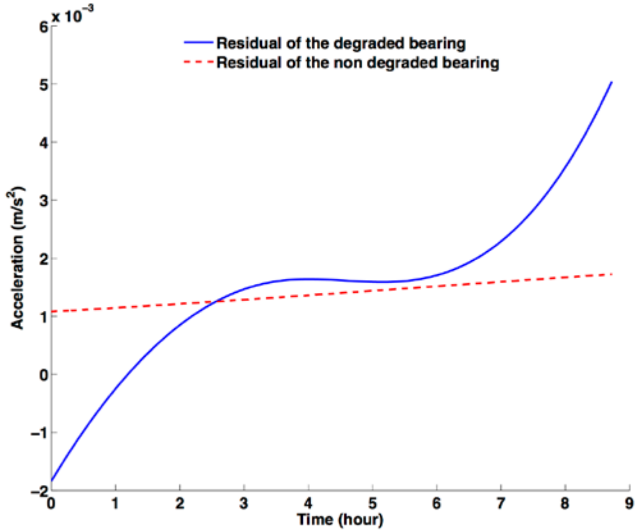
Trend construction

Non degraded



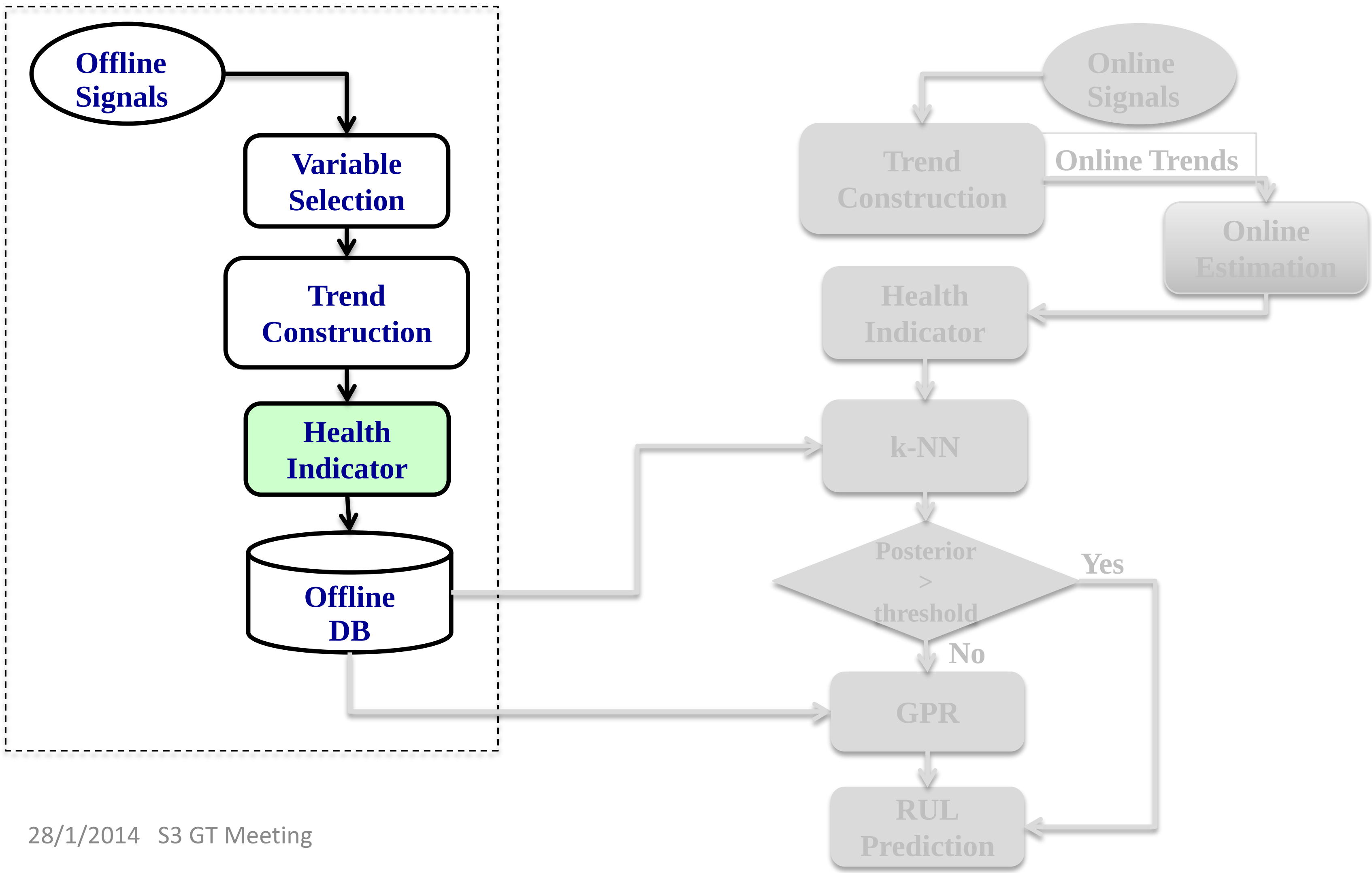


Degraded



The method

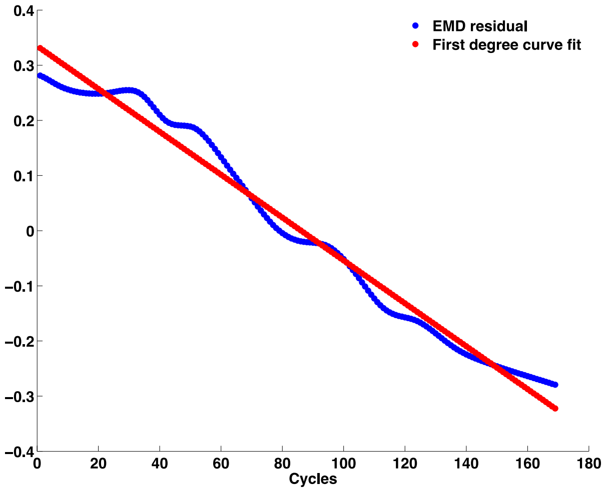
Overall scheme



The method

Health indicator

- 4 features are extracted
  - Two coefficients of a linear regression curve fit of the signal until time “t”.
  - Mean of the signal until time “t”.
  - Variance of the signal until time “t”.

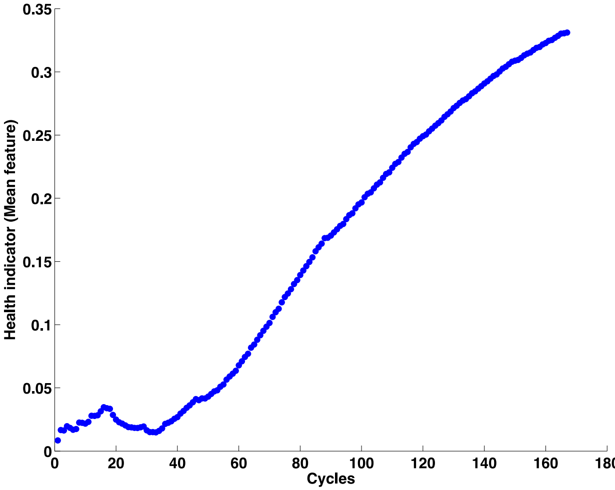


The method

Health indicator

- 4 features are extracted
  - Two coefficients of a linear regression curve fit of the signal until time “t”.
  - Mean of the signal until time “t”.
  - Variance of the signal until time “t”.

- The result is a health indicator.



The method

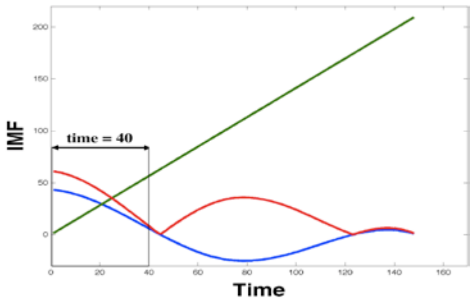
Health indicator

- 4 features are extracted
  - Two coefficients of a linear regression curve fit of the signal until time “t”.
  - Mean of the signal until time “t”.
  - Variance of the signal until time “t”.

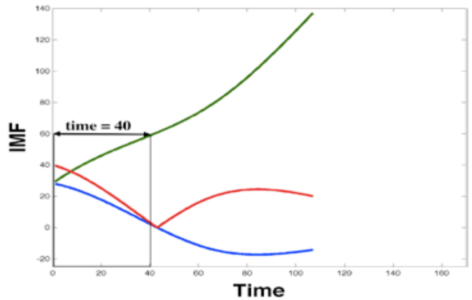
- The result is a health indicator.

- The features are labeled according to the EOL of each trend.

