prelert®

Elastic London User Group Meetup

Anomaly Detection Using Elasticsearch

Dr. Stephen Dodson, CTO, Founder, Prelert

May 19, 2016

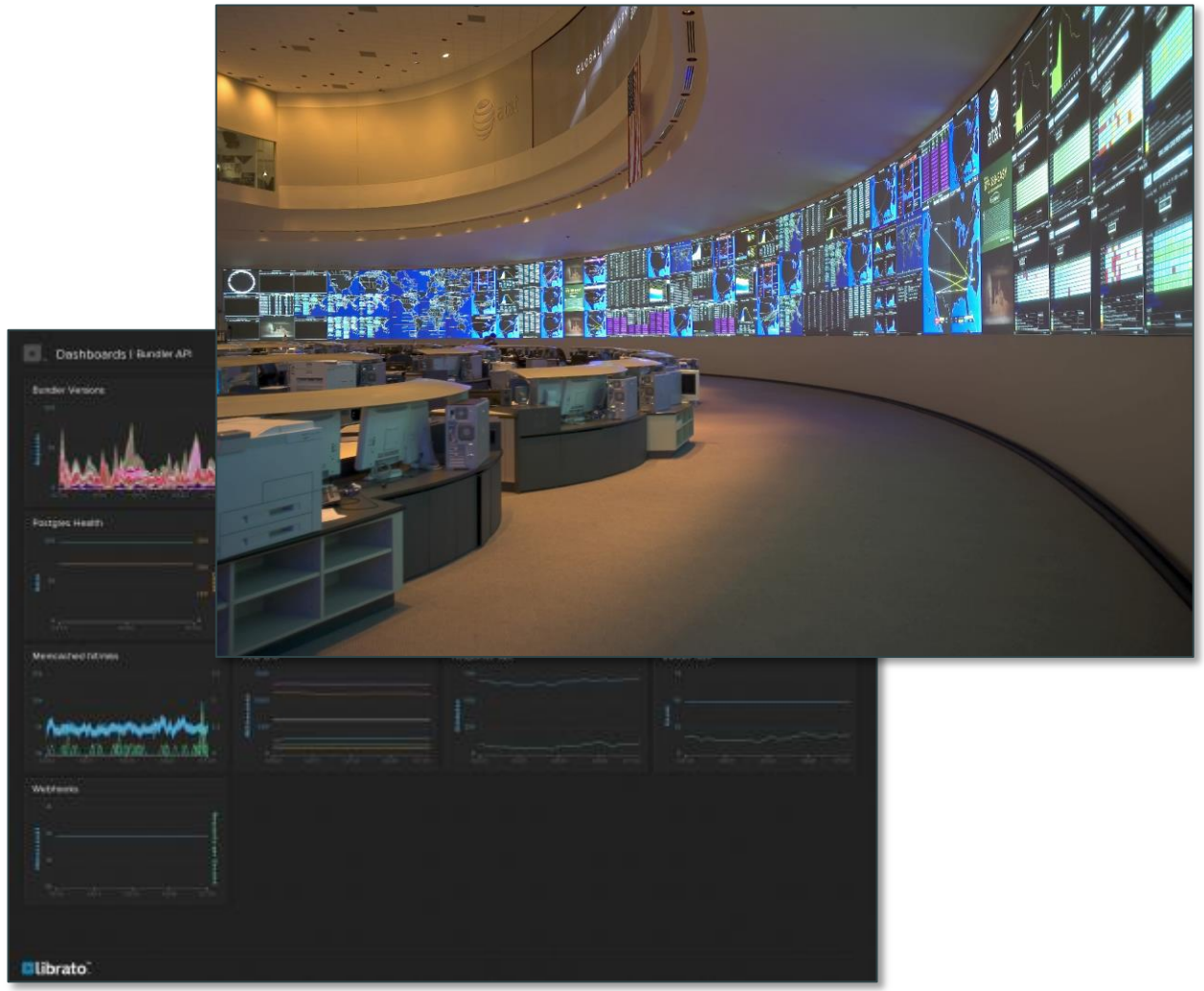
Overview

- **Prelert Overview**
  - *Elastic partner*
  - *Recently released "Behavioral Analytics for the Elastic Stack" Product*
    - *Application built using Kibana's UI framework*
  - 100+ customers + OEMs with CA, Bluecoat + others
    - IT Operations, IT Security, Retail analytics, IoT etc..
  - Note – the dev team in London did the work behind the product! I am only presenting!
- **Why is anomaly detection useful**
- **Deep dive into some anomaly detection approaches**
  - (but due to time constraints this will probably only describe univariate metric modeling in detail)
- **DEMO**

Problem Overview

Problem Overview – IT Operations

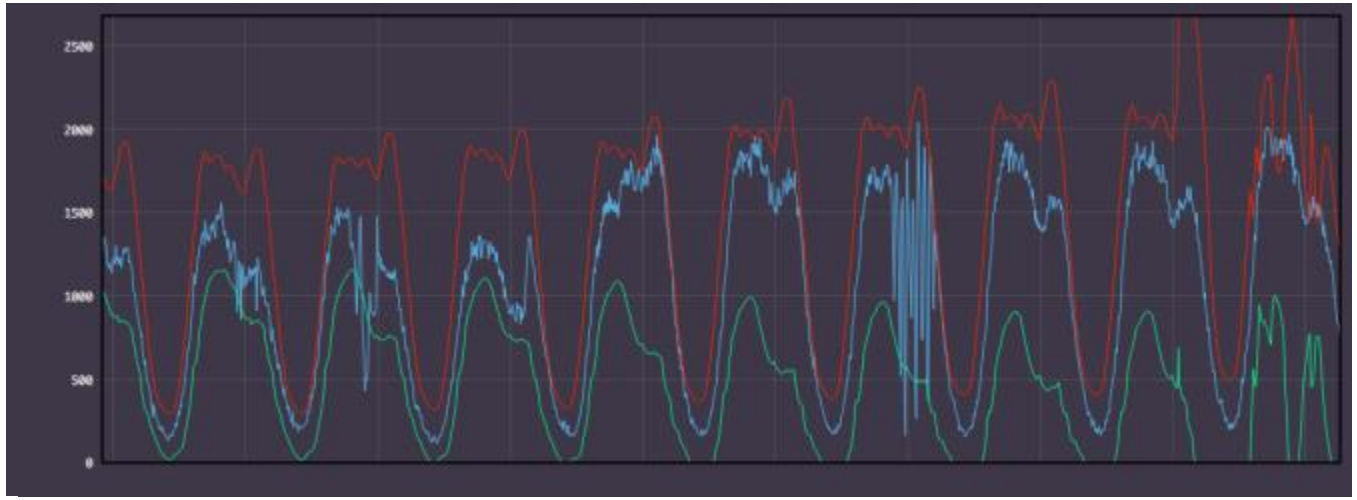
- **Operation Center Example**
- **Questions:**
  - Are my systems behaving normally?
  - If not, why?
- **Requires collection and storage of machine data (e.g. Elasticsearch)**
- **Requires understanding normal system behavior, *and when system is abnormal (anomalous)***

