Predicting WCET Trends in Long-lived Real-time Applications

22nd International Conference on Reliable Software Technologies
Ada-Europe 2017

Xiaotian Dai and Alan Burns

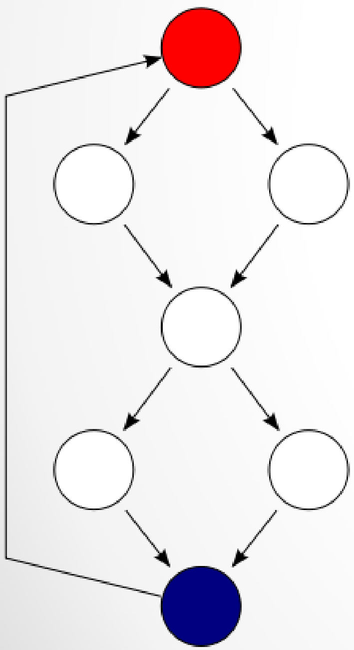
Real-time Systems Group

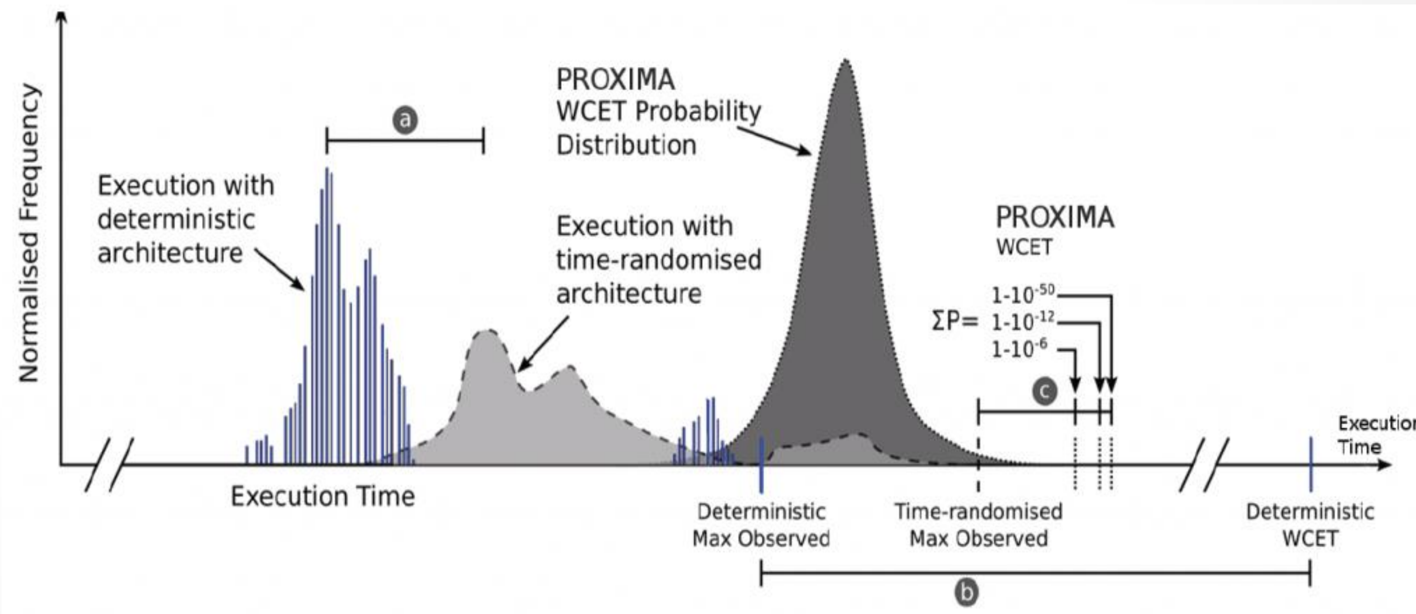
University of York, UK

Contents

Background | Motivation

- Worst-case execution time (WCET)
  - is important in timing analysis (DO-178 and ISO26262)
  - static and measurement-based
  - pWCET





[1] Wilhelm, Reinhard et al. "The worst-case execution-time problem - overview of methods and survey of tools." (2008).

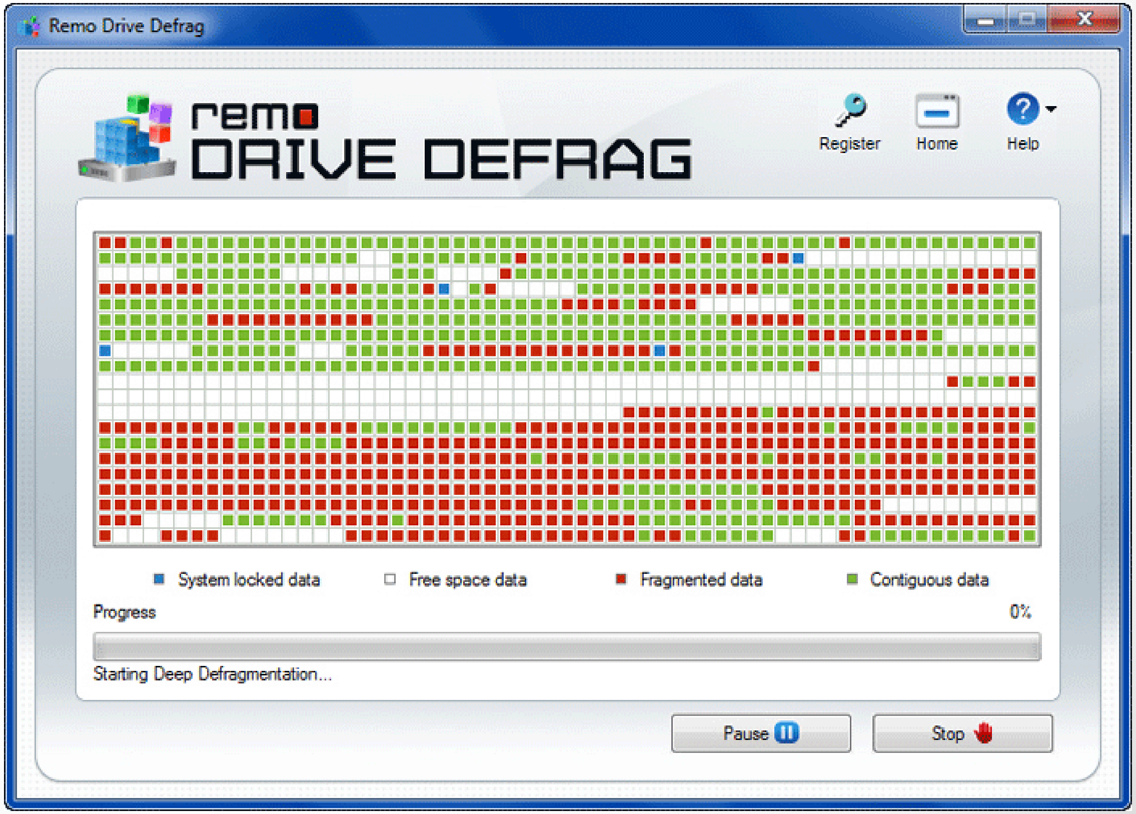
Background | Motivation

- Current understanding of WCETs:
  - A theoretical boundary exists, if designed and programmed with constrained models.
  - Known as a static, upper-bound value of execution times

[1] Wilhelm, Reinhard et al. "The worst-case execution-time problem - overview of methods and survey of tools." (2008).