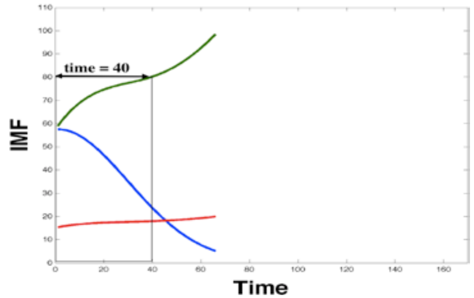
EOL = 148



EOL = 107

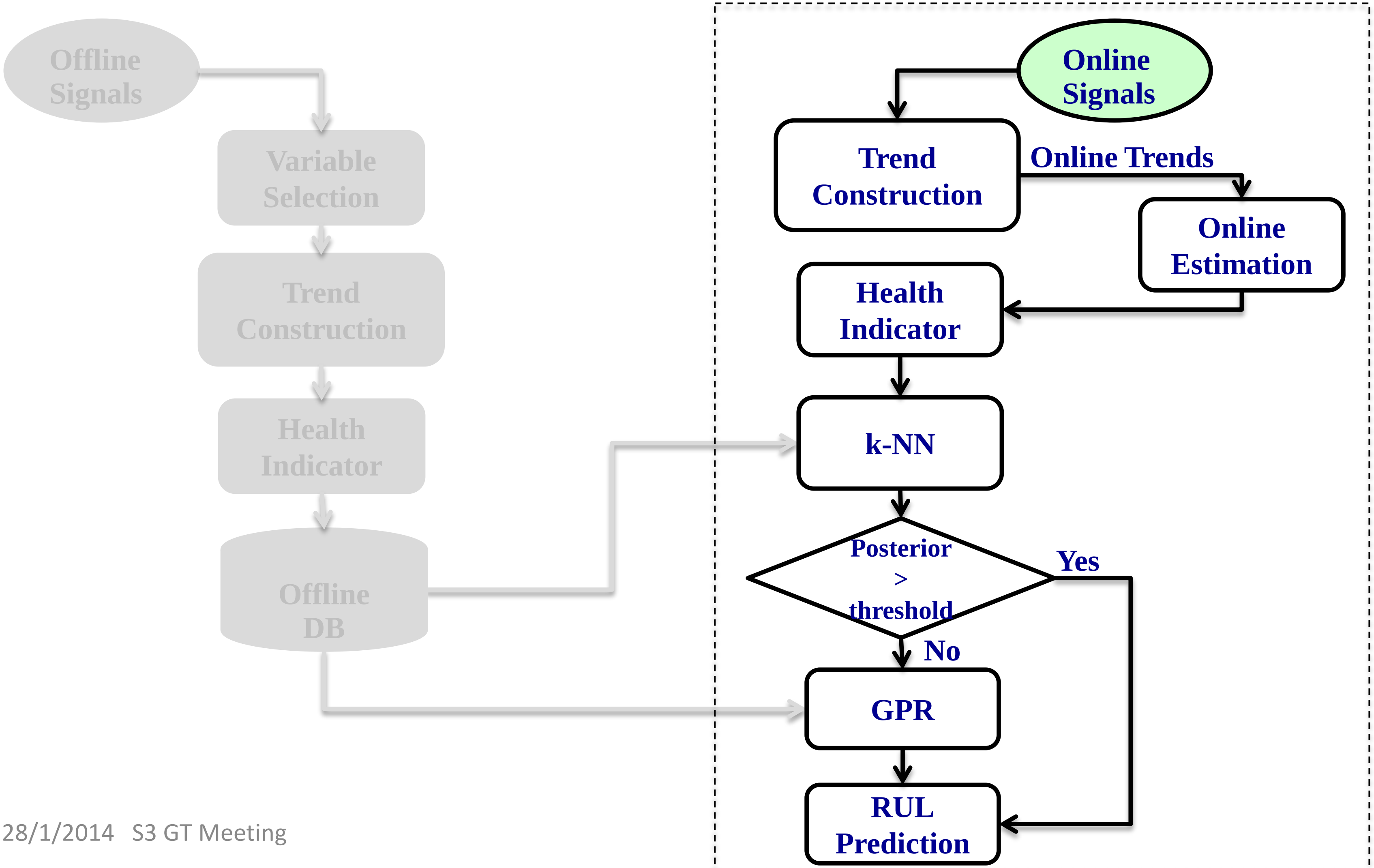


EOL = 66



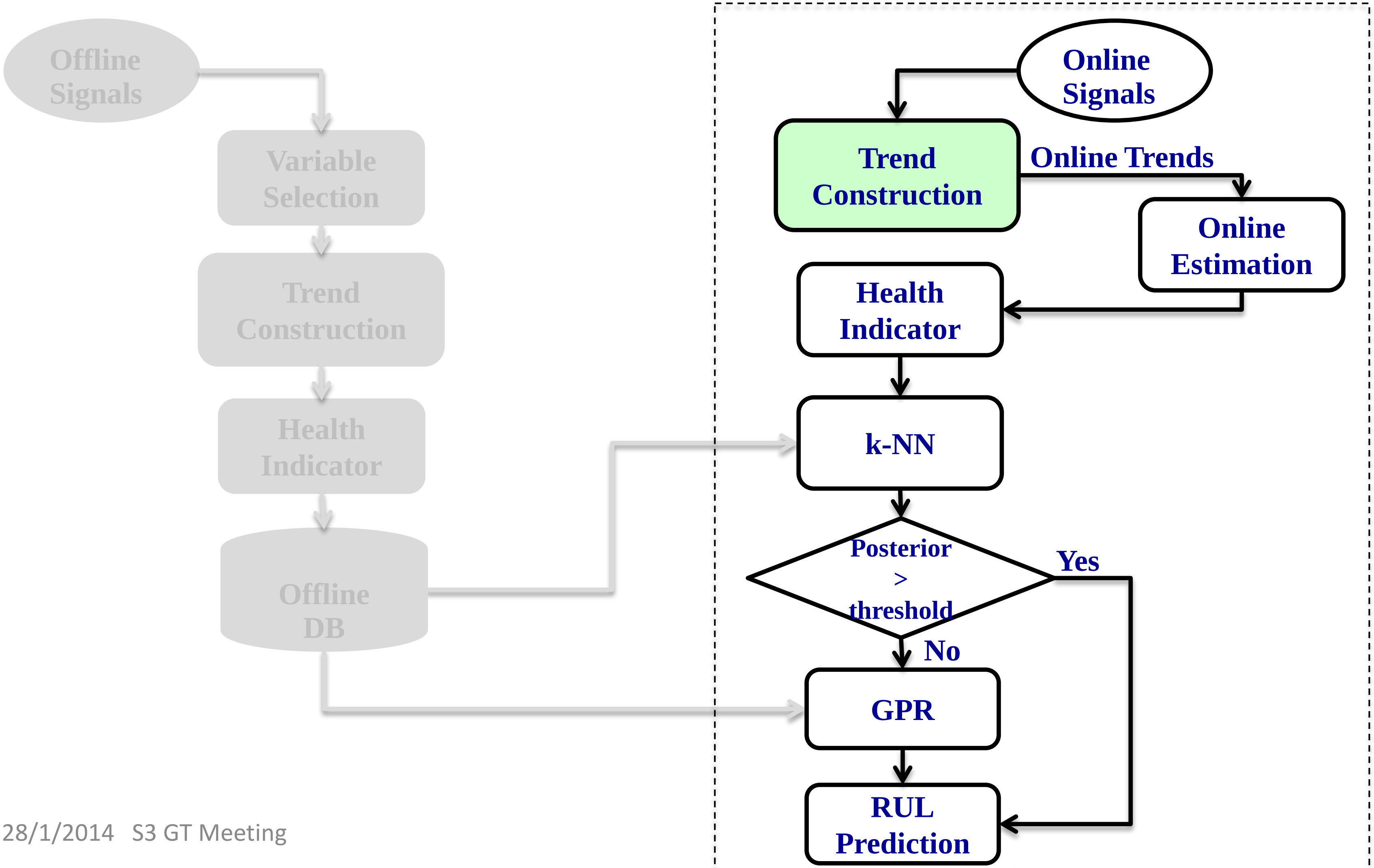
The method

Overall scheme



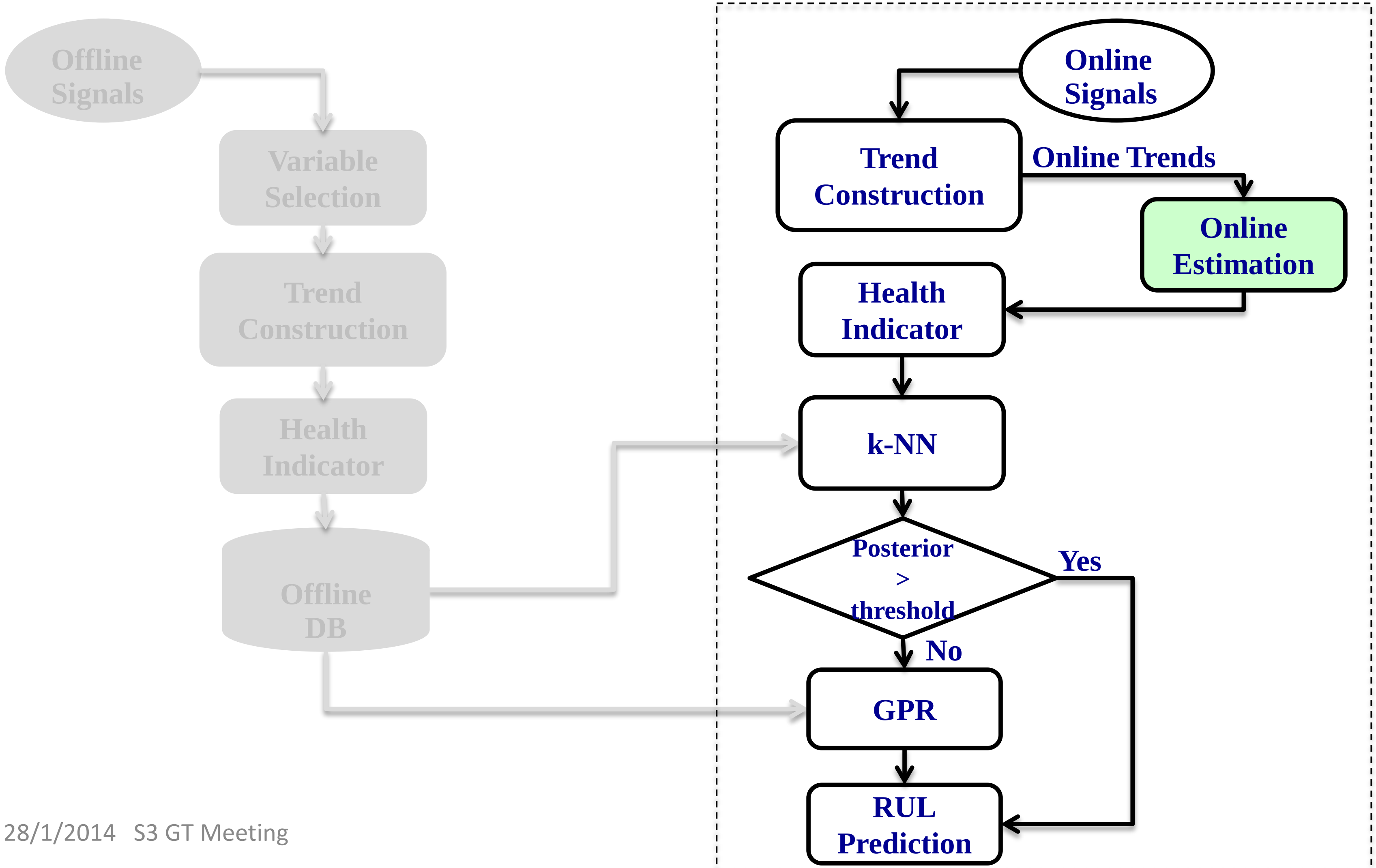
The method

Overall scheme



The method

Overall scheme



The method

<u>State estimation: Bayes filter</u>

- Critical components are dynamic systems that possess internal state which can characterize the system health.
- Internal state can not be measured directly.
- Sensory data are used to deduce the internal state.
- The evolution of the state and measurements are governed by probabilistic laws.

The method

State estimation: Bayes filter

- Markovian internal state will be denoted as x_t and the sensory data are denoted as z_t at time t.
- Two main quantities need to be observed:
  - State transition probability: $p(x_t \mid x_{t-1})$
  - Measurement transition probability: $p(z_t \mid x_t)$
- Probability over state variable x_t will be denoted as:
  - $p_{posterior}(x_t) = p(x_t \mid z_{1:t})$.
- Prediction probability distribution denoted as:
  - $p_{prior}(x_t) = p(x_t \mid z_{1:t-1})$

The method

State estimation: Bayes filter

- Bayes filter: a general algorithm for calculating the prior and posterior probabilities.