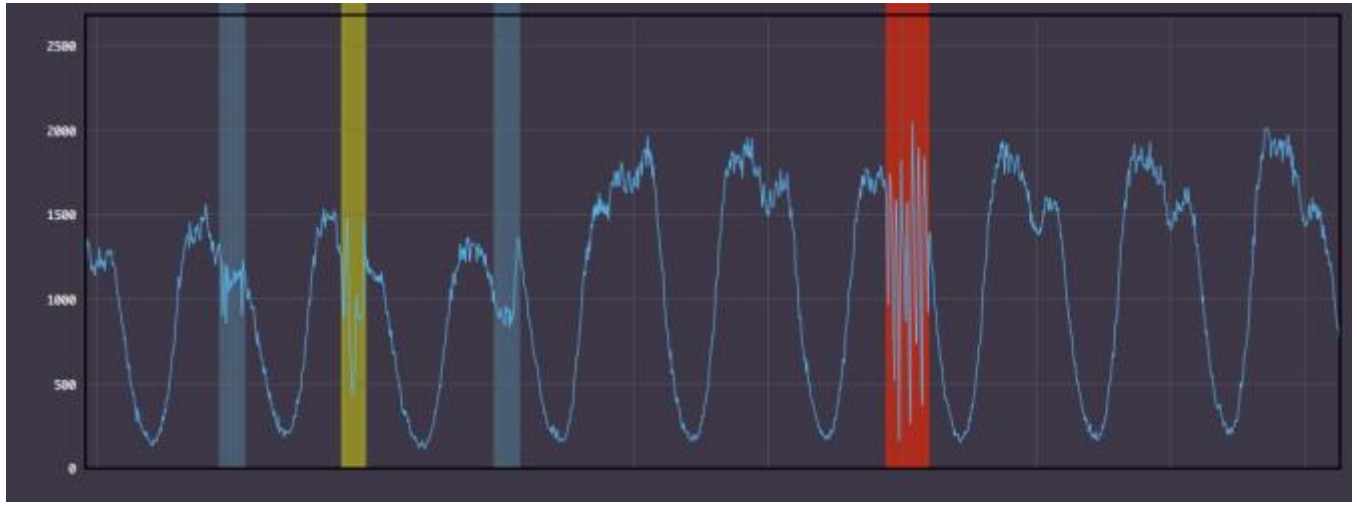


prelert®

Retail Customer Example

Alert when KPI (orders/min) changes

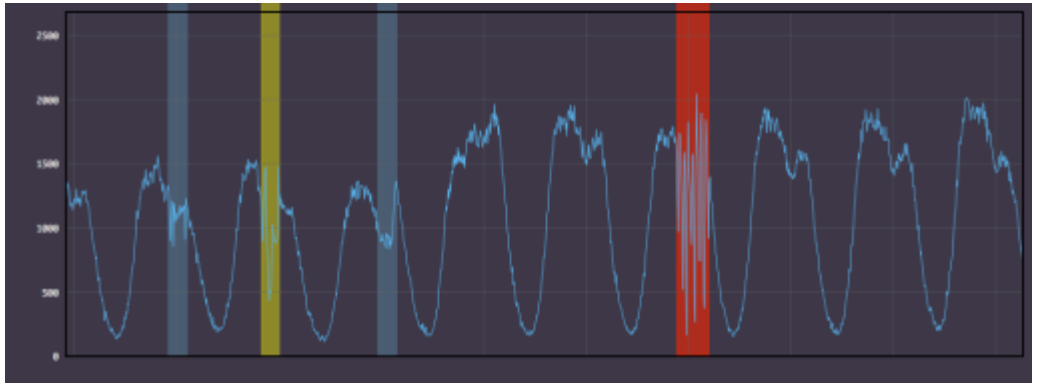




prelert®

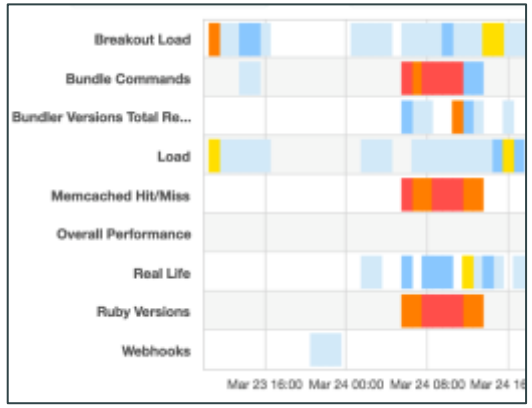
Retail Customer Example – KPI to Probable Causes

- **Identify anomalies in:**
  - Key Performance Indicators (KPI's) e.g. Orders per minute, Response time etc.
  - Backend metrics and log files e.g. products, points of sale, applications, systems