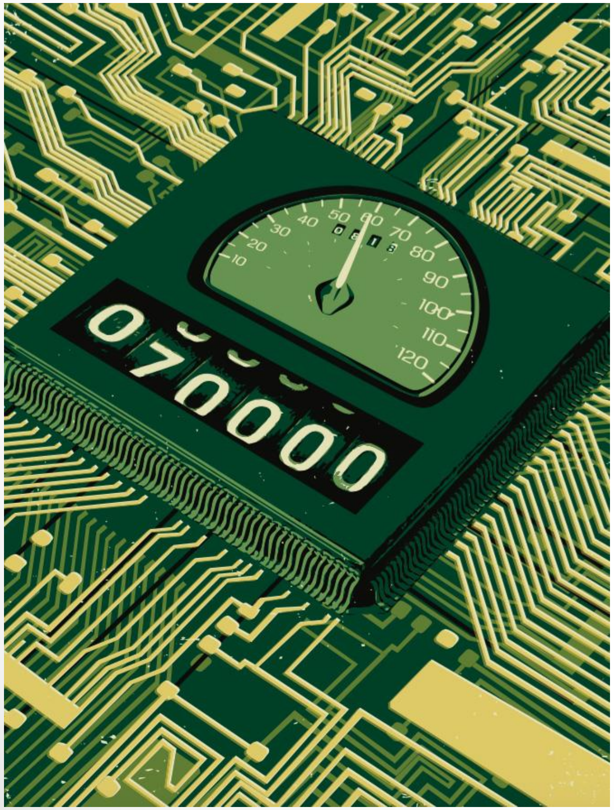
Issues | Motivation

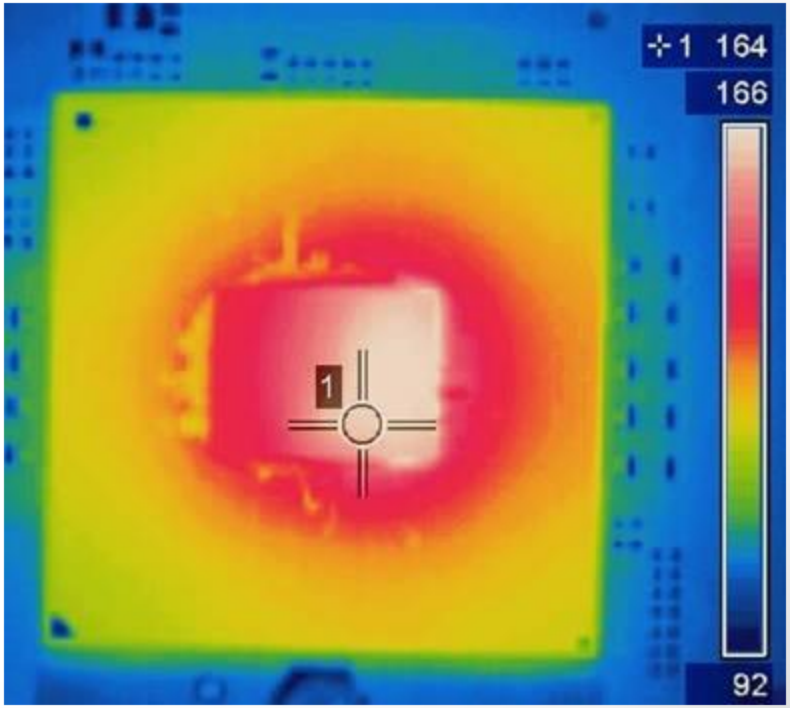
- Data accessing time ↑
  - relevant data growth
  - hard disk fault/fragmented



Issues | Motivation

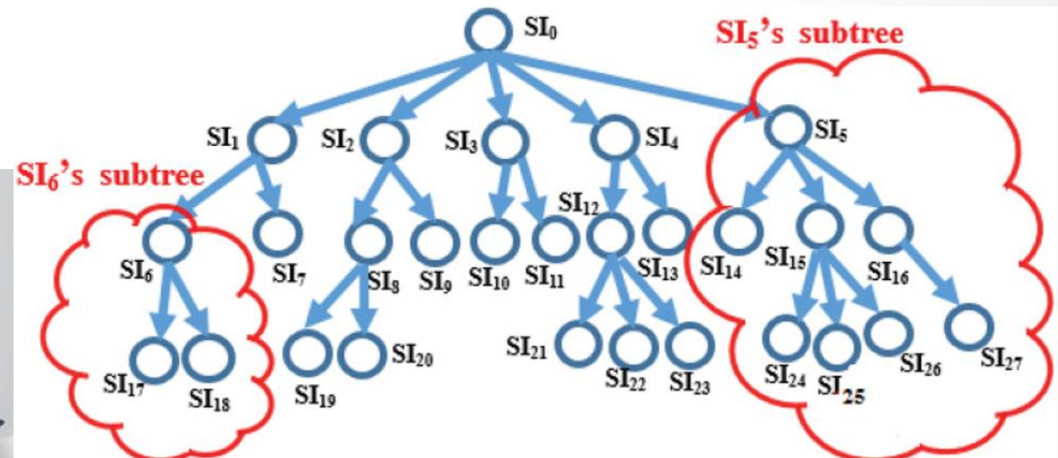
- Hardware ageing: computer systems age just like humans
  - CPU **transistor ageing**: fundamental speed ↓
  - Thermal performance decreased: lacking maintenance





Issues | Motivation

- Emerging systems
  - Self-adaptive systems: increased software complexity
  - Machine that learns and evolves, e.g., autonomous robots



Issues (continue) | Motivation

- Contribute **negative** and **non-deterministic** effects on WCETs.
- Subtle in a short period, but noticeable in long-term.
- Traditional WCET analysis could solve this by giving a very **pessimistic** boundary.

- A new perspective on WCET:
  <u>a dynamic view</u> of WCET (dWCET), as an extension of traditional WCET analysis.

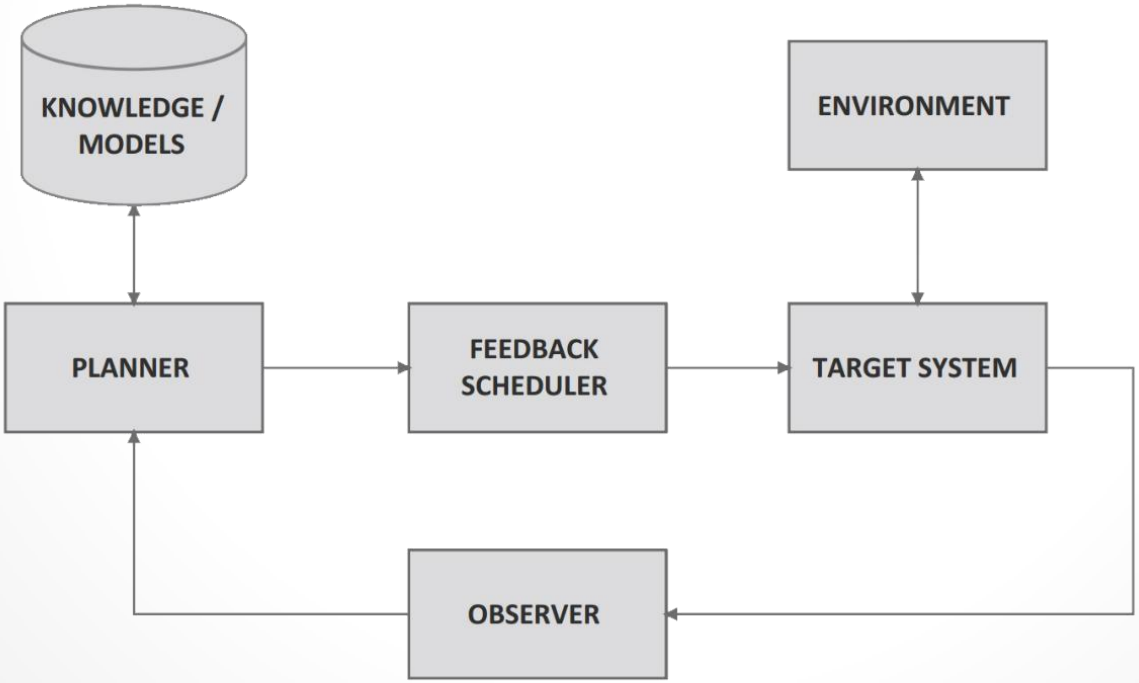
dWCET | Motivation

- Run-time modelling of WCET.
- Enhanced Parametric WCET:
  WCET = f(t, system changes[, mode, state, input, …]).

- Pro 1: Early detection of potential timing errors, and achieve graceful degradation.
- Pro 2: Utilize resources better (with feedback scheduling) .

Adaptive Feedback Scheduling

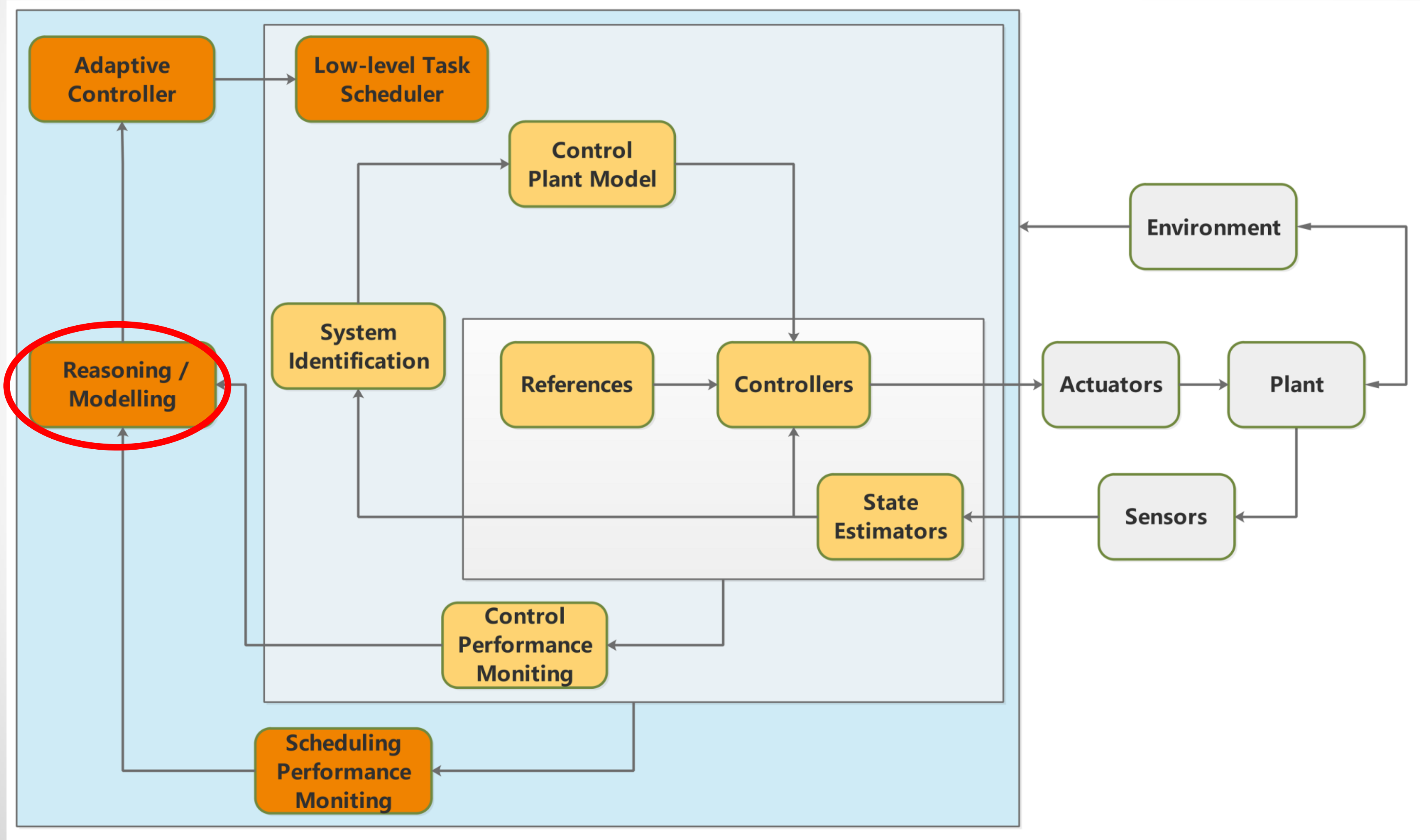
- A variation of Feedback Control Scheduling (FBS)
- Adaptive
  - ability to handle unexpected events
  - understanding of the system increases
-