$$
\begin{aligned}
&Input: p_0(x_{t-1}), z_t \\
&Output: p_{posterior}(x_t) \\
&\forall x_t \\
&\quad p_{prior}(x_t) = \int p(x_t \mid x_{t-1}) p(x_{t-1})\, dx \\
&\quad p_{posterior}(x_t) = \eta\, p(z_t \mid x_t)\, p_{prior}(x_t) \\
&end
\end{aligned}
$$

- Estimates the probability distribution recursively from the data.

The method

State estimation: Bayes filter

- Can be implemented in different ways such as:
  - Kalman filter.
  - Extended Kalman filter.
  - Particle filter.
  - Histogram filter.

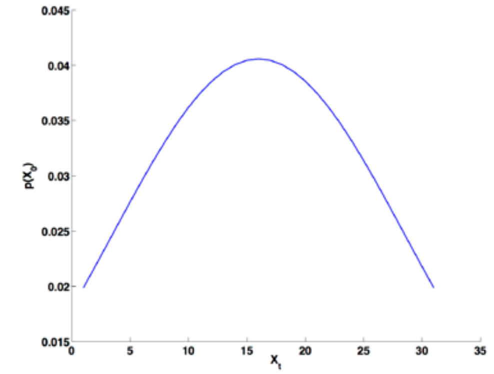
- Any implementation requires knowing three probability distributions:
  1. Initial probability.
  2. Measurement transition probability.
  3. State transition probability.

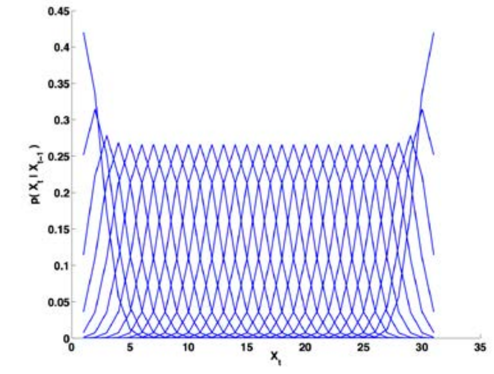
The method

State estimation: Bayes filter

- Known as histogram filter for continuous states.

$$
\begin{aligned}
&Input : \{p_{k,t-1}\}, z_t \\
&Output : \{p_{k,t}\} \\
&\forall x_t \\
&\quad \bar{p}_{k,t} = \sum_i p(X_t = x_k \mid X_{t-1} = x_i) p_{i,t-1} \\
&\quad p_{k,t} = \eta p(z_t \mid X_t = x_k) \bar{p}_{k,t} \\
&end
\end{aligned}
$$

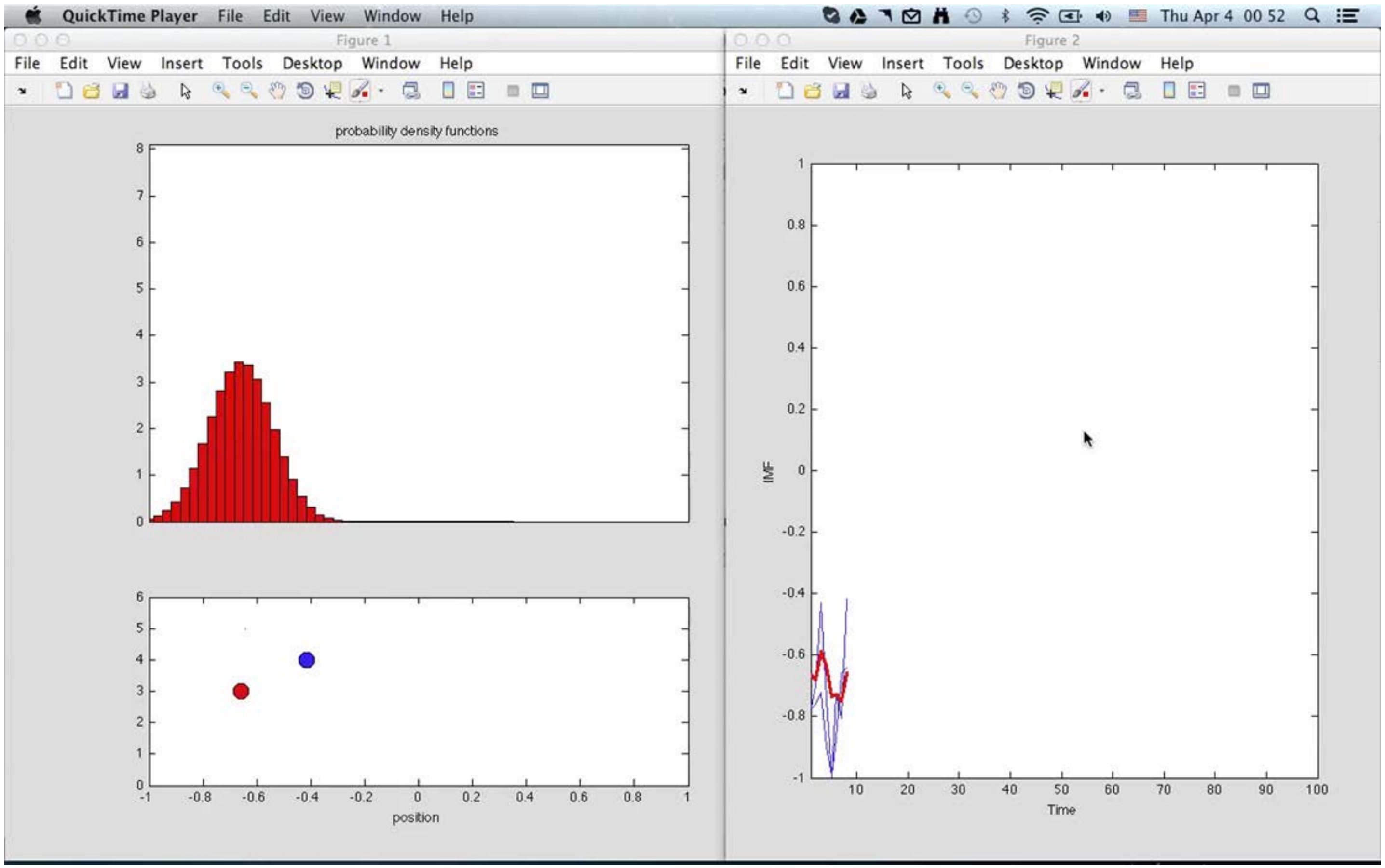




- Decomposes the state space into many regions and represents the posterior by a histogram.