Automatically correlate key performance indicator anomaly to probable causes

Backend metrics/logs anomalies



prelert®

Anomaly Detection Methods

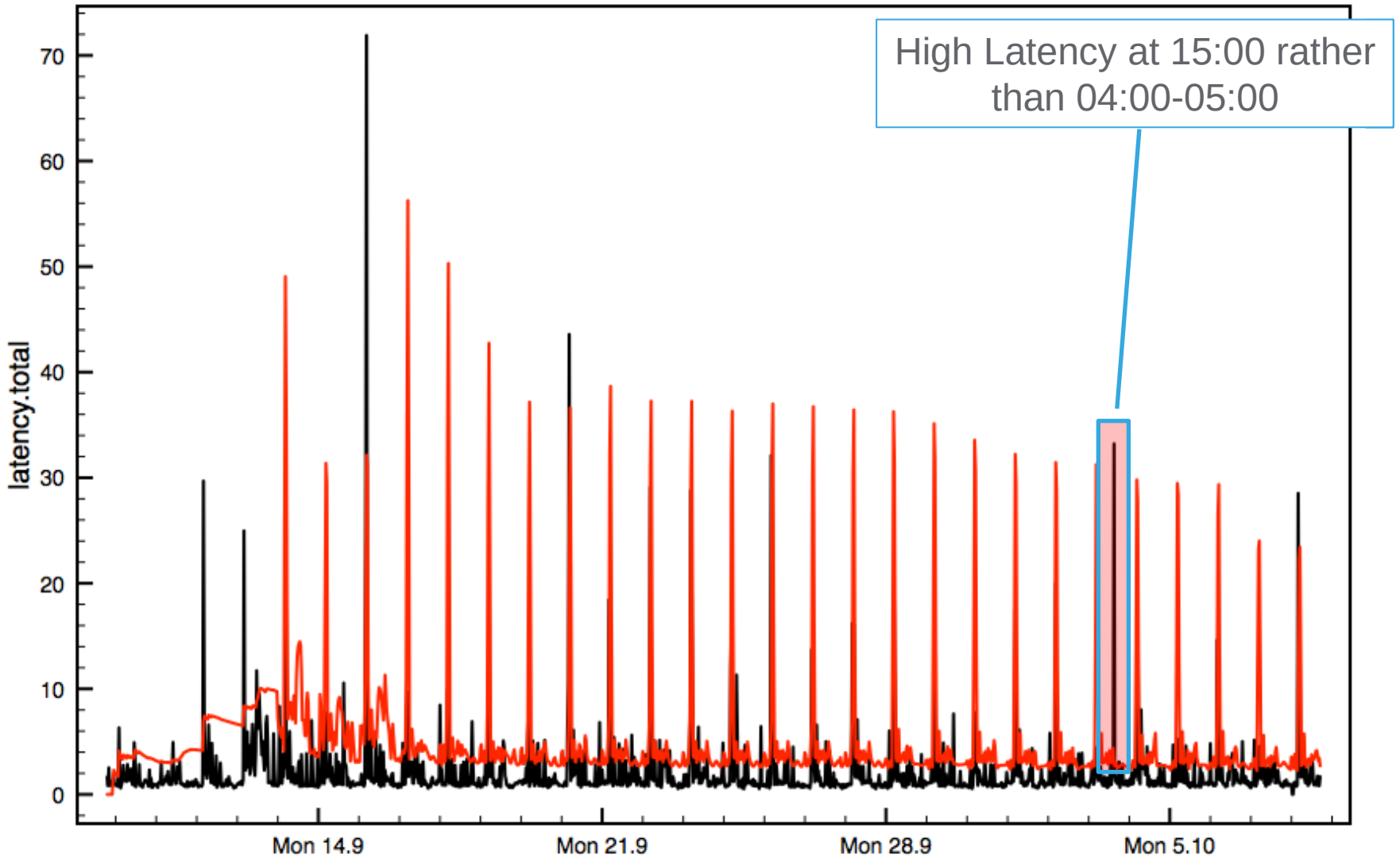
prelert®

Anomaly Detection Methods

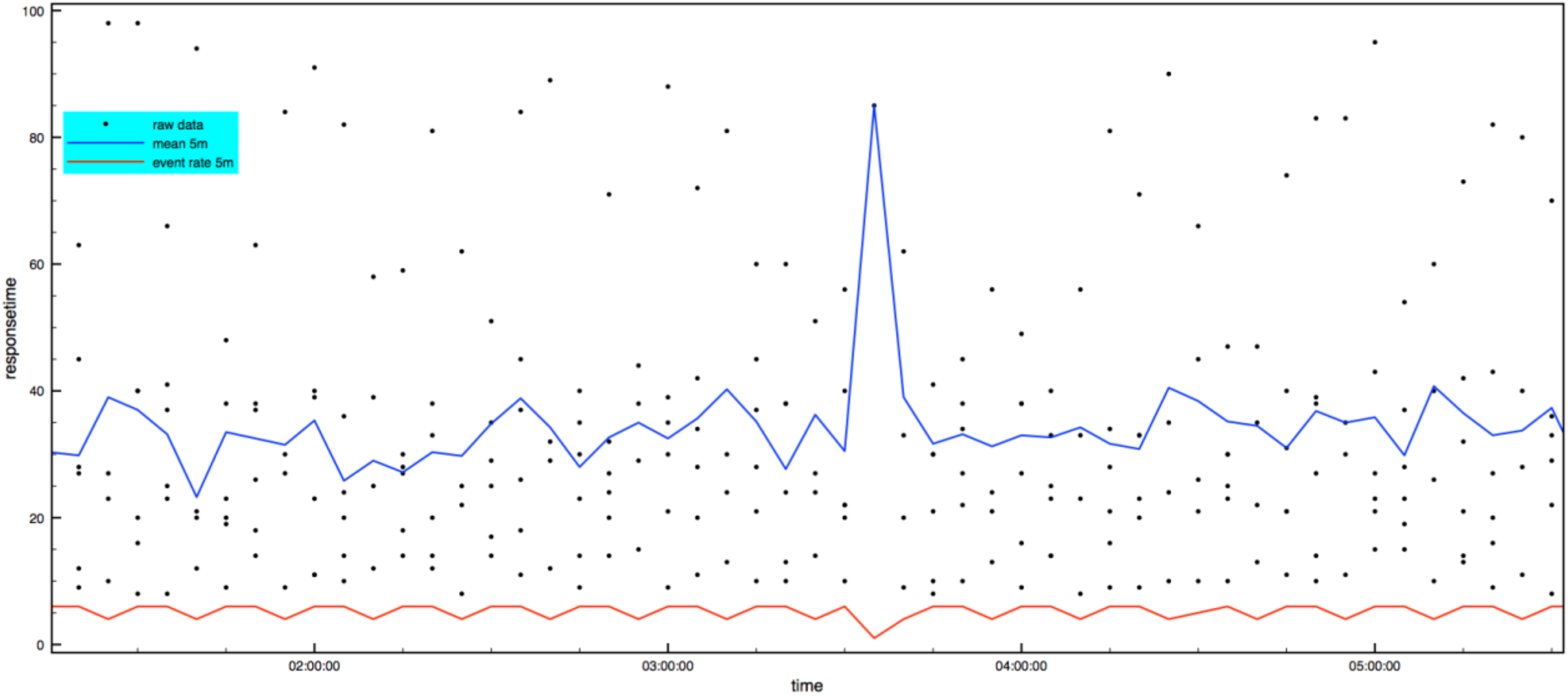
- **Goal: identify anomalies in time series data**
  - Eyeballing charts
  - Static thresholds
  - Dynamic thresholds
    - Behavioral analytics: learn behavior of data from the data

Eyeballing charts

- **Limited to a subset of available information**
  - Can’t look at all data
- **Can be misleading**
  - Different aggregation strategies can look very different
- **24x7 coverage is expensive**