[1] Lu, Chenyang, et al. "Feedback control real-time scheduling: Framework, modeling, and algorithms." (2002).

In Practice | A-FBS

- A-FBS uses with an adaptive control system:



Advantages | A-FBS

- Explicitly monitoring and modelling the system.
- Handling uncertainties in run-time executions.
- Increase system resilience: automated the process of (proactive) fault tolerance.
- Dynamic resource allocation: run-time optimization of scheduling.

What's Next?

- The activation of system changes/degrades will be propagated in the system and reflects on WCETs.
- There are many ways we can model dynamics in WCETs.
- In this initial study, we consider one of these: trends in WCET.
- Use a linear model to describe trend.

Trend Identification

- Many techniques in the literature:
  - AR / ARMAX
  - Regression Analysis
  - Non-parametric
  - EVT
  - Neural Network
  - Decision Tree Regression
  - …
- but not all of them fit our case:
  - data points are execution times
  - distribution is not known
  - few prior knowledge
  - need a long-term prediction

Methods | Trend Identification

- Non-parametric Methods
  - TSE: Theil-Sen Estimator
- Regression Analysis
  - OLS: Ordinary least-squares regression (OLS-regression)
- Extreme-value Theory
  - EVD: Generalized Extreme-value distribution
- Machine Learning Methods
  - SVR: Support Vector Regression

- These methods have never been used to analysis trends in WCETs. How to evaluate?

[1] Sen, P.K., “Estimates of the regression coefficient based on Kendall’s tau” (1968).
[2] Basak, Debasish et al. “Support Vector Regression.” (2008).
[3] Kotz, S.. “Extreme Value Distributions: Theory and Applications.” (2016).

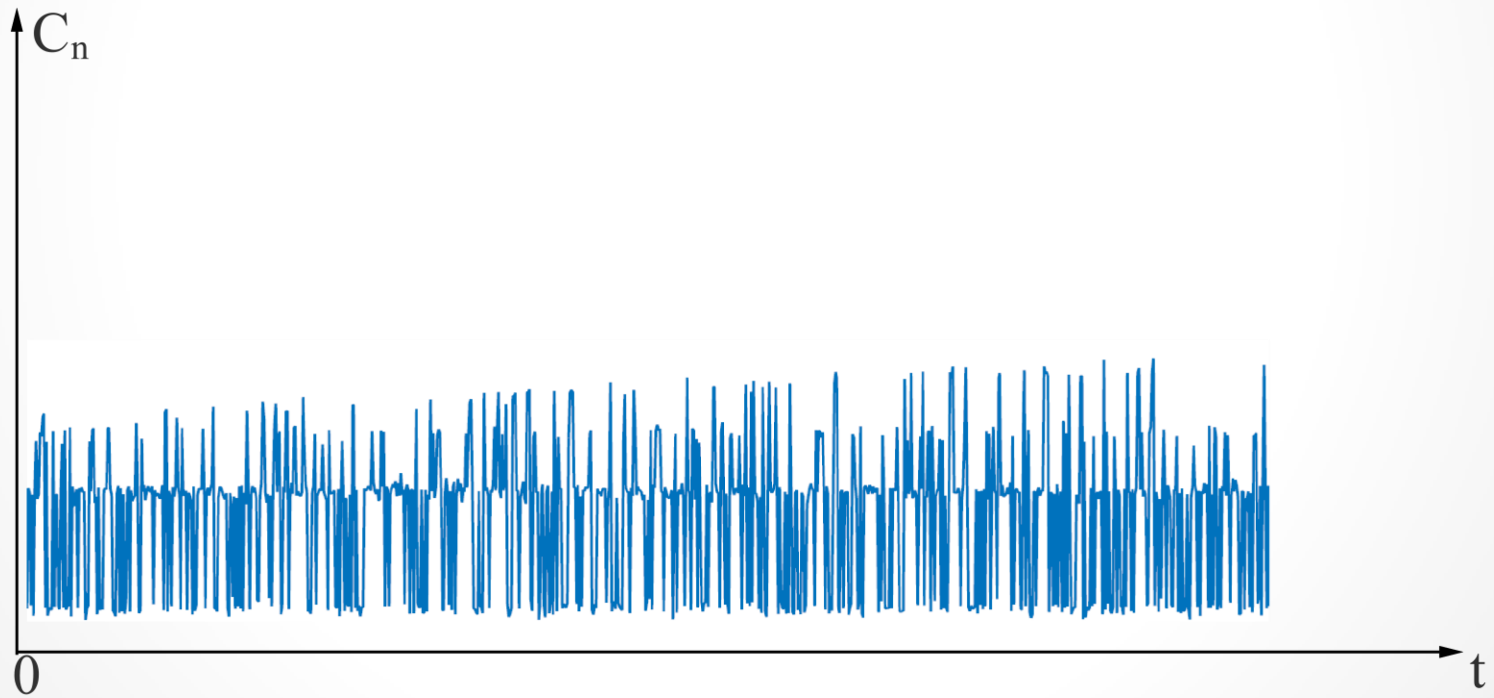
Dataset | Evaluation

- Use synthetic data to make it evaluable.
- One observation represents a high watermark of run-time executions.
- Markov model with multiple dominated paths.
- An increasing trend only in the worst-case path.

Group	Subgroup	Dataset Index	Data Size	Increasing Trend
A	A1	1 - 10	5,000	0%
B	B1	11 - 20	5,000	1%
	B2	21 - 30	2,500	2%
	B3	31 - 40	1,667	3%
	B4	41 - 50	1,250	4%