The method

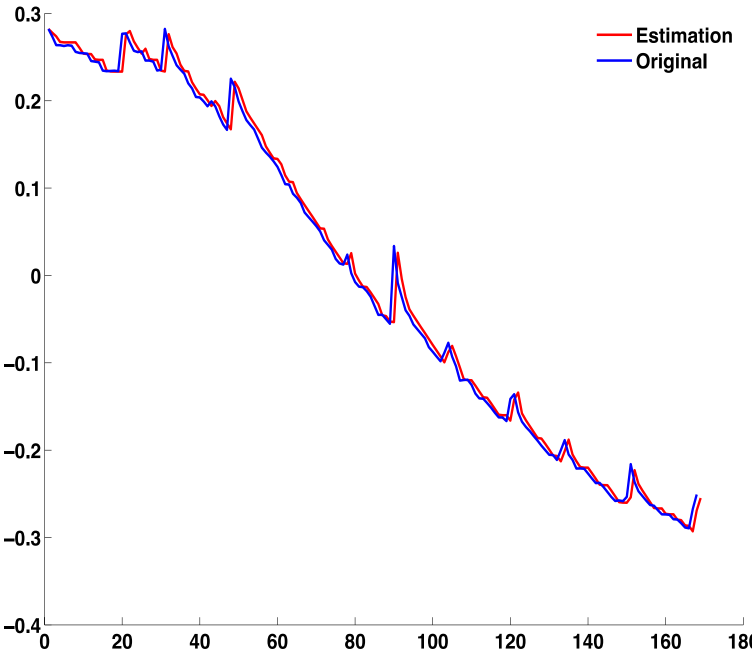
State estimation: Bayes filter



The method

State estimation: Bayes filter

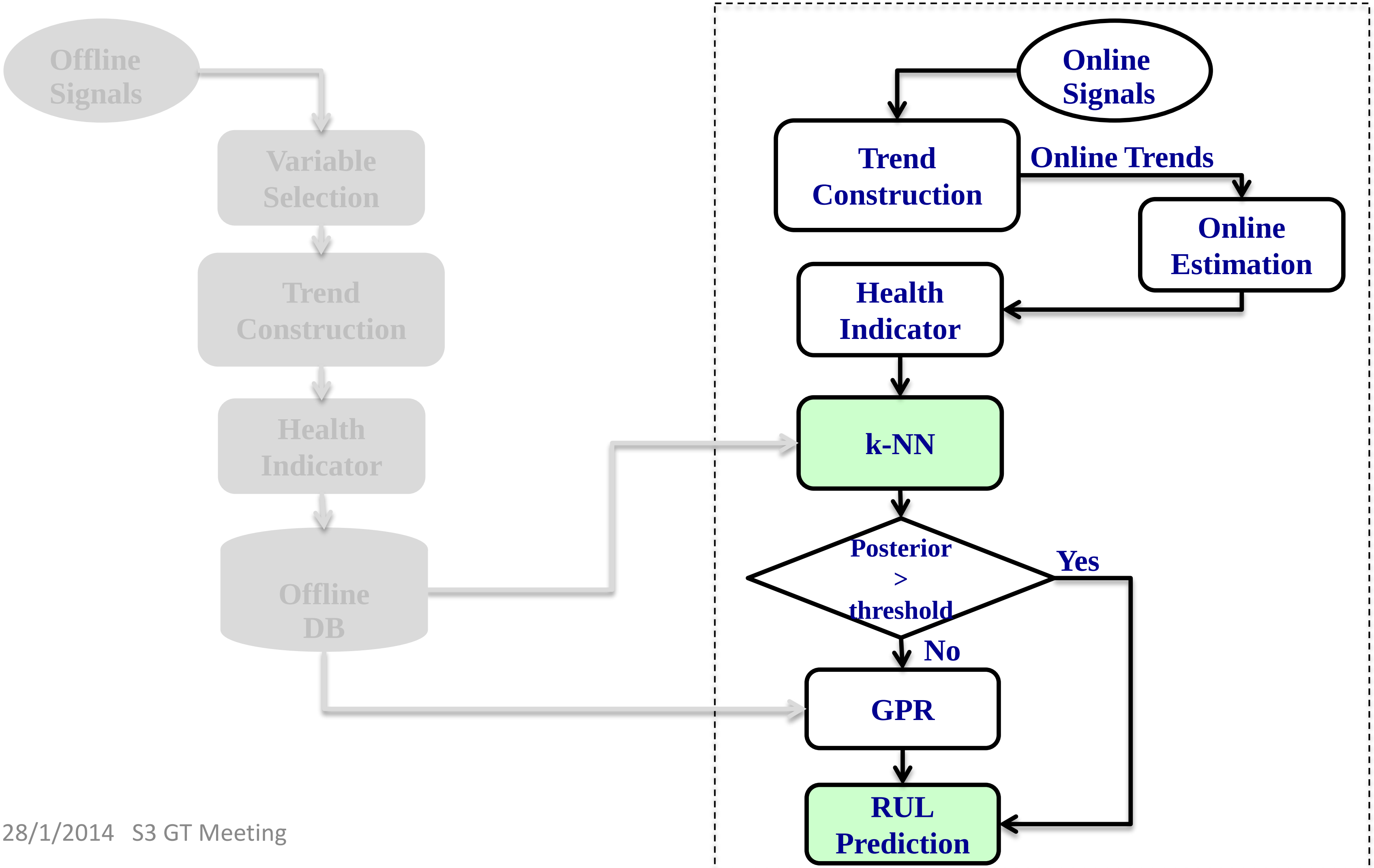
- Example of estimating the trend of the projected capacity and voltage variables at discharge.



- RMS = 0.0148

The method

Overall scheme

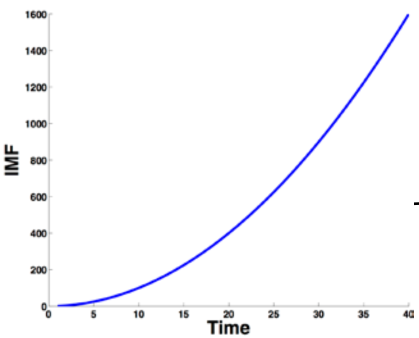


The method

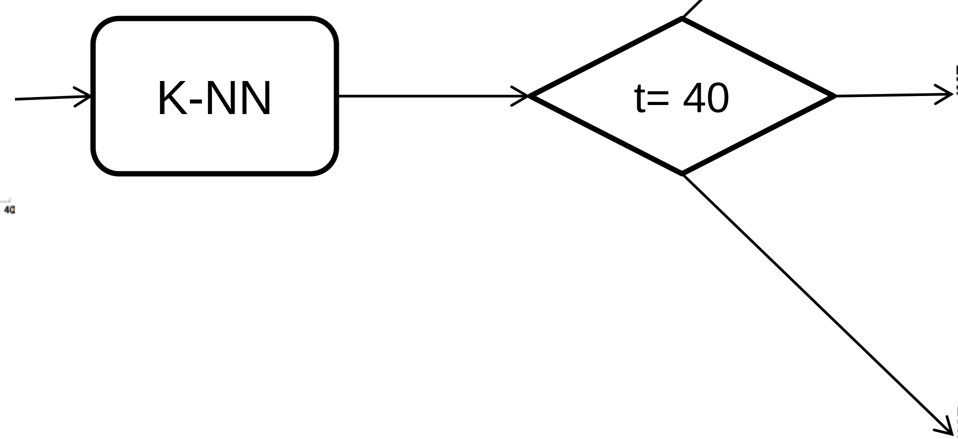
Online selection (k-NN)

- A k-NN classifier for objects based on closest training examples in the feature space.

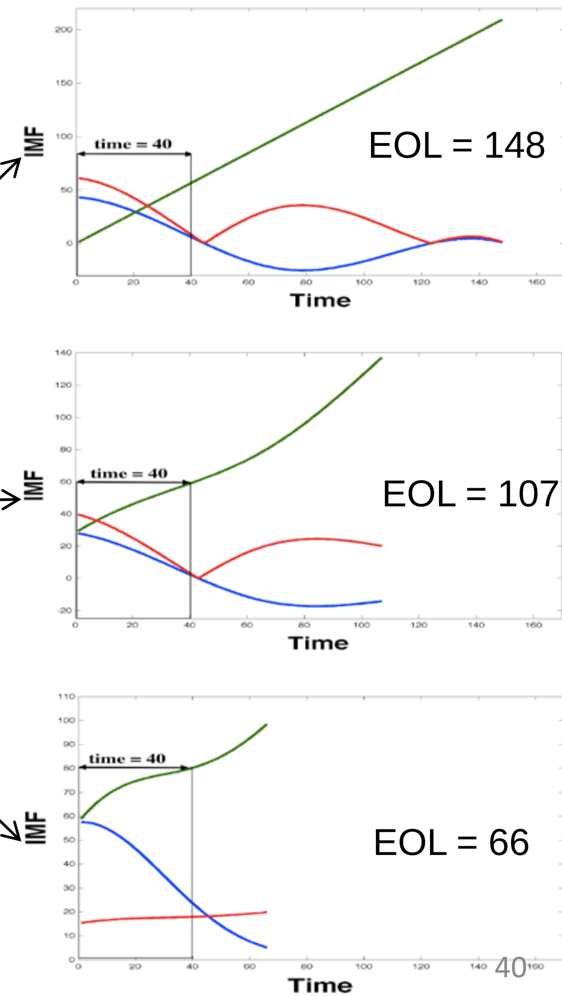
$$p(C_k|a) = \frac{p(a|C_k) \times p(C_k)}{p(a)} = \frac{K_k}{K}$$