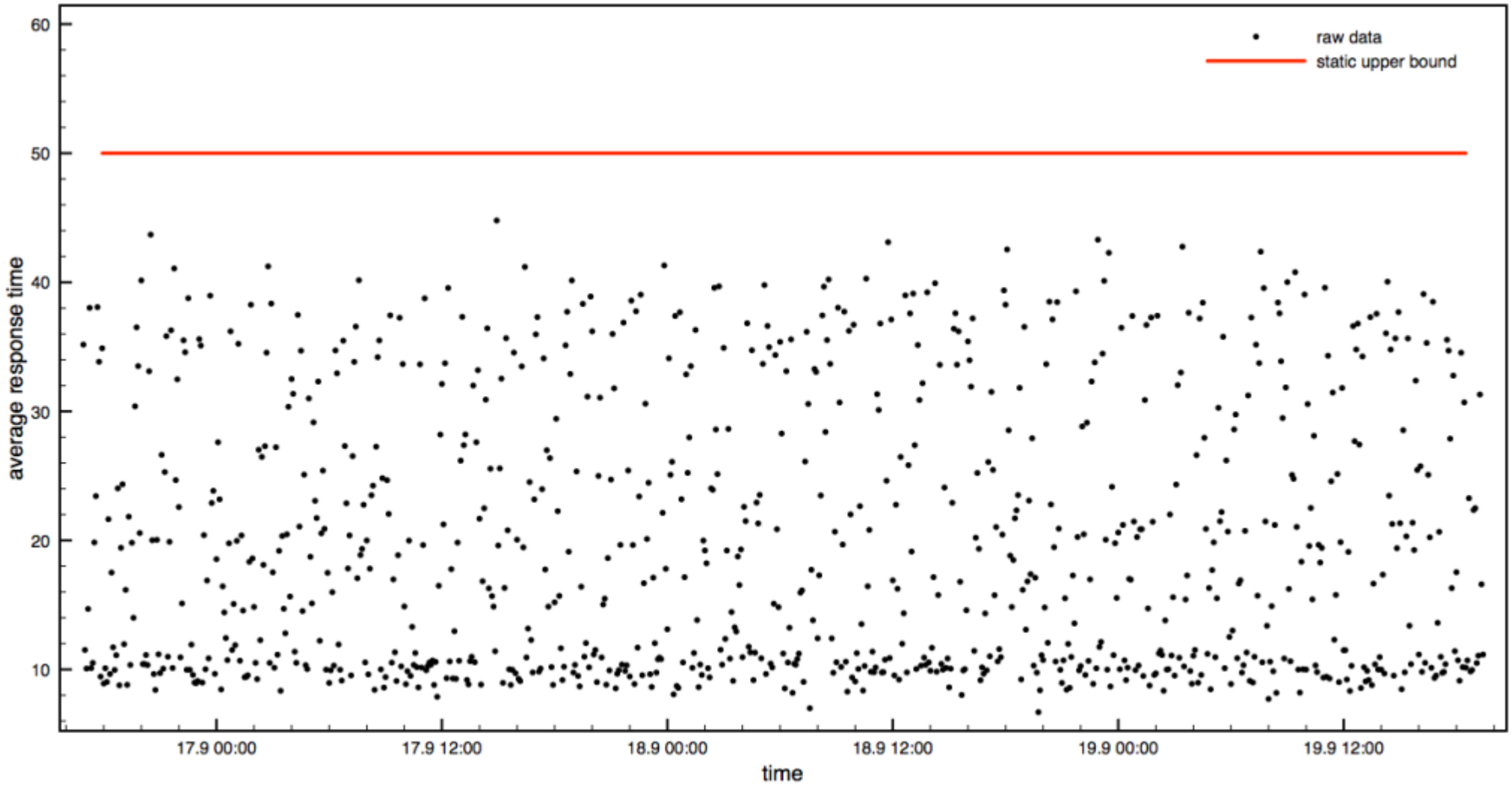


Problem Overview – IT Operations



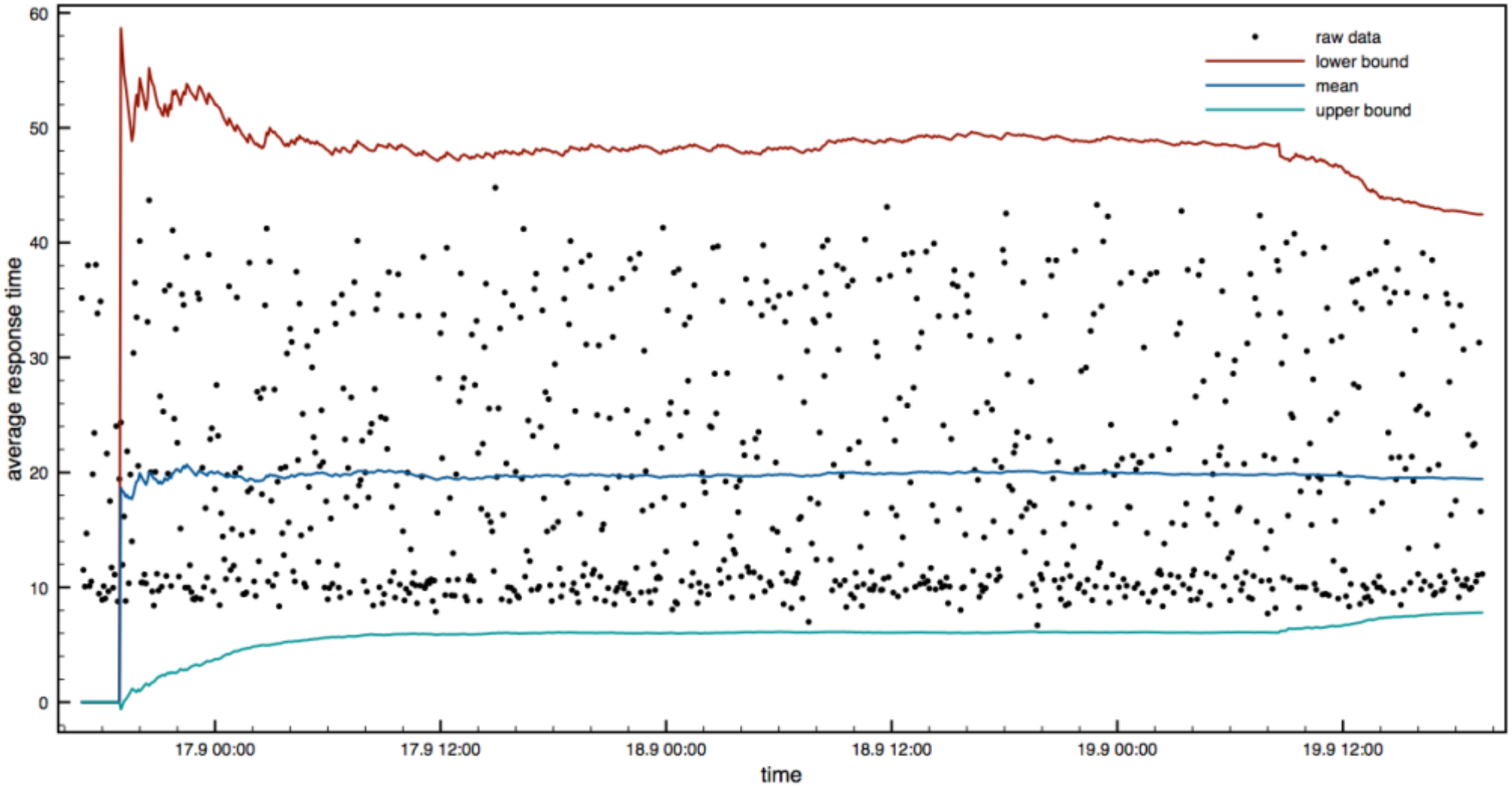
prelert®

Static Thresholds



prelert®

Dynamic Thresholds

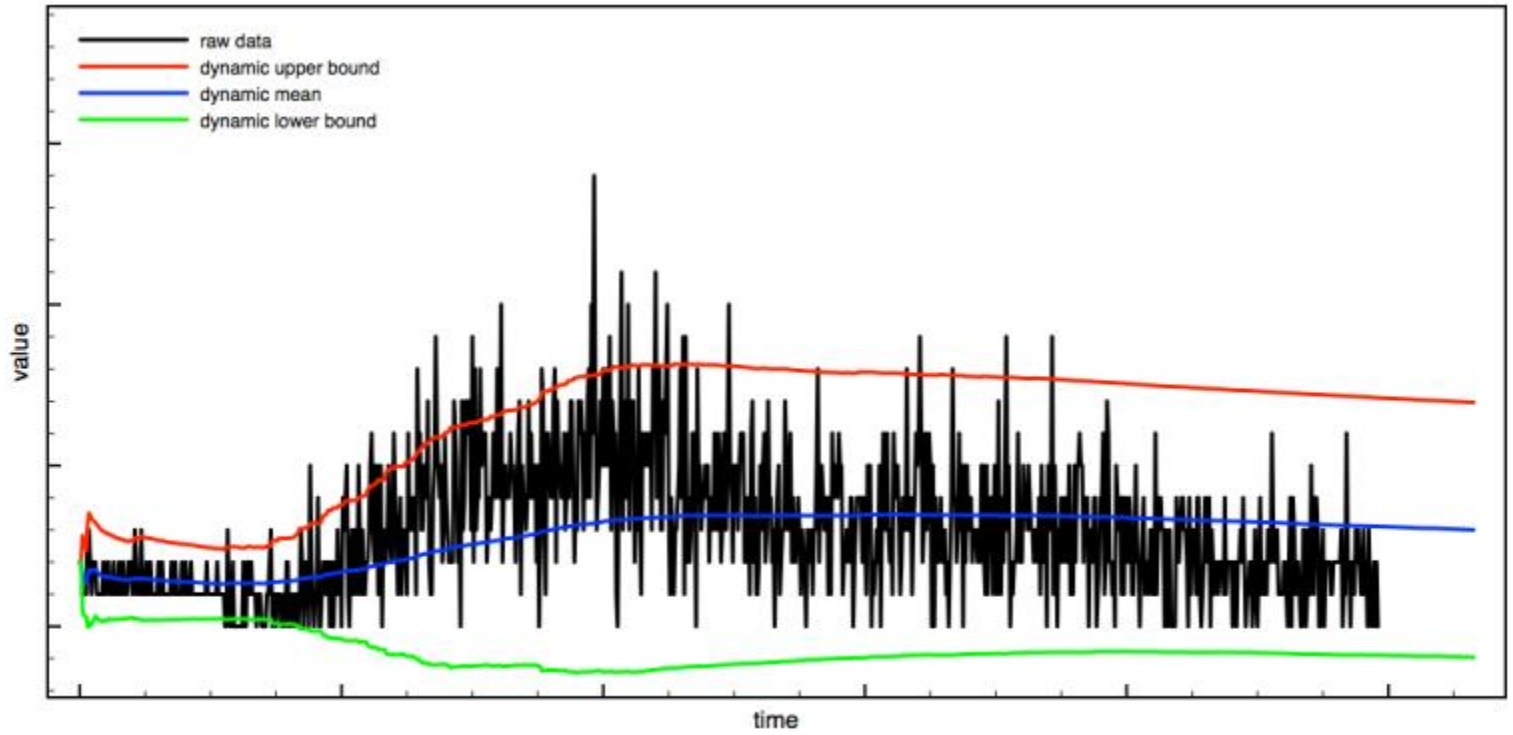


prelert®

Dynamic Thresholds

How to model data?

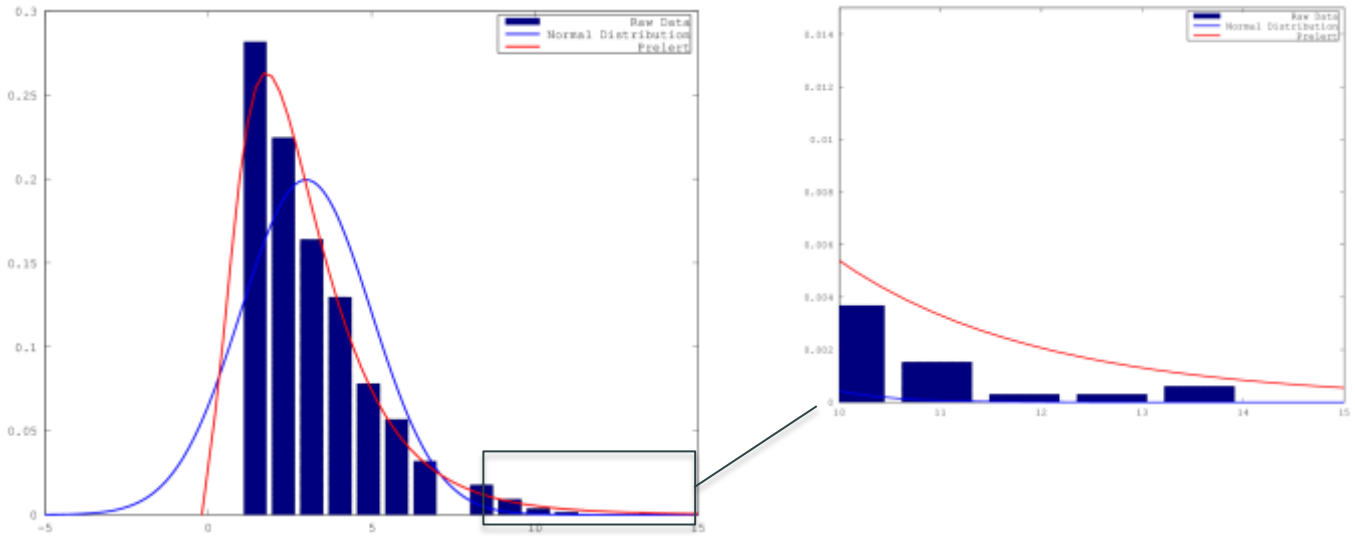
- **General Approach:**
  - Unsupervised machine learning
  - Create a predictive model for the distribution of feature values at given time as a function of time, based on the historical values we have seen to date
  - Use predictive model to calculate probability (anomalousness) of a new value
- **First approach – assume data fits a Normal Distribution (anomalies where value > mean + 2.5 sd)**