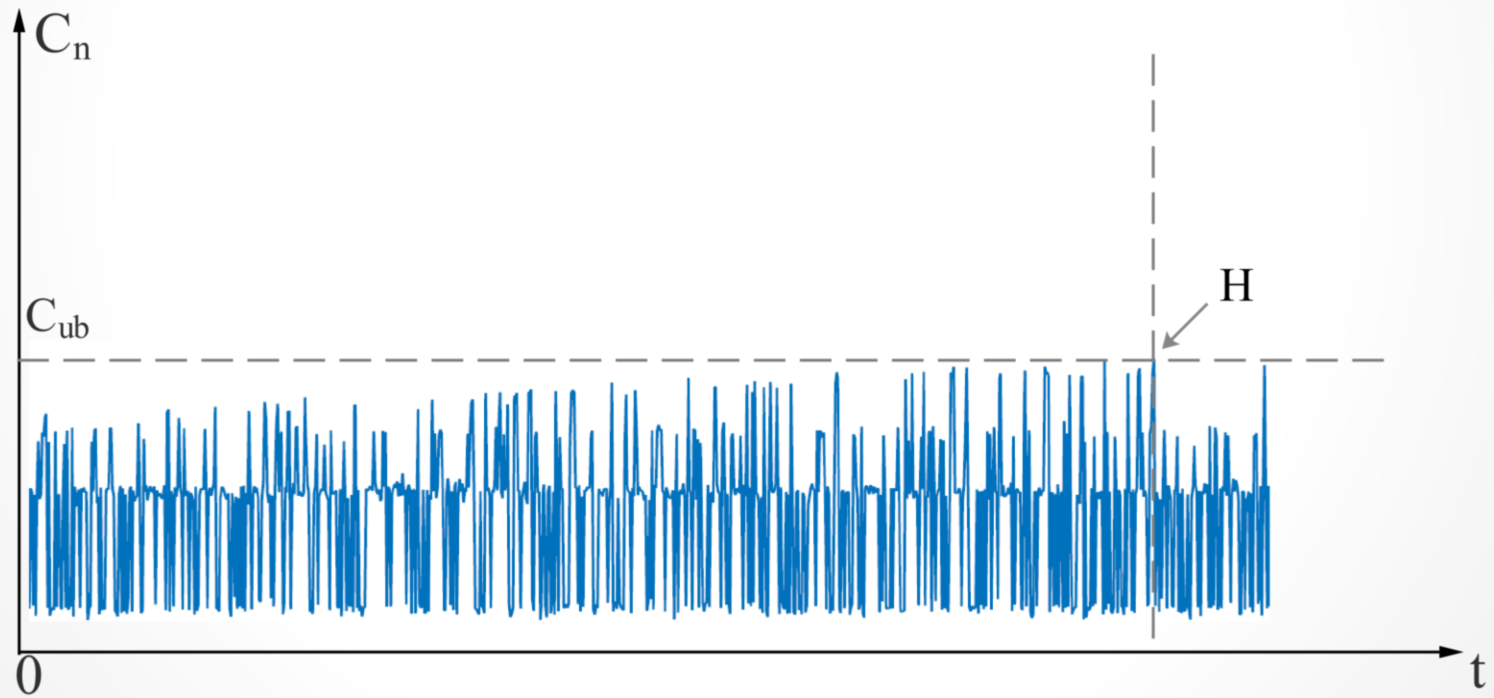
Evaluation

- The Evaluation Framework



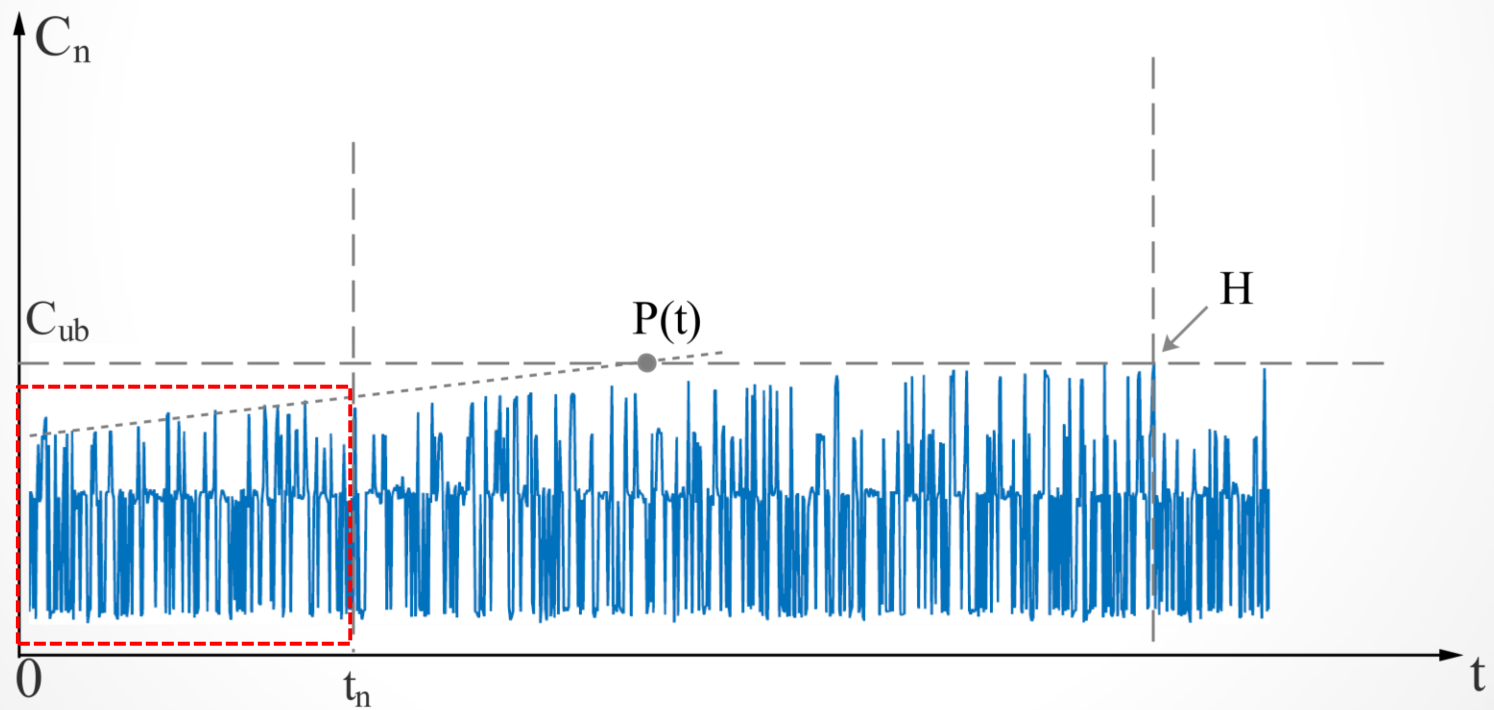
Evaluation

- The Evaluation Framework



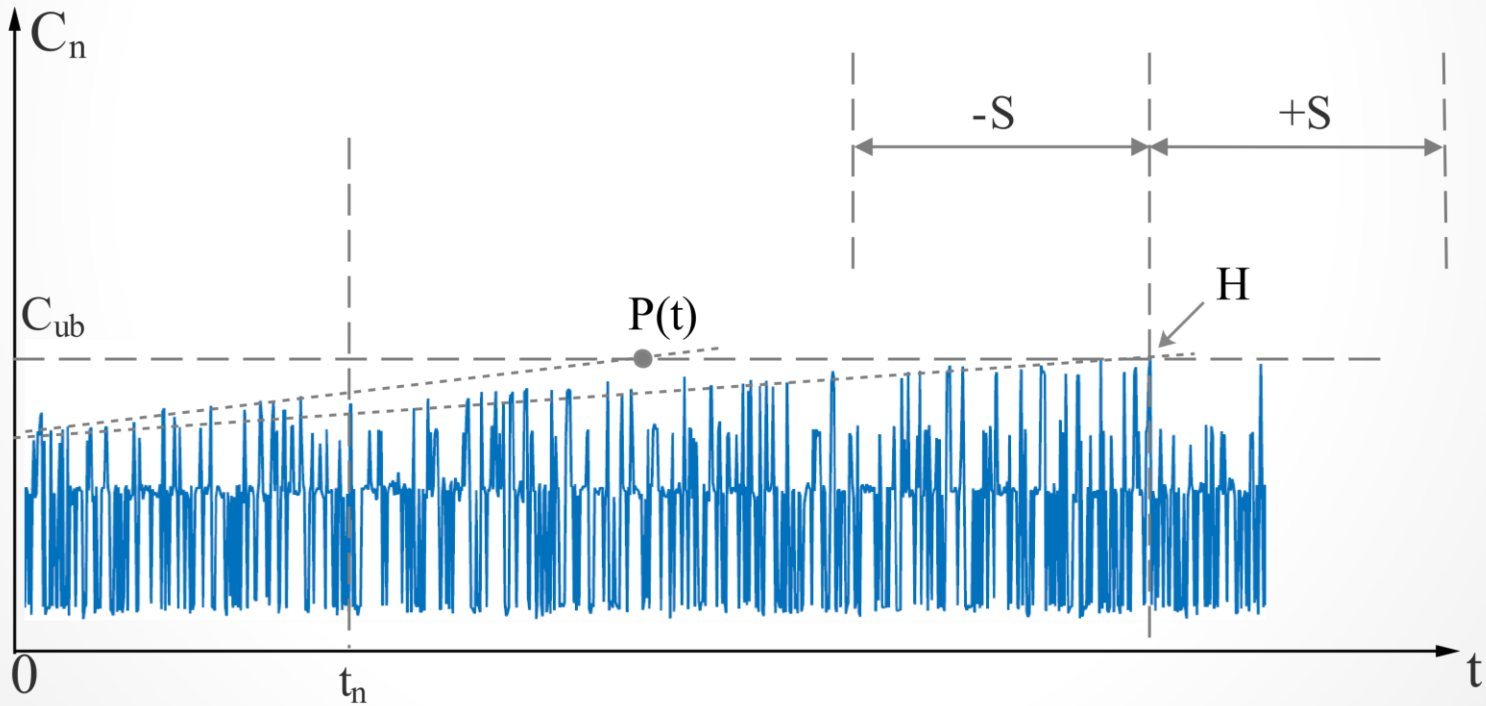
Evaluation

- The Evaluation Framework