Online trend at time = 40



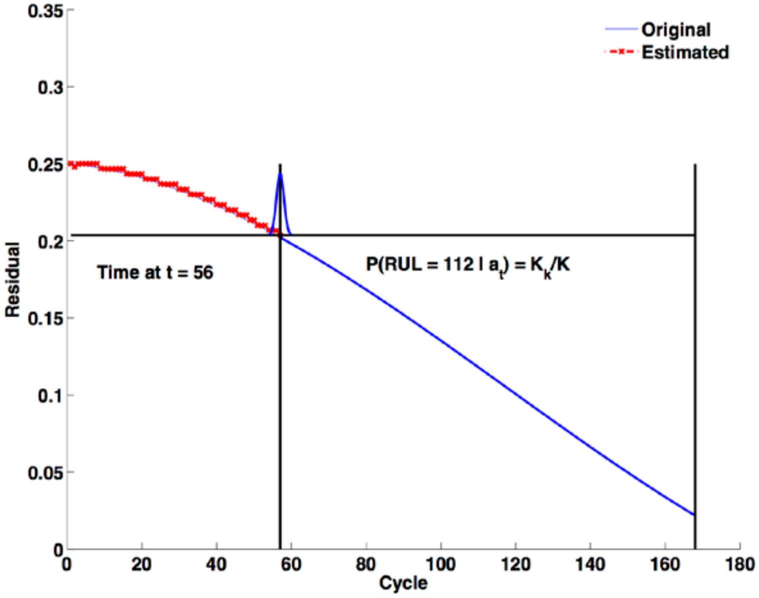
Offline trends



The method

RUL prediction

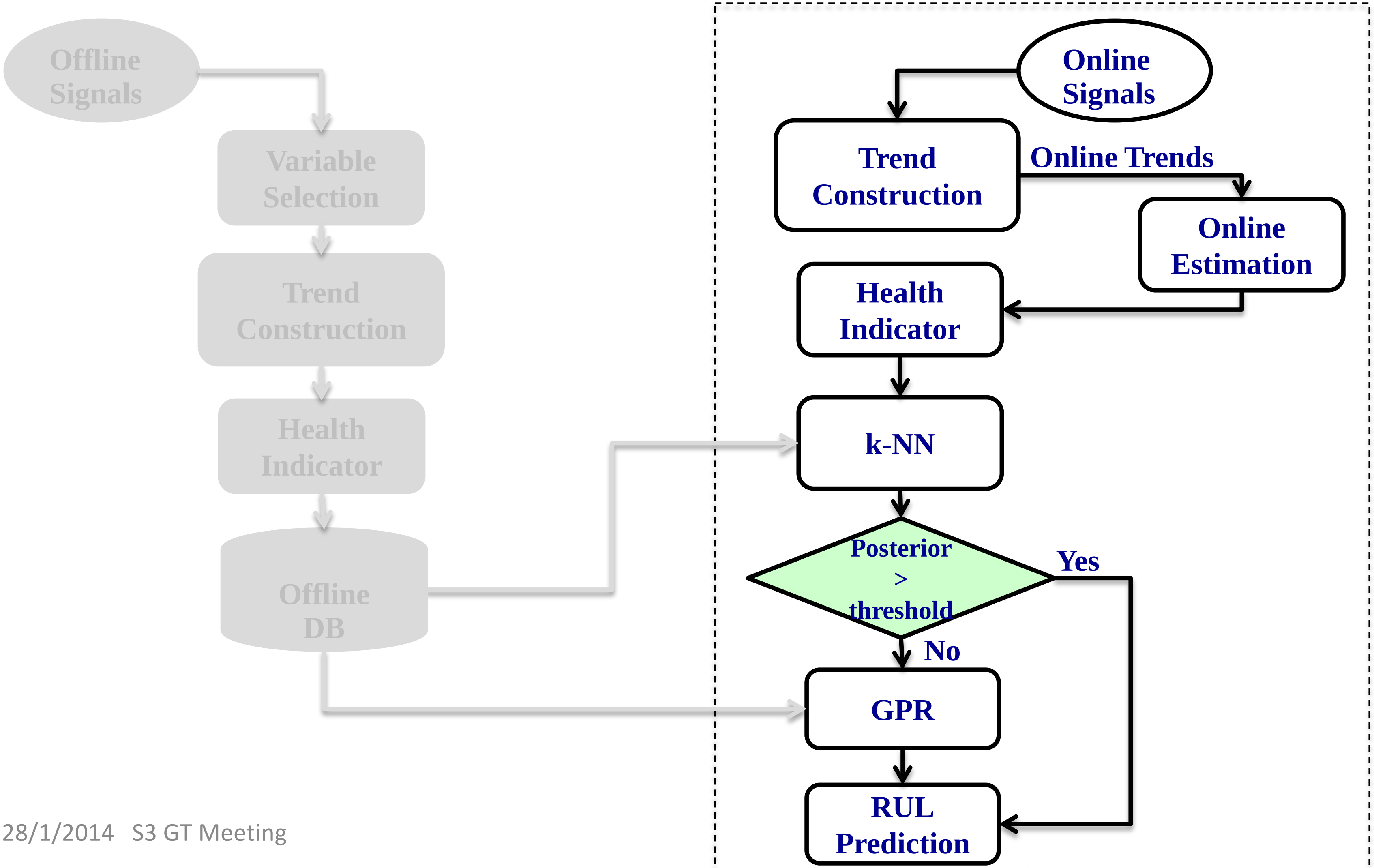
- Using Euclidean distance.
  - Between online signal and database till time “t”.
  - Efficient for this particular problem.



$$d(q-p)=\sqrt{(q_1-p_1)^2+(q_2-p_2)^2+....+(q_n-p_n)^2}=\sqrt{\sum_{i=1}^{n}(q_i-p_i)^2}$$