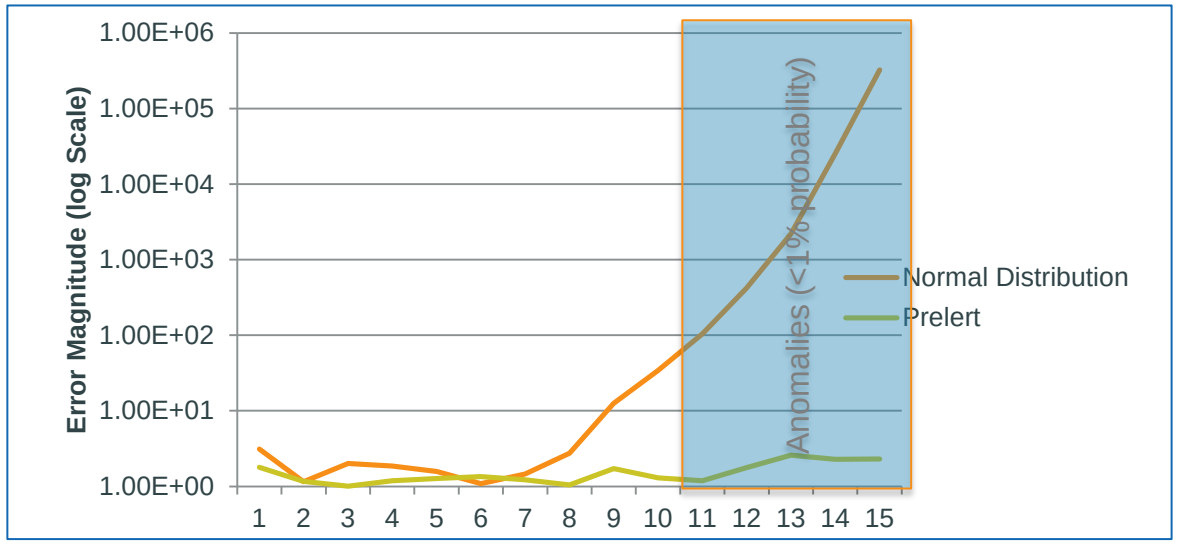


prelert®

Dynamic Thresholds

How to model data?





Value	Reference Probability*	Probability Using Normal Distribution	Probability Using Prelert
1	1	0.319498	0.560764
2	0.71835734	0.620394	0.834183
3	0.493717438	0.996086	0.49202
4	0.329757892	0.613486	0.27675
5	0.200122587	0.314752	0.157113
6	0.121973644	0.132196	0.0906333
7	0.065277352	0.0448887	0.0532552
8	0.033404842	0.0122142	0.0318895
9	0.033404842	0.00264629	0.0194524
10	0.015629789	0.000454401	0.012078
11	0.006435795	6.16E-05	0.00762581
12	0.002758198	6.58E-06	0.00489119
13	0.001225866	5.53E-07	0.00318382
14	0.000919399	3.65E-08	0.00210125
15	0.000612933	1.89E-09	0.00140482
16	?	7.64E-11	0.000950655
17	?	2.42E-12	0.000650663
18	?	6.01E-14	0.000450111
19	?	1.17E-15	0.000314514
20	?	1.77E-17	0.000221852

* Reference Probability is not ‘exact’ but empirical estimate

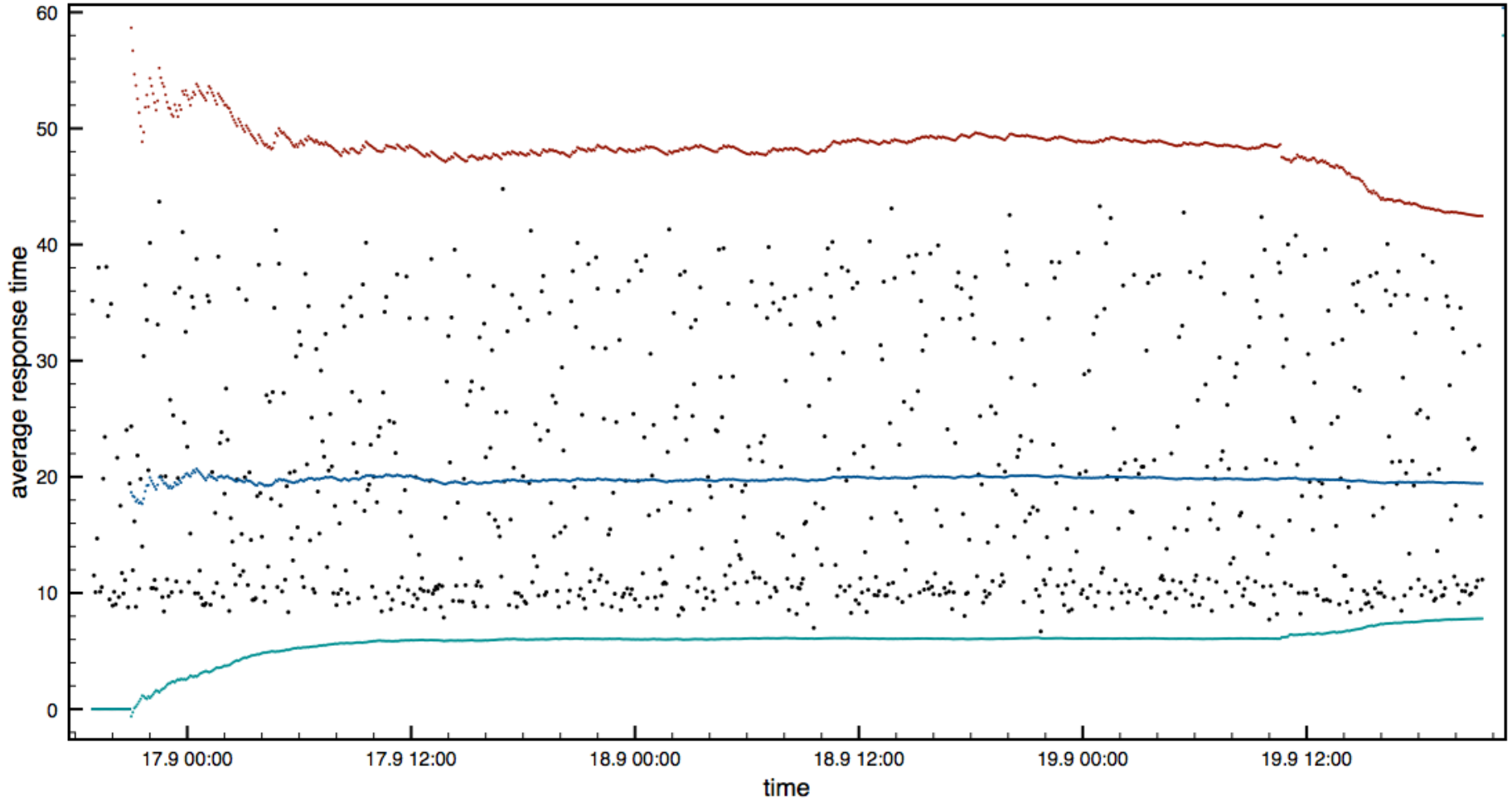
prelert®

Behavioral Analytics
Requirements and Constraints

- **Lots of different techniques for anomaly detection and time series analysis from mechanical Turk to deep learning**
- **How to choose an appropriate method?**