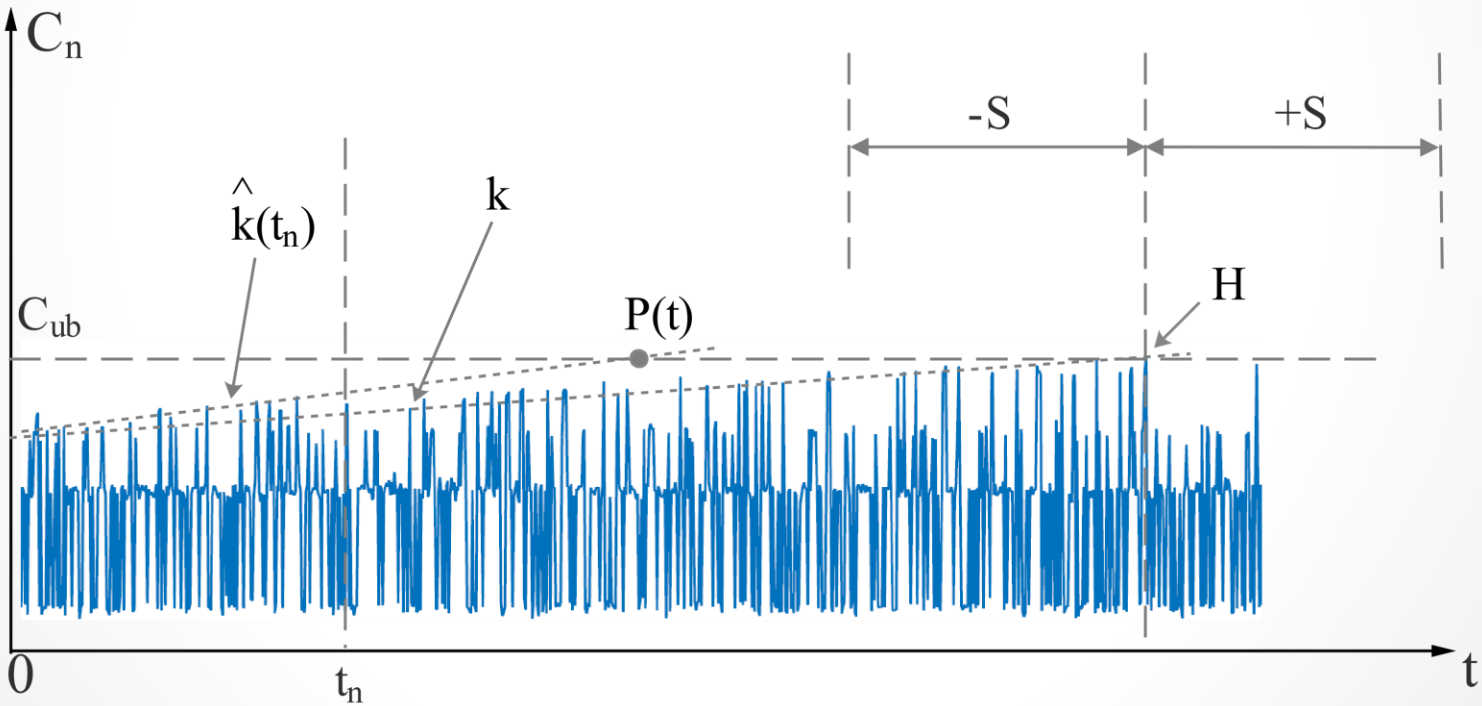
Evaluation

- The Evaluation Framework



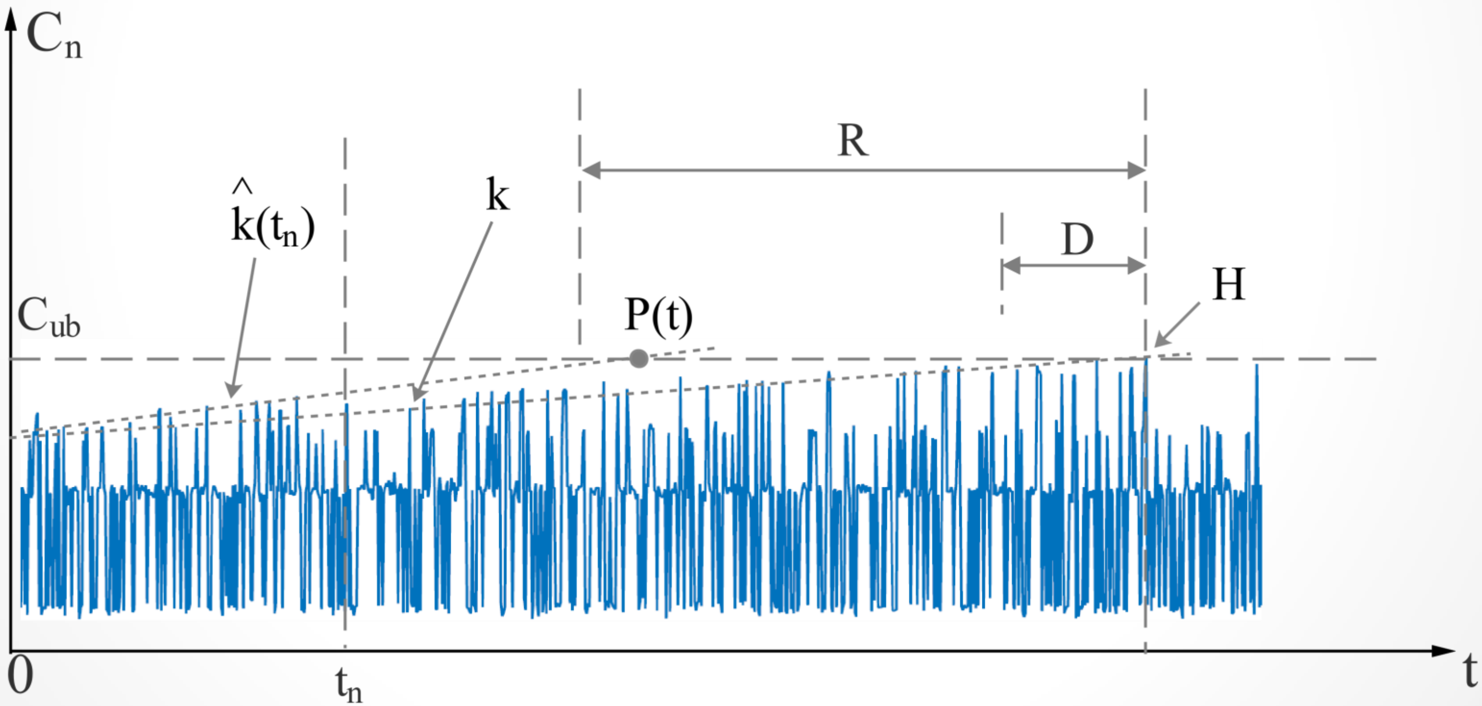
Evaluation

- The Evaluation Framework



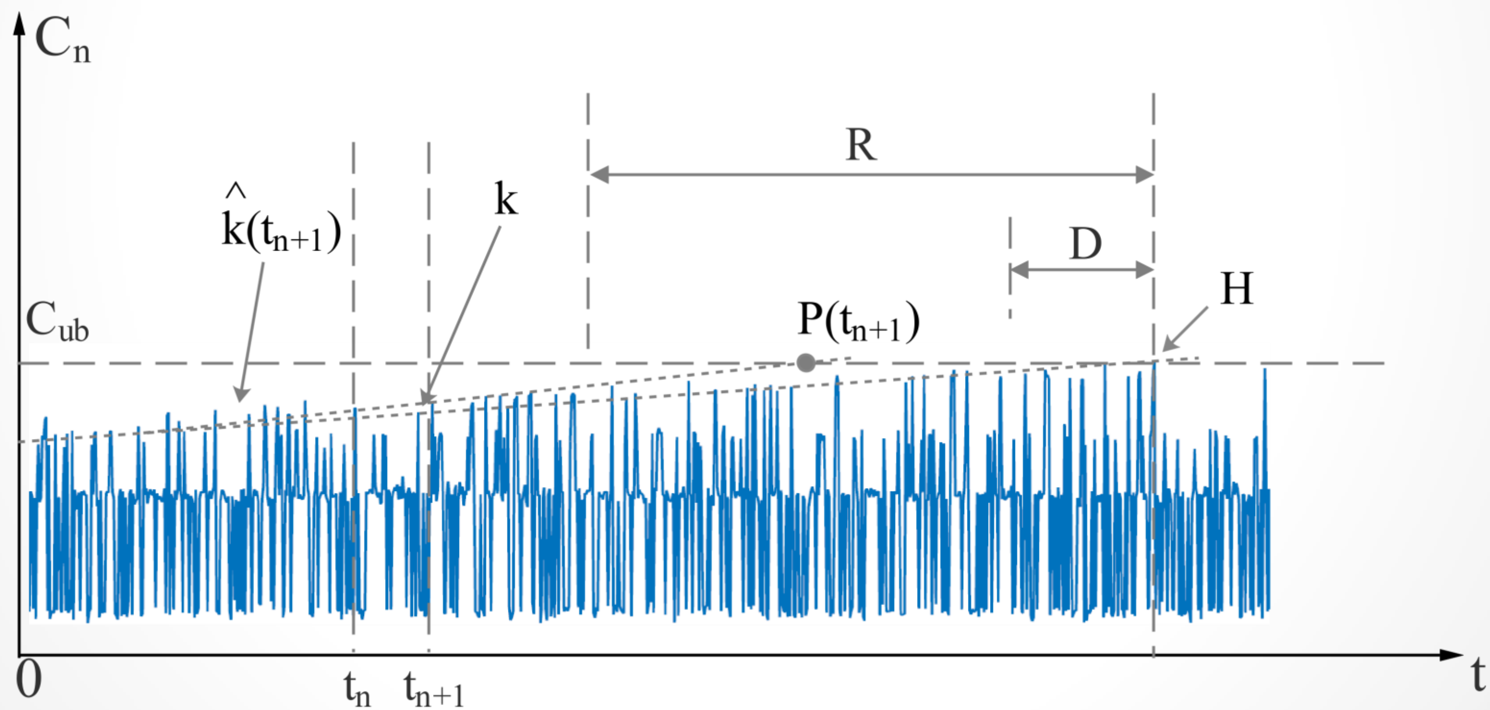
Evaluation

- The Evaluation Framework