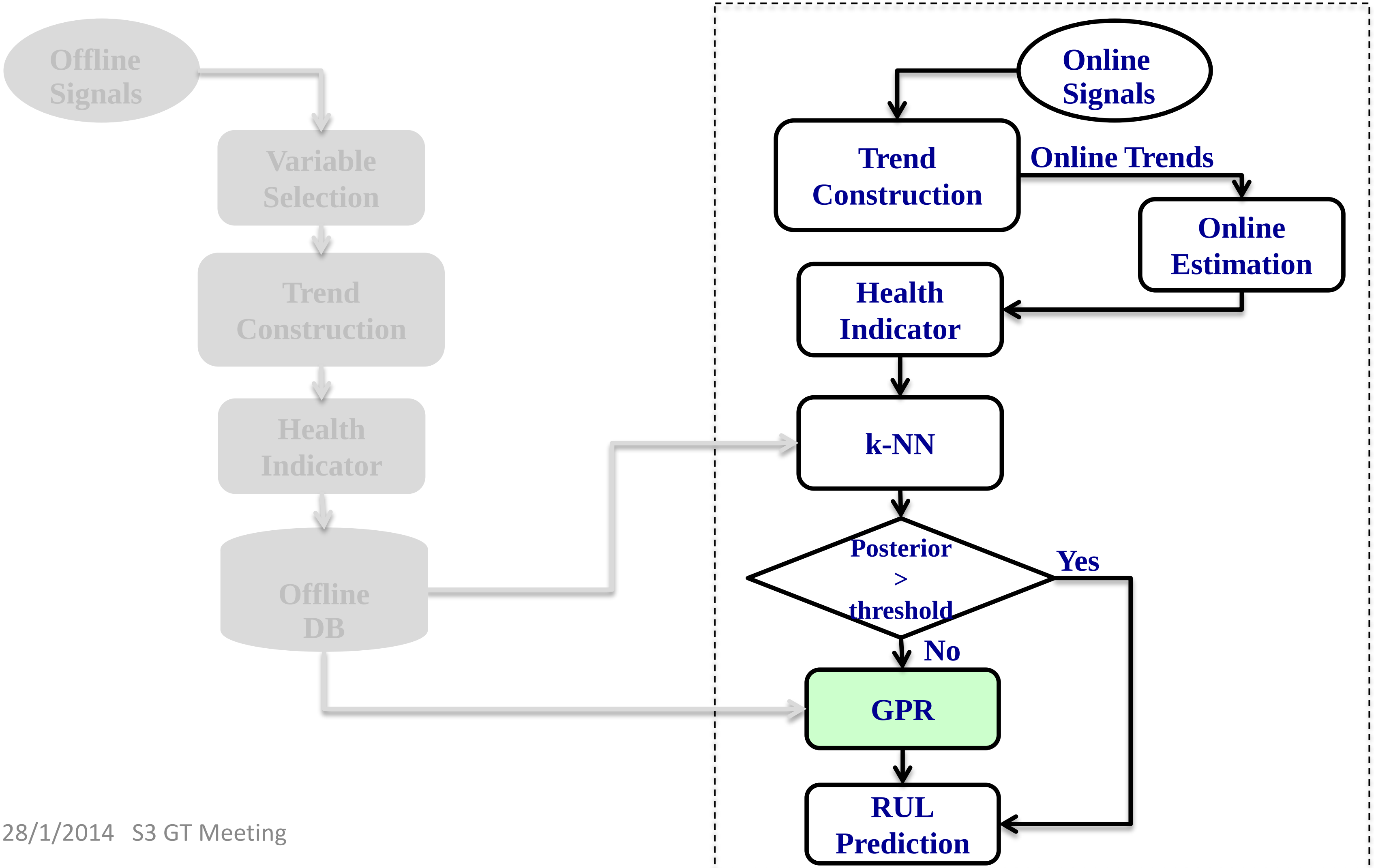
The method

Overall scheme



The method

Overall scheme

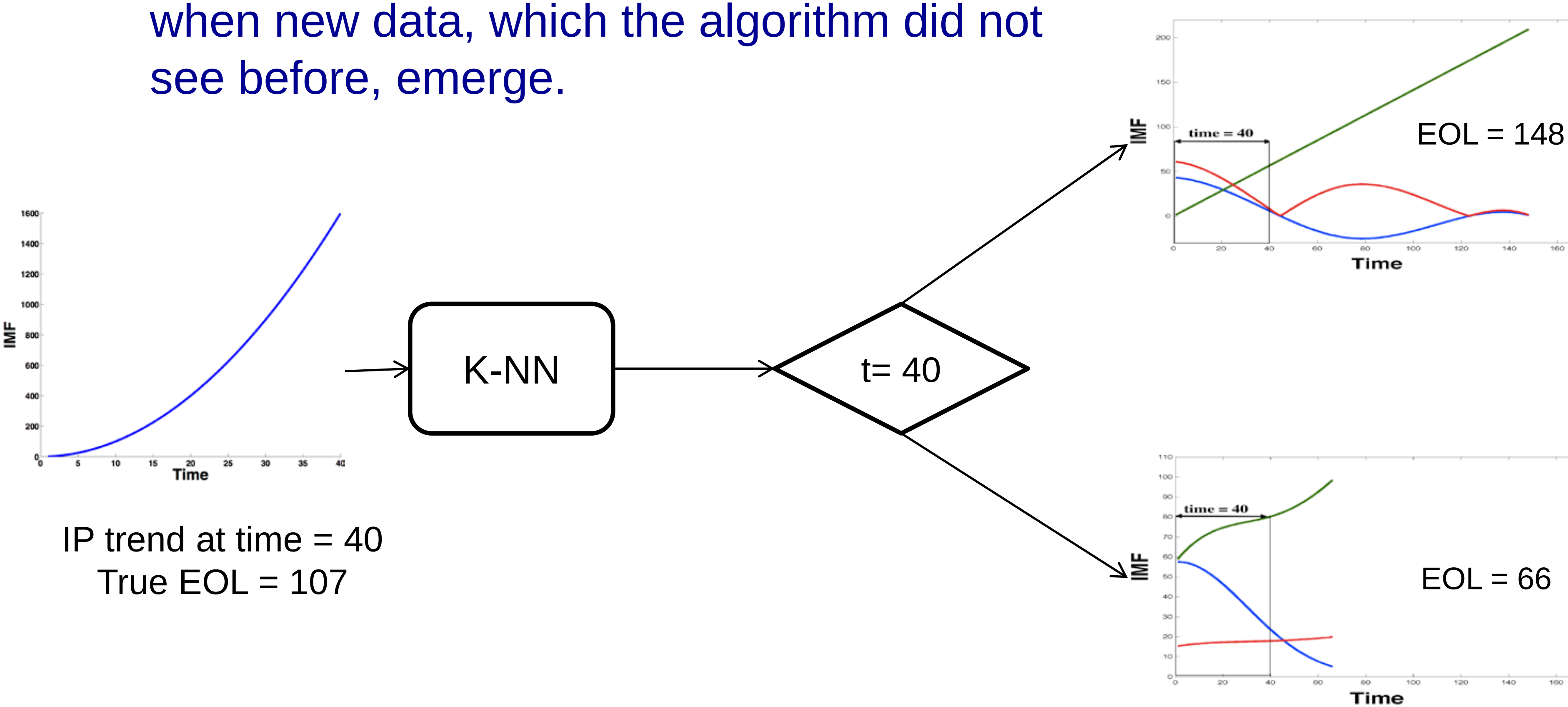


The method

Gaussian process regression (GPR)

- The classification error tends to be very high when new data, which the algorithm did not see before, emerge.

Offline trends



The method

Gaussian process regression (GPR)

- To map from online samples ***x*** to the most similar group of offline trends ***y*** GPR is proposed.

- GPR defines the prior for output ***f(x)*** in form of distribution over functions specified by Gaussian process (GP) and Gaussian noise:

$$y = f(x) + N(0, \sigma_n^2)$$

- GP function ***f(x)*** is specified by a mean function ***m(x)*** and covariance function ***k(x,x´)*** collected for all possible pairs of the input vector ***x***.

The method

Gaussian process (GP)

- The posterior probability distribution can be written as:

$$\begin{bmatrix} y \\ y^* \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} \mu \\ \mu^* \end{bmatrix}, \begin{bmatrix} K & K^{*T} \\ K^* & K^{**} \end{bmatrix} \right)$$

- The best estimate for y^*, i.e. RUL, is the mean of this distribution

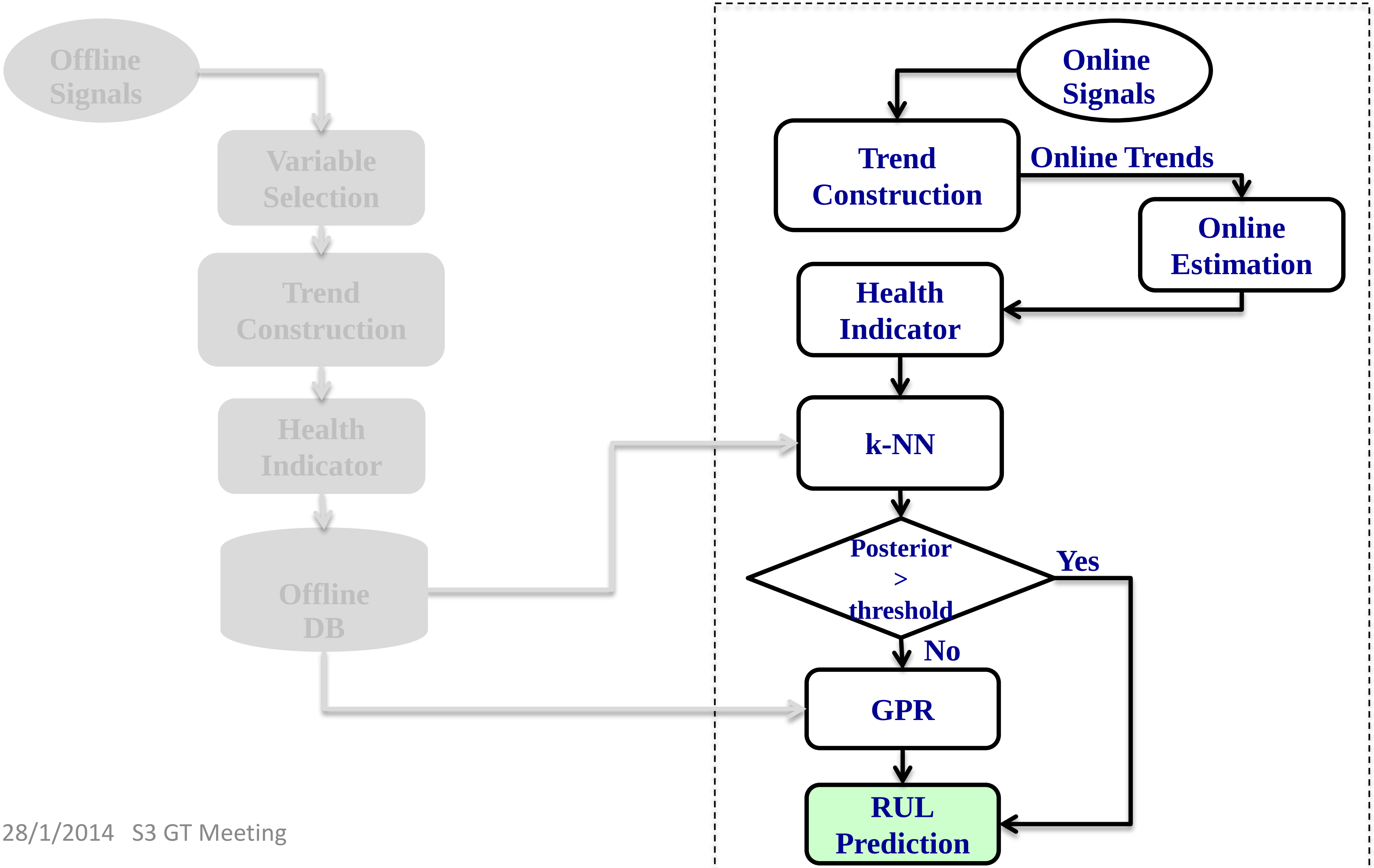
$$\overline{y^*} = \mu^* + K^* K^{-1}(y - \mu)$$

- The uncertainty in the estimate is represented in the variance.