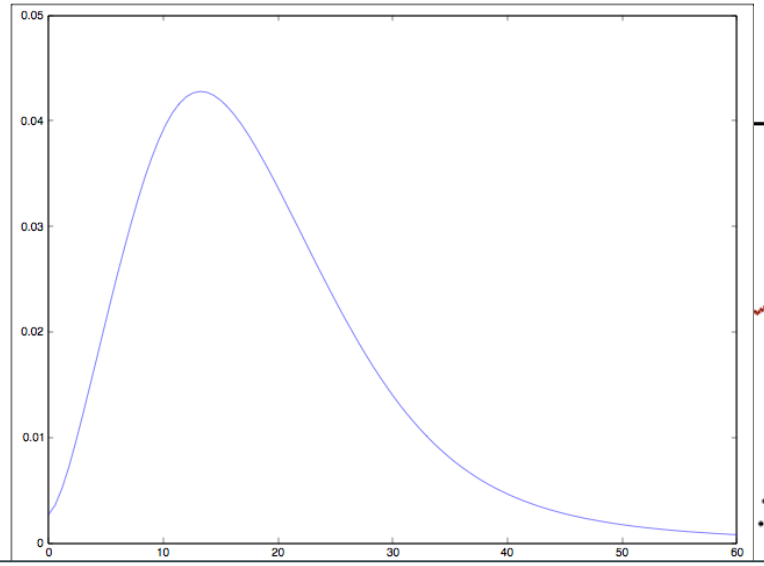
- **Requirements and constraints**
  1. *Real-time*
  2. *Accurate*
  3. *Robust*
  4. *Easy to use*
  5. *Scalable*

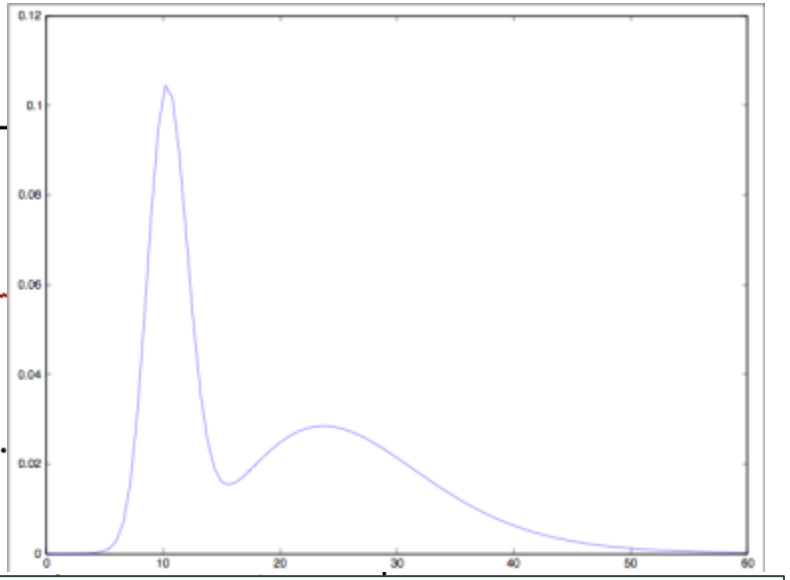
Behavioral Analytics
Learning Distributions