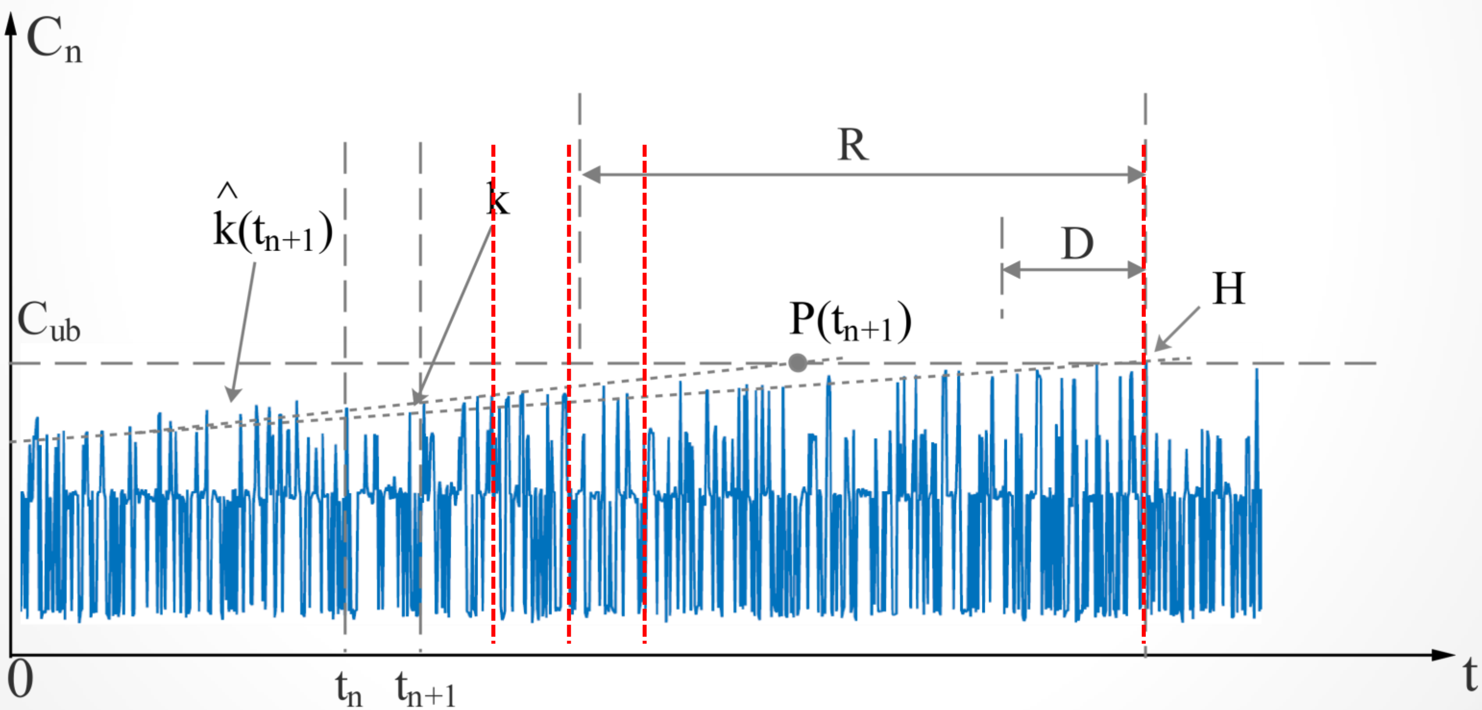
Evaluation

- The Evaluation Framework



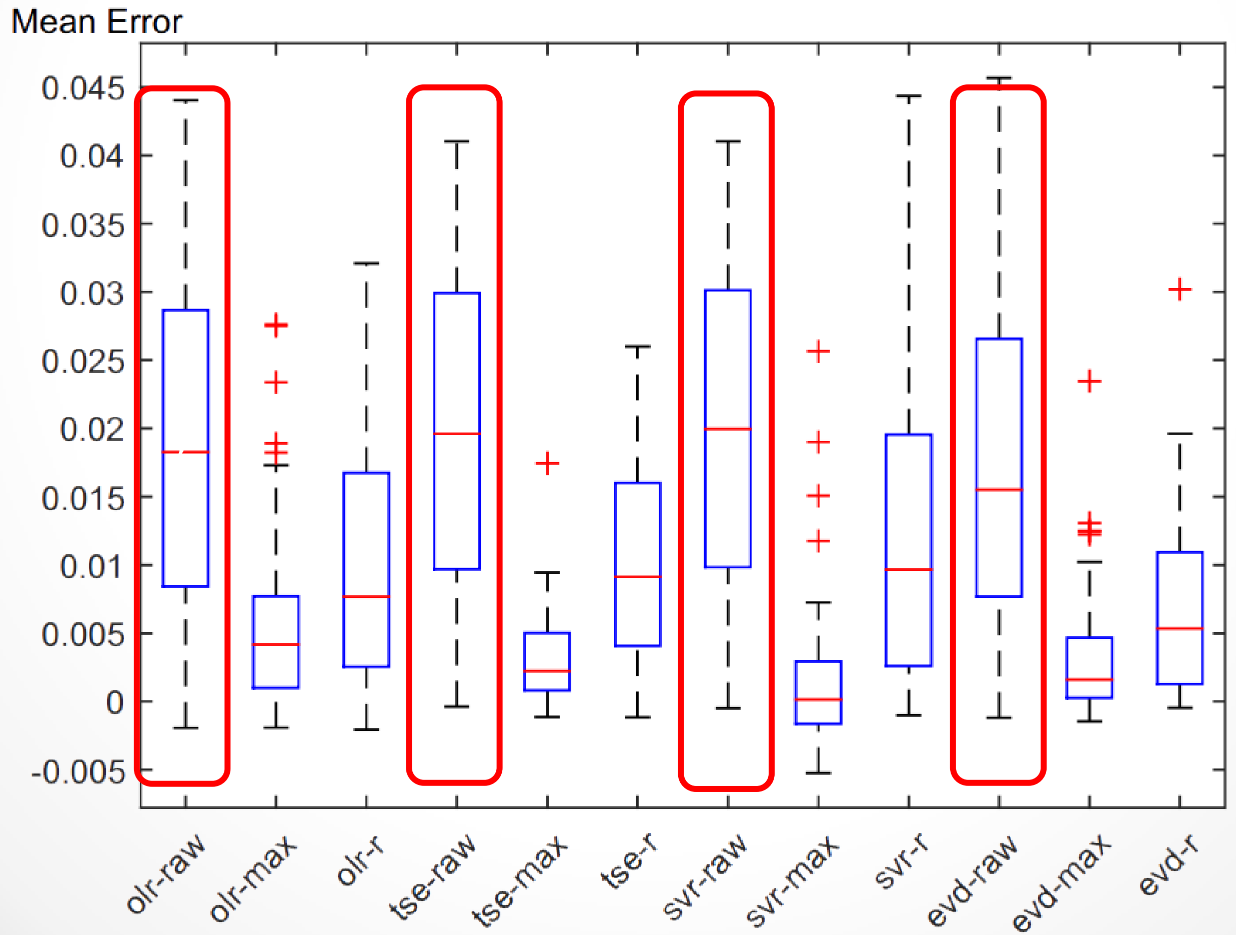
Evaluation

- The Evaluation Framework



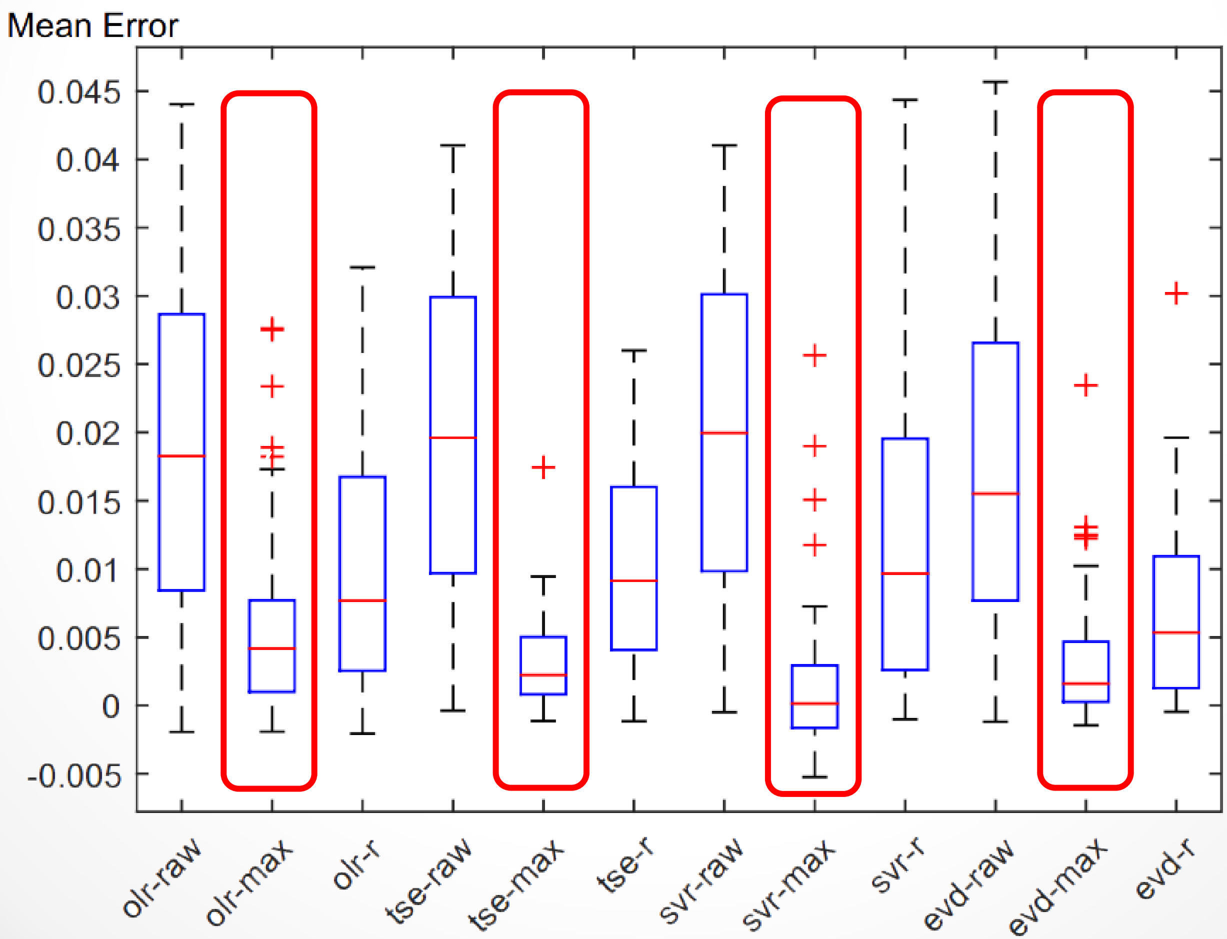
Pre-processing | Evaluation