$$\mathrm{var}(y^*) = K^{**} - K^* \mathrm{K}^{-1} K^{*T}$$

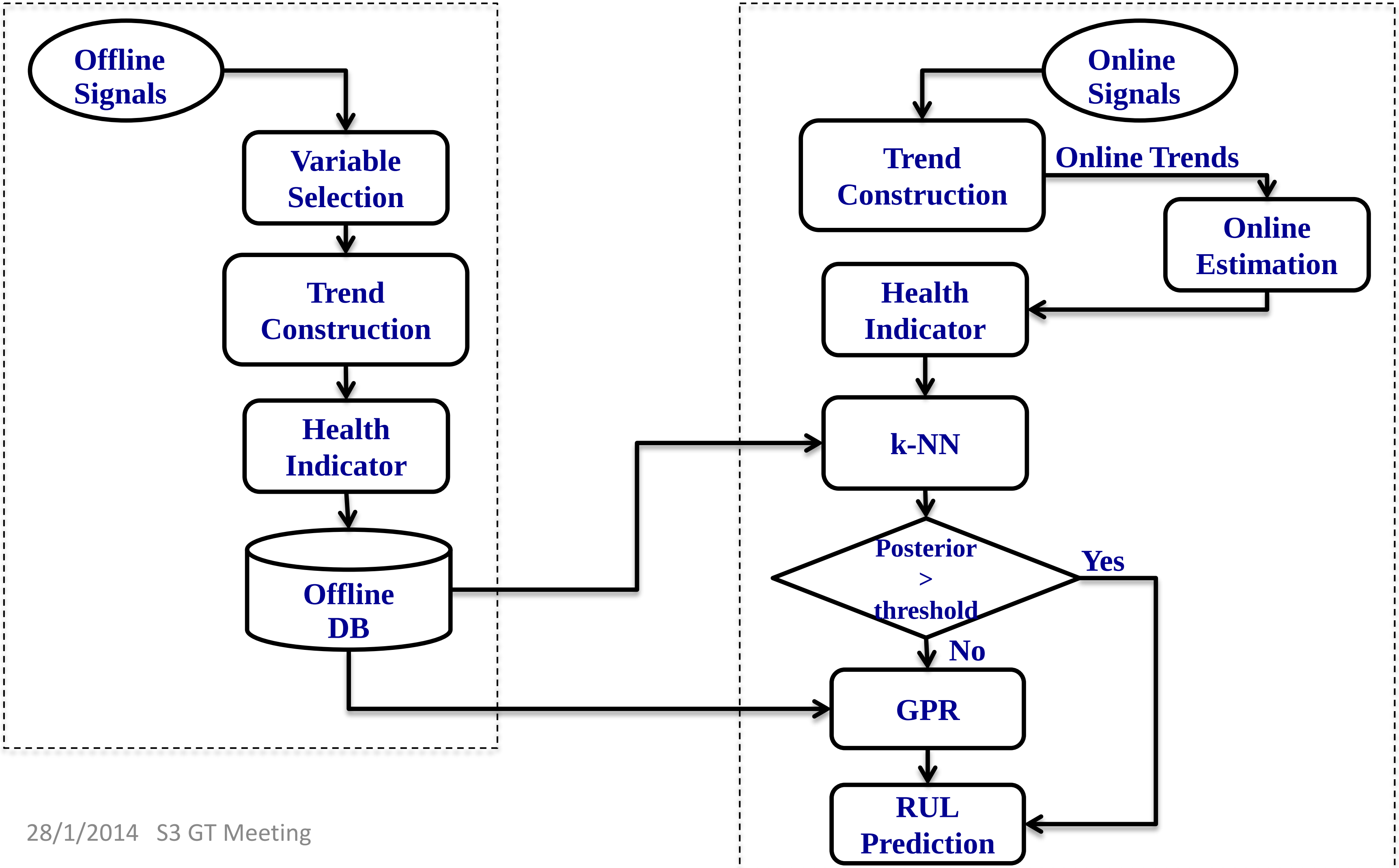
The method

Overall scheme



The method

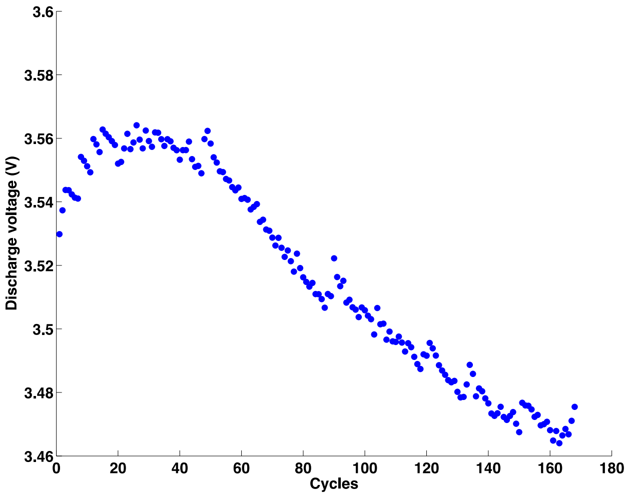
Overall scheme

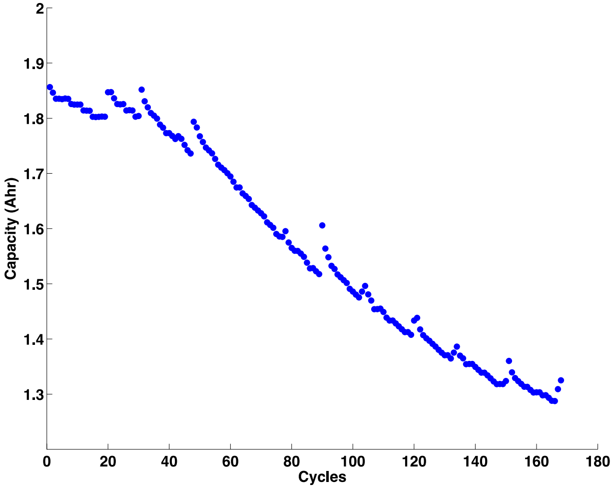


Applications and results

NASA batteries

- 34 datasets are used
- 11 signals are used in the experiment
  - 5 charging.
  - 6 discharging.
- Two interesting relations are selected
  - Discharging voltage
  - Capacity

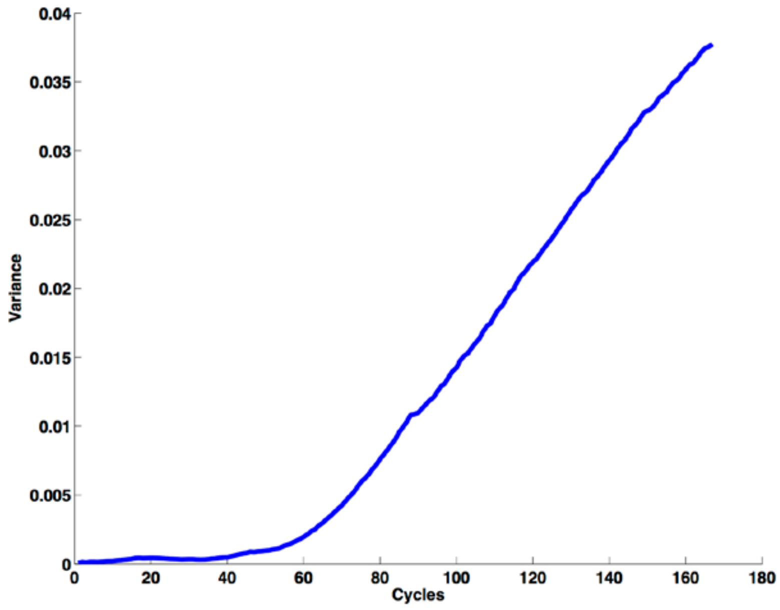


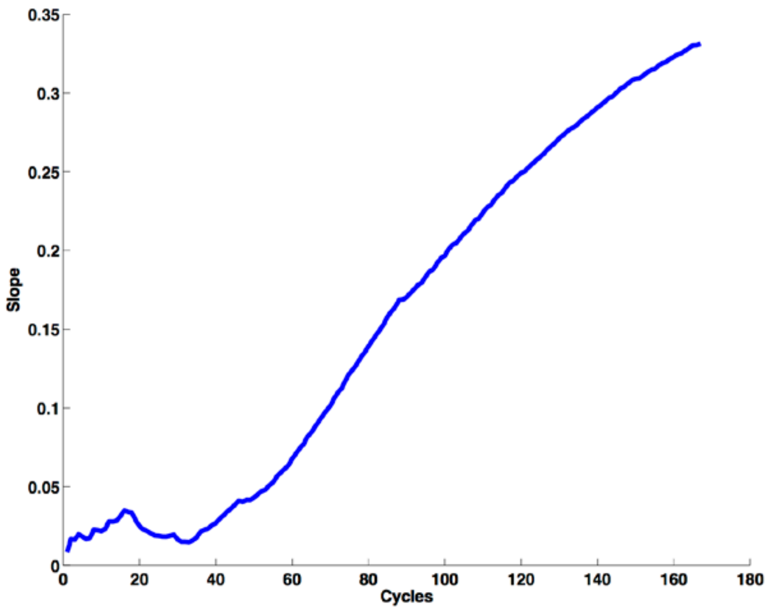


Applications and results

NASA batteries

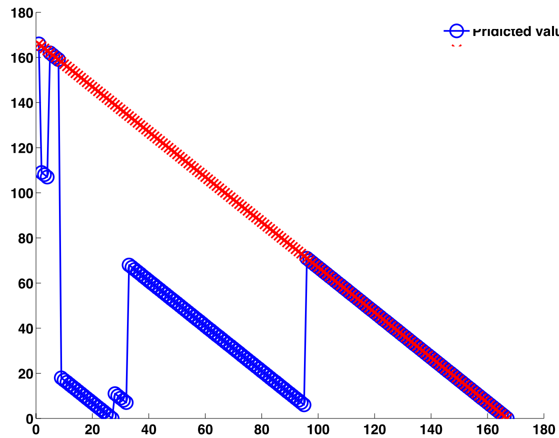
- Two health indicators correlated with the capacity
  - Variance
  - Slope