prelert®

Behavioral Analytics
Learning Distributions



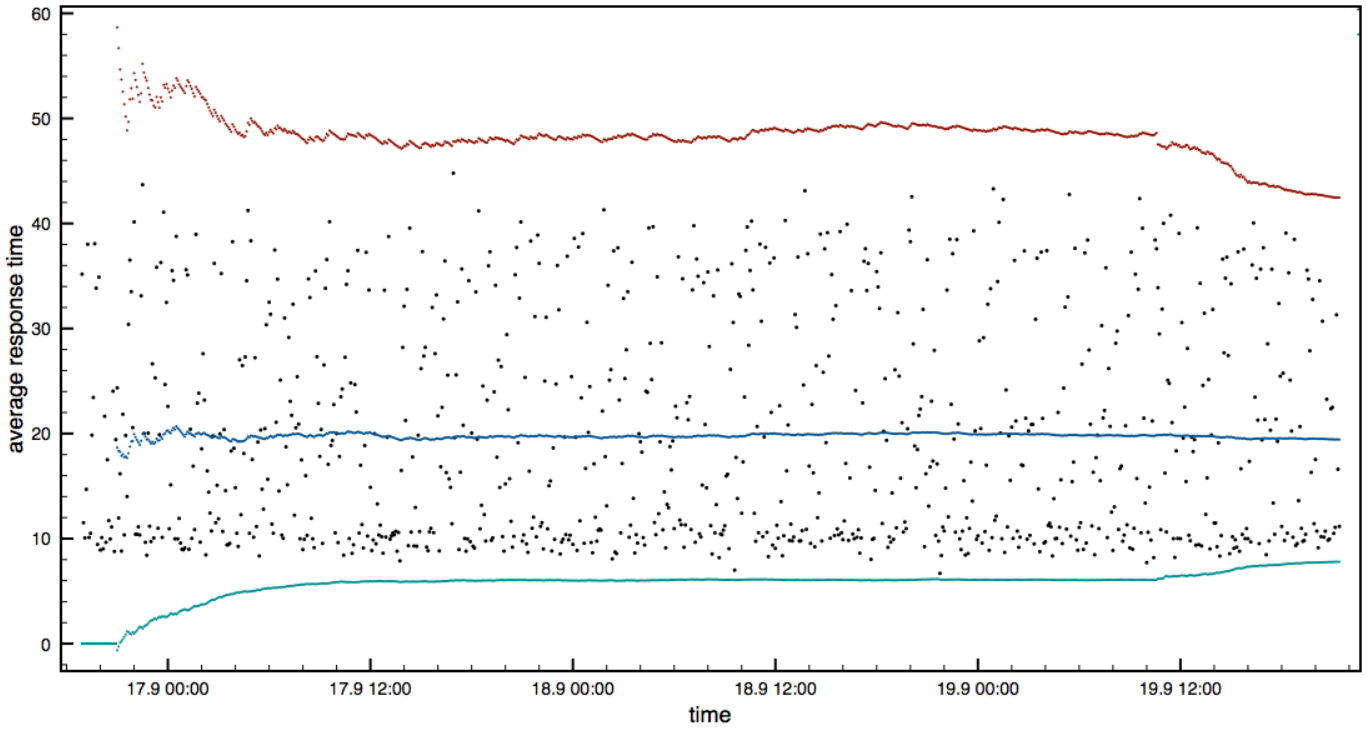


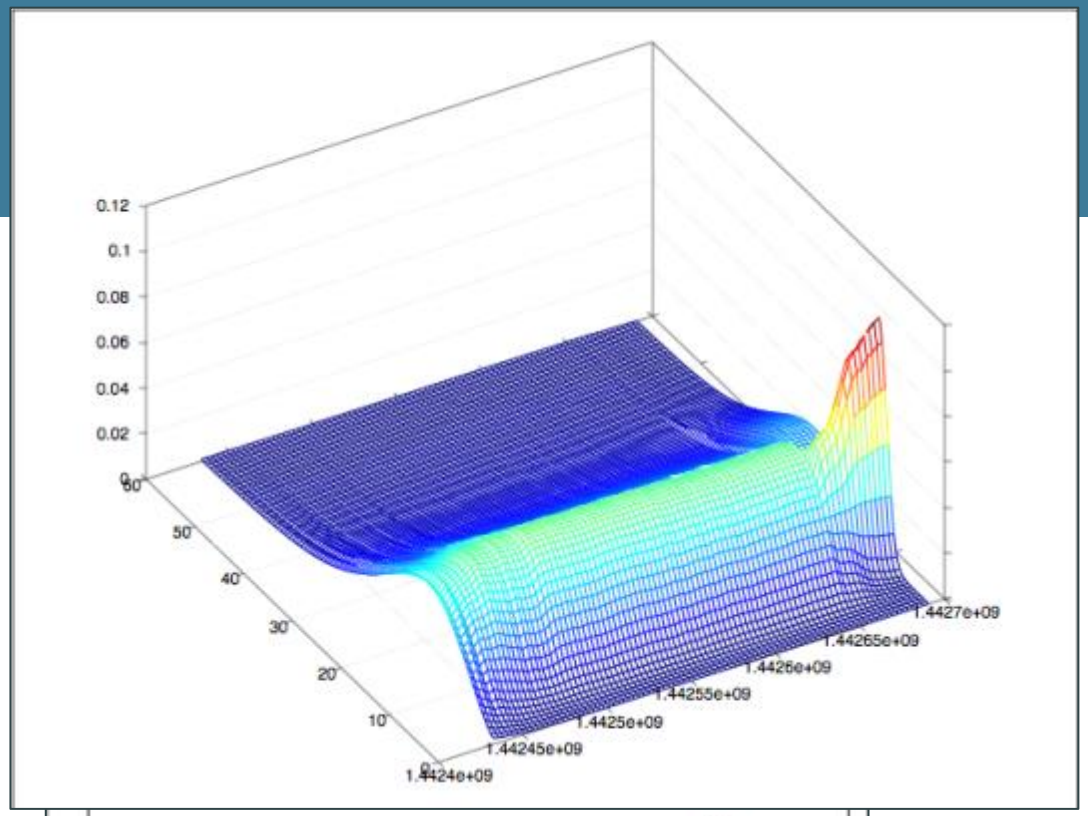
one-of-n:
weight 0.25 gamma mean = 18.38113 std = 14.32349
weight 0.25 log-normal mean = 18.30681 std = 10.18887
weight 0.25 normal mean = 18.36997 std = 13.22049
weight 0.25 multimodal:
weight 0.25 # samples 3.664231 weight 0.25 weight 1 one-of-n:
weight 0.25 weight 1 weight 0.3333333 gamma mean = 18.38113 std = 14.32349
weight 0.25 weight 1 weight 0.3333333 log-normal mean = 18.31183 std = 10.18451
weight 0.25 weight 1 weight 0.3333333 normal mean = 18.38113 std = 13.18854

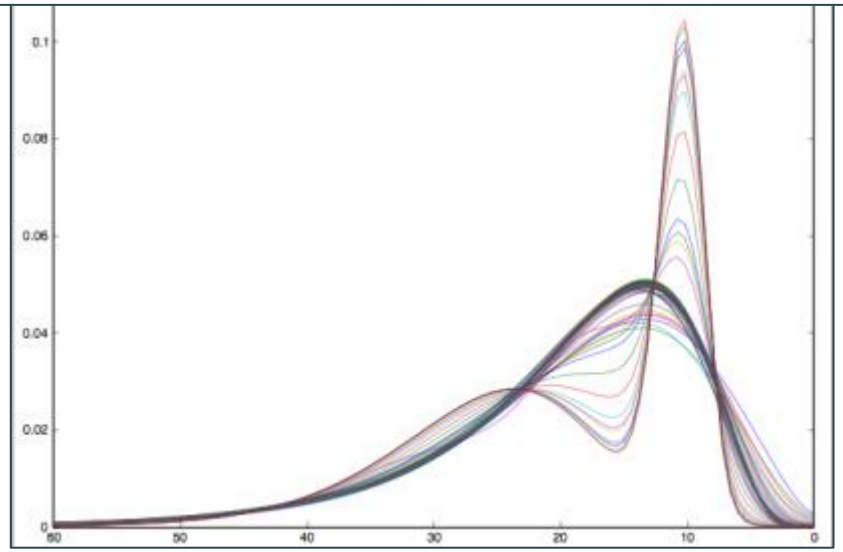
one-of-n:
weight 0.9872225 multimodal:
weight 0.9872225 # samples 114.2597
weight 0.9872225 weight 0.4498603 one-of-n:
weight 0.9872225 weight 0.4498603 weight 0.2418992 gamma mean = 10.52094 std = 1.728486
weight 0.9872225 weight 0.4498603 weight 0.2711693 log-normal mean = 10.52424 std = 1.745909
weight 0.9872225 weight 0.4498603 weight 0.4869315 normal mean = 10.52094 std = 1.722452
weight 0.9872225 weight 0.5501397 one-of-n:
weight 0.9872225 weight 0.5501397 weight 0.3222726 gamma mean = 26.6806 std = 8.068258
weight 0.9872225 weight 0.5501397 weight 0.5707544 log-normal mean = 26.69812 std = 8.290151
weight 0.9872225 weight 0.5501397 weight 0.1069731 normal mean = 26.6806 std = 8.097122

prelert®

Behavioral Analytics
Learning Distributions

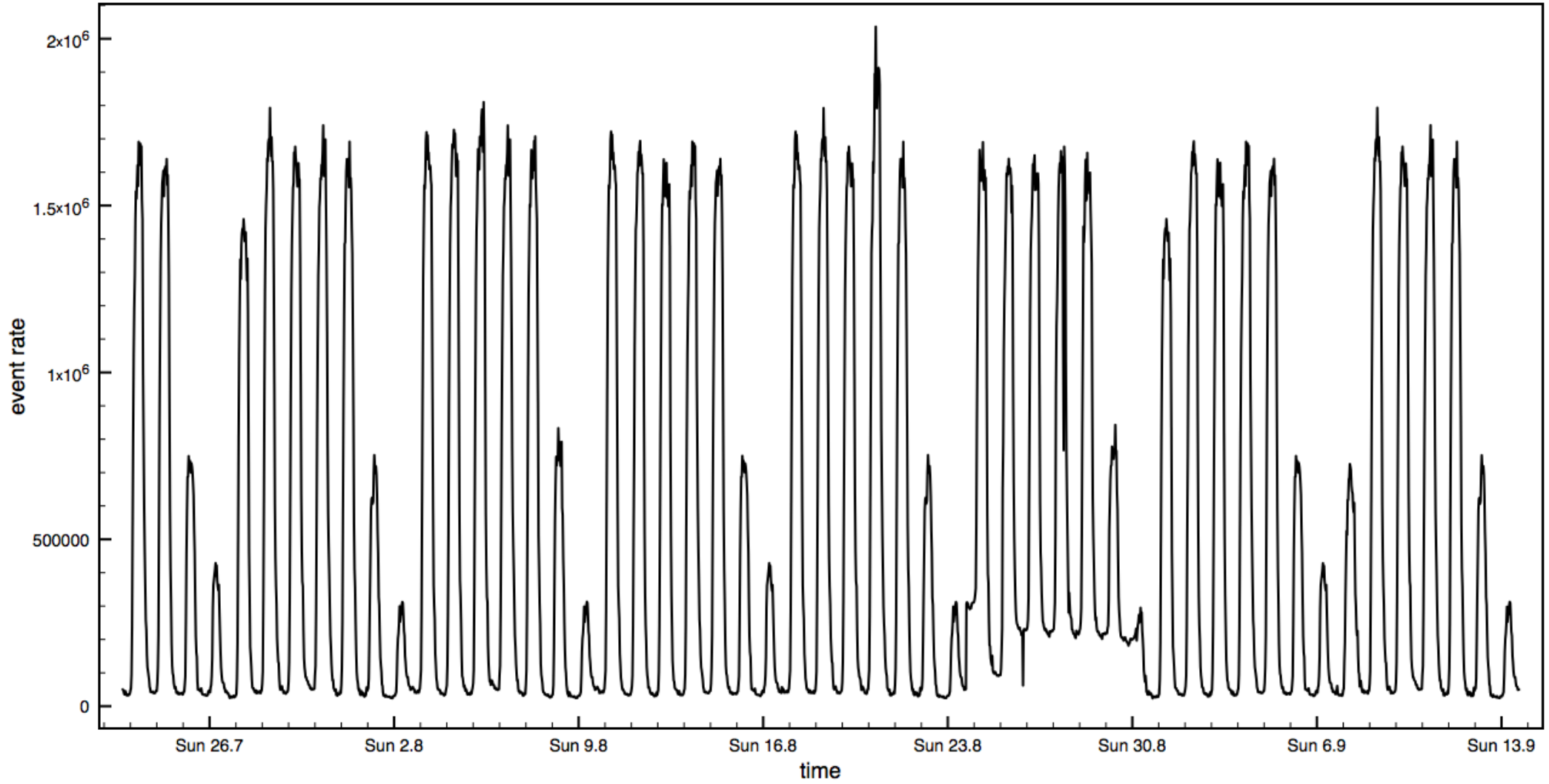






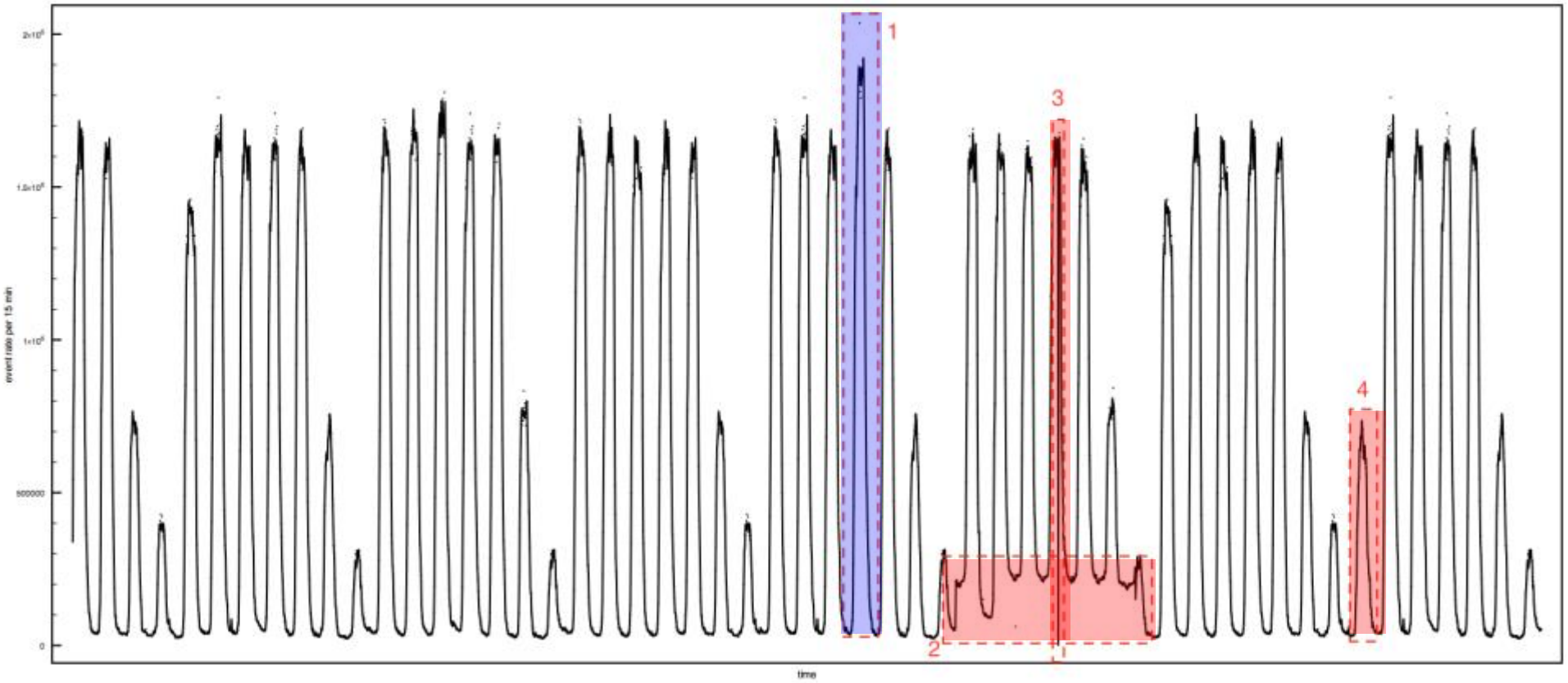
prelert®

Behavioral Analytics
Learning Periodicity



prelert®

Behavioral Analytics
Learning Periodicity

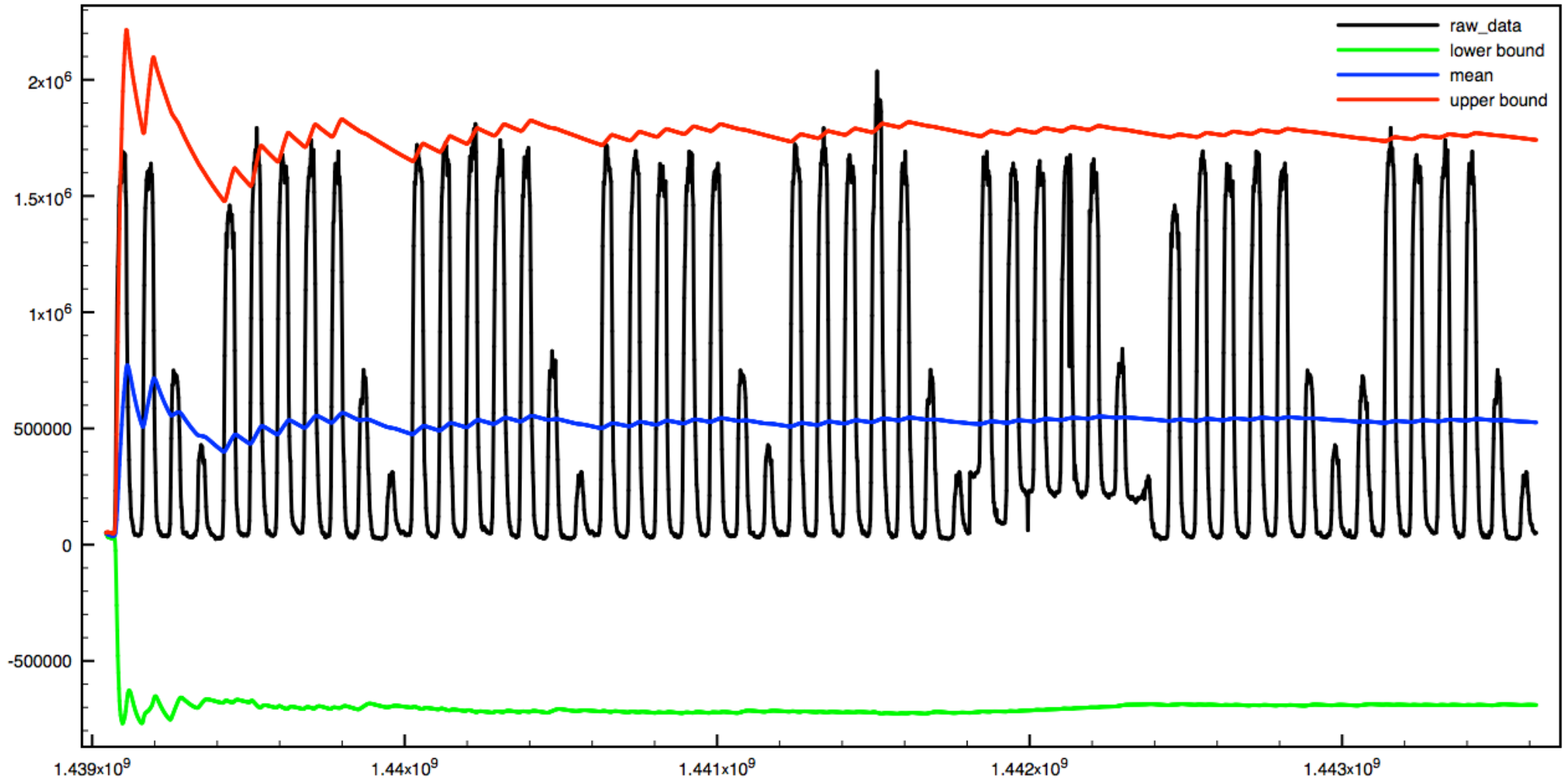


1 – Slightly higher rate
2 – Abnormal base rate over entire week
3 – Significant drop in rate
4 – Unusually low rate compared to typical Mondays

prelert®

Behavioral Analytics