- Evaluated with **raw**, block **max**ima and **r**-largest
- Mean (absolute) error of trend magnitude



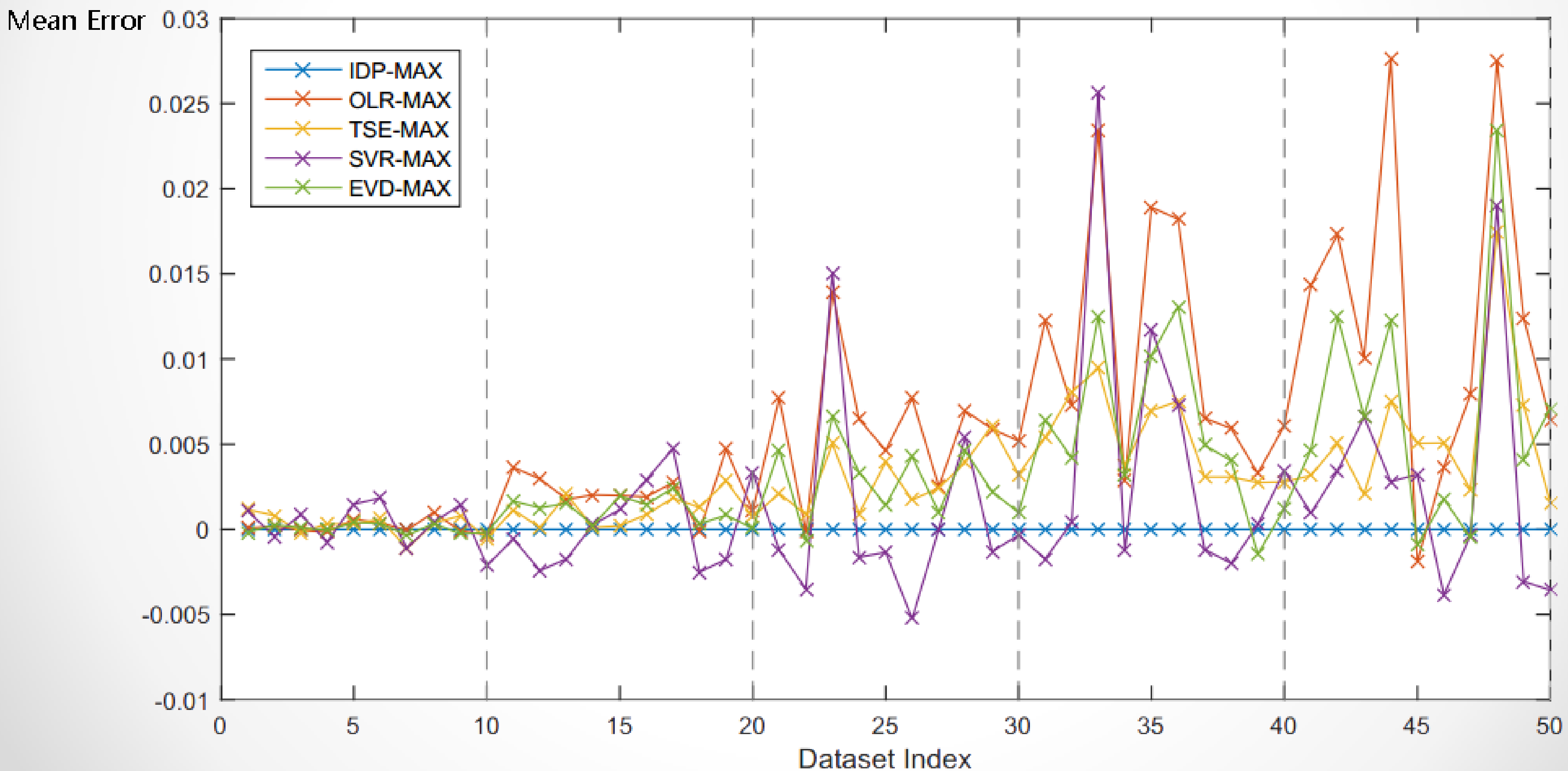
Pre-processing | Evaluation

- Evaluated with **raw**, block **max**ima and **r**-largest
- Mean (absolute) error of trend magnitude