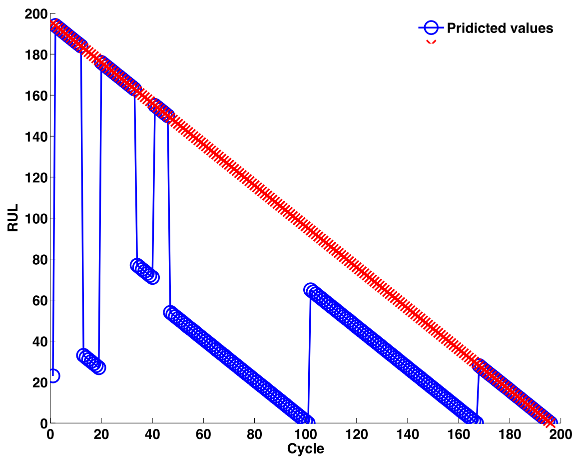


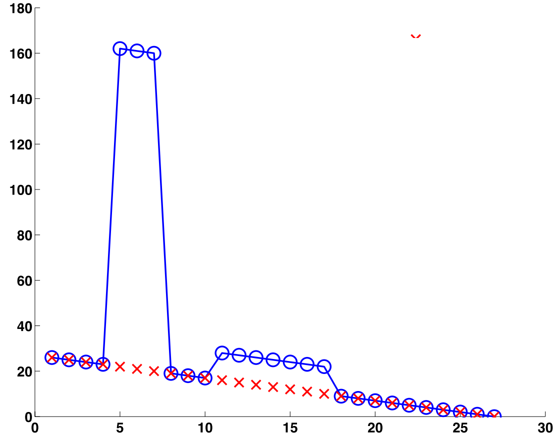


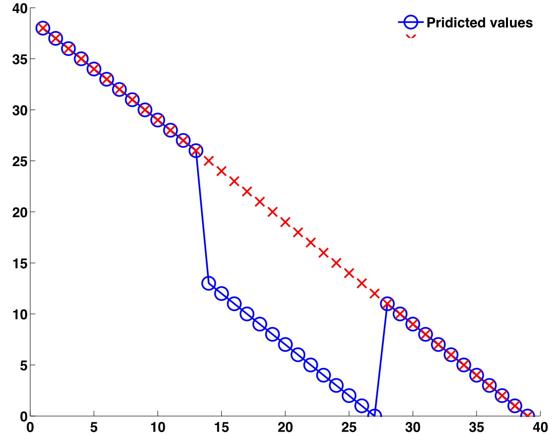
Applications and results

NASA batteries









Applications and results

NASA batteries

- Three fold cross validation.
- Percentage error calculated.
- Three data sets were unique.

	Fold #1	Fold #2	Fold #3	Total error (Avg.)
Without unique data	27.11%	23.73%	18.01	22.94%
With unique data	38.75%	37.11%	36.89%	37.58%

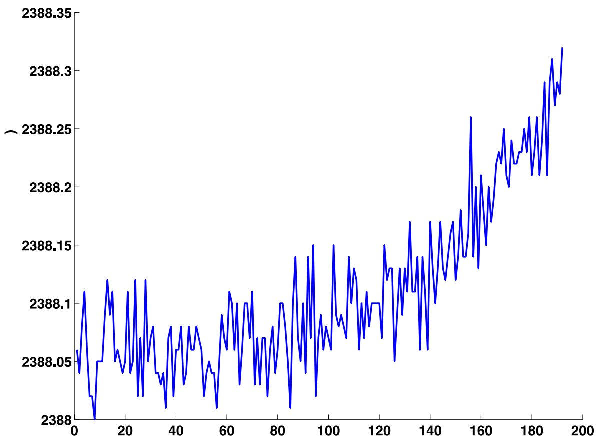
Applications and results

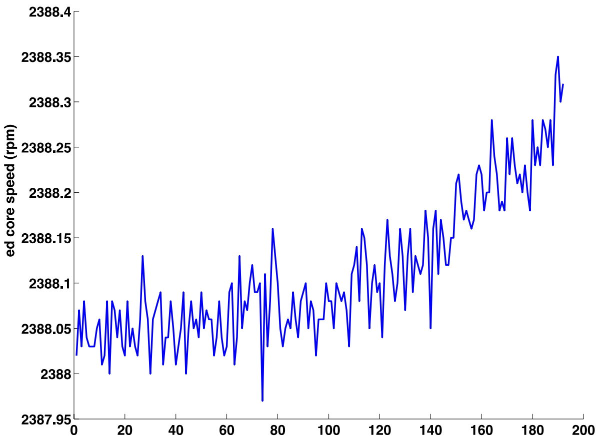
NASA Turbofan engine data

- 2 datasets are used
  - 100 engine data for training
  - 100 engine data for testing

- 21 signals are used in the experiment
  - Total temperature at fan inlet
  - Pressure at fan inlet
  - Demanded fan speed ….

- Two interesting relations are selected
  - Physical core speed
  - Corrected core speed



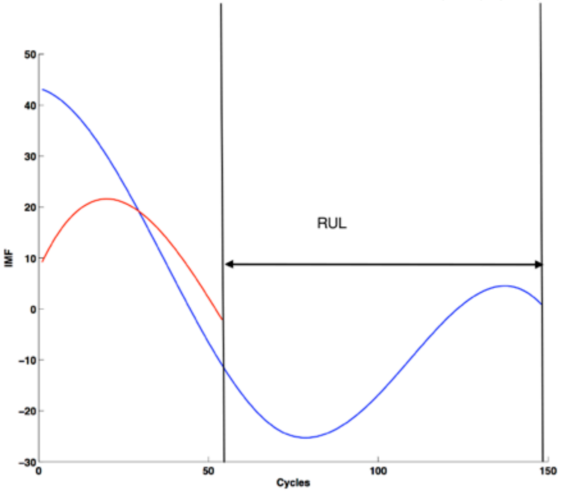


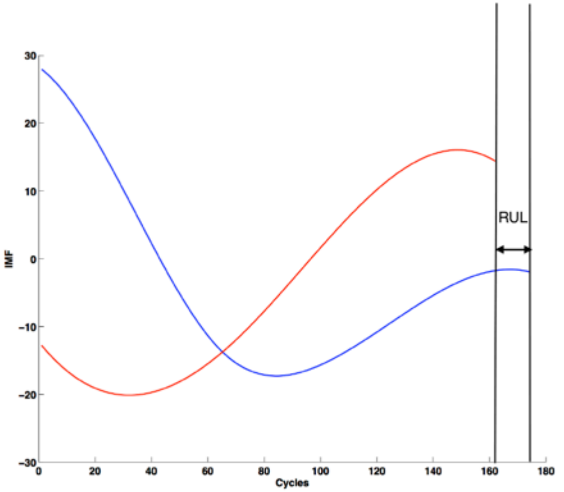
Applications and results

NASA turbofans

- All training data used for training and testing for test.
- Percentage error calculated.
- Only one prediction at the pre-specified critical time.

k-NN	GPR	Integrated
12.74%	11.48%	11.41%





Conclusion & future work

- A data-driven prognostic method for condition assessment and RUL prediction is proposed.
- The method can be identified as direct RUL Bayesian learning approach.
- The uncertainty about the measurements and the predictions represented by conditional probability.
- Two health indicators shown to be correlated with degradation mechanism.

Conclusion & future work

- Apply the method on data sets with variable operating conditions.
- Test the method after introducing maintenance interventions.
- Test the proposed method on new application.
- Explore other classification/regression models.

References

[1] ISO 13381-1. Condition monitoring and diagnostics of machines- Prognostics- Part1: General guidelines, 2004.

[2] D.A. Tobon-Mejia, K. Medjaher, and N. Zerhouni. CNC machine tool's wear diagnostics and prognostics by using dynamic bayesian networks. Mech. Syst. and Signal Proc., 28: 167-182, 2012.

[3] Enrico Zio and Francesco Di Maio. A data-driven fuzzy approach for predicting the remaining useful life in dynamic failure scenarios of a nuclear system. Rel. Eng. and Safety Syst., 95: 49-57, 2010.

[4] Andrew K.S. Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20: 1483-1510, 2006

[5] B. Saha and K. Goebel (2007). "Battery Data Set", NASA Ames Prognostics Data Repositoryhttp://ti.arc.nasa.gov/project/prognostic-data-repository], NASA Ames, Moffett Field, CA

[6] A. Saxena and K. Goebel (2008). "C-MAPSS Data Set", NASA Ames Prognostics Data Repository,http://ti.arc.nasa.gov/project/prognostic-data-repository], NASA Ames, Moffett Field, CA.

Questions?