Dataset Sensitivity | Evaluation

- All methods use block maxima
- Subgroups are separated by dashed lines

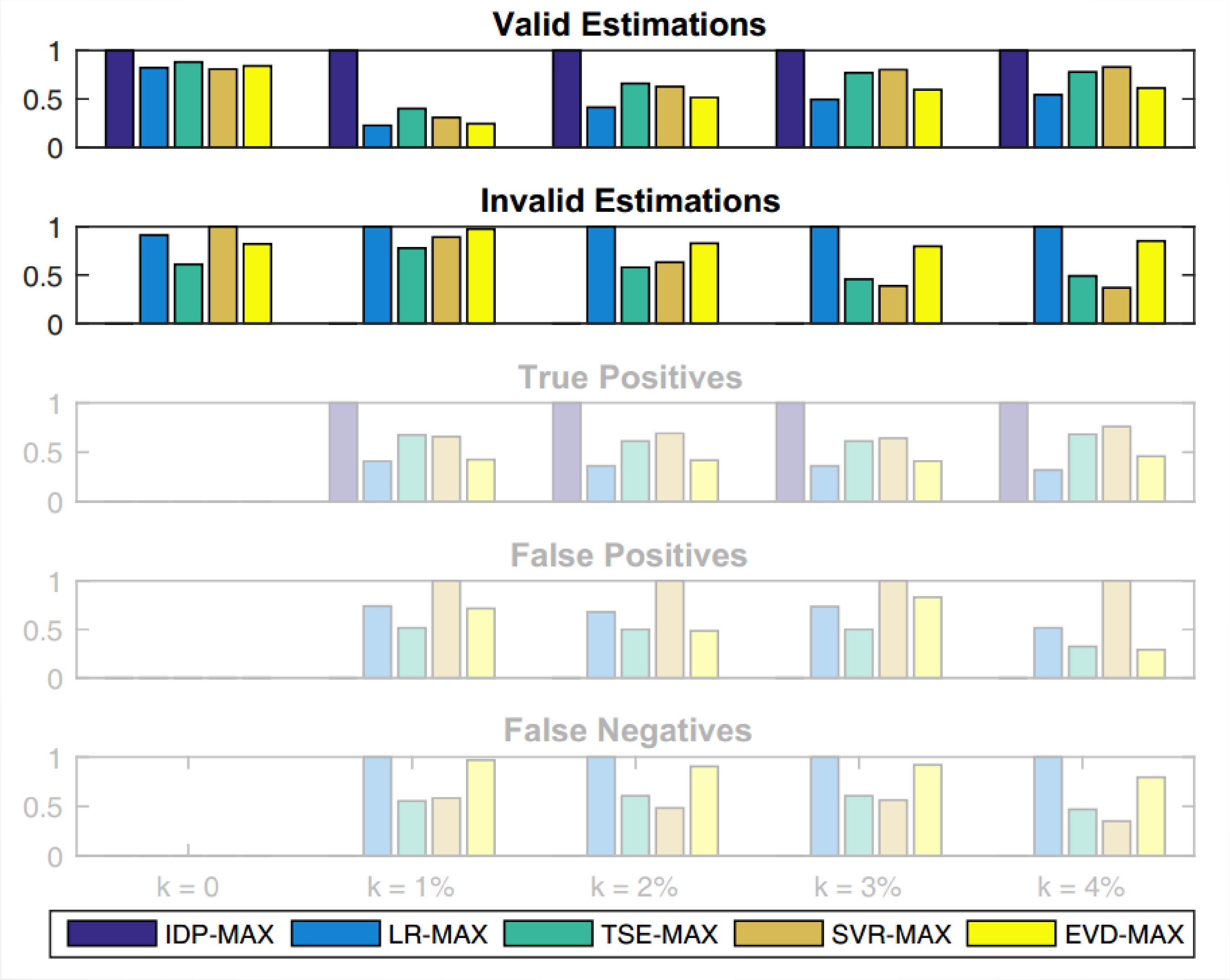


Trend Error | Evaluation

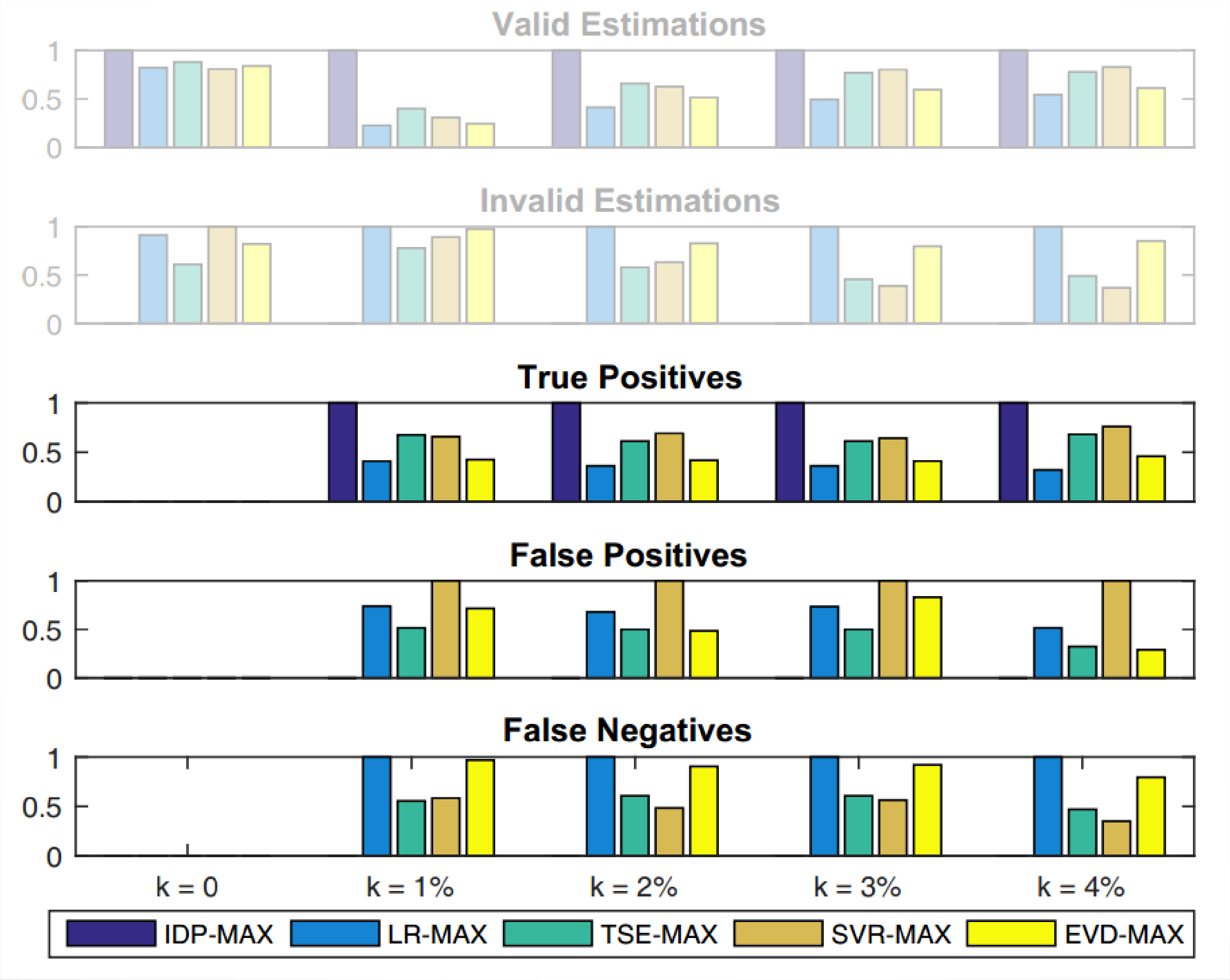
- Evaluate Trend error (= actual k - predicted k):

	Minimum	Median	Mean	Maximum	σ
olr-max	-1.91	4.16	6.31	27.64	7.21
olr-r	-2.07	7.68	10.59	32.10	9.31
tse-max	-1.12	2.23	3.07	17.45	$\boxed{3.27}$
tse-r	-1.15	9.14	9.91	26.00	7.72
svr-max	-5.24	$\boxed{0.15}$	$\boxed{1.60}$	25.65	5.71
svr-r	-1.00	9.65	12.72	44.36	12.74
evd-max	-1.46	1.60	3.40	23.47	4.75
evd-r	-0.45	5.34	6.86	30.20	6.77

Normalized Performance | Evaluation



Normalized Performance | Evaluation



Mean Penalties | Evaluation

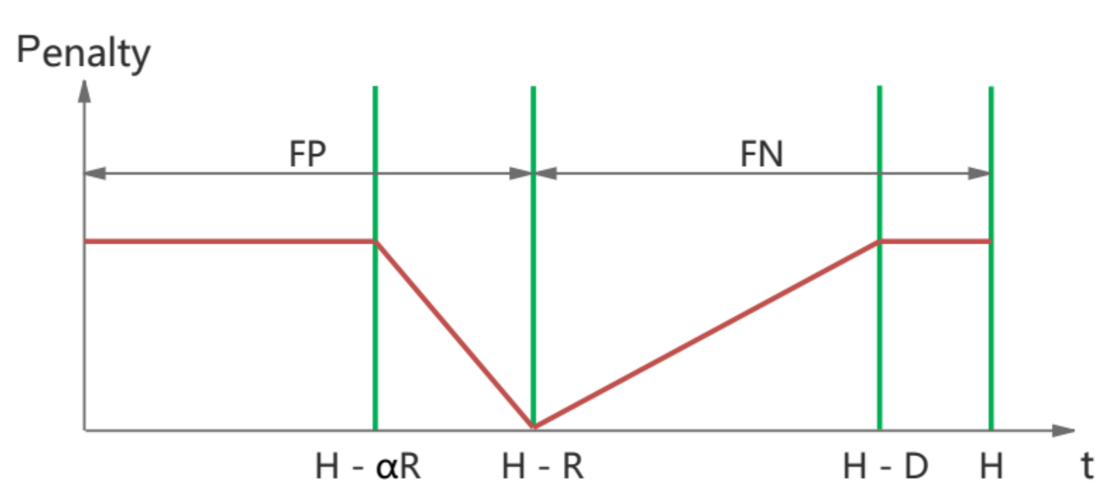


Table 4. Mean penalties over all datasets for each prediction method

	OLR	TSE	SVR	EVD
raw	62	62	62	62
maxima	58.28	29.02	42.26	49.68
r-largest	58.2	53.82	77.58	55.76

Conclusion

- Introduced dWCET and A-FBS
- Evaluated data pre-processing methods
- Result is sensitive to datasets
- Best two methods: svr-max and tse-max

- Future work
  - More dedicated dataset: e.g., with non-linear trend
  - Other analysis: anomaly detection, pattern recognition
  - Multiple variables + PCA
  - Evaluate with real-world data

Thank You for your attention!

Any Question/Comment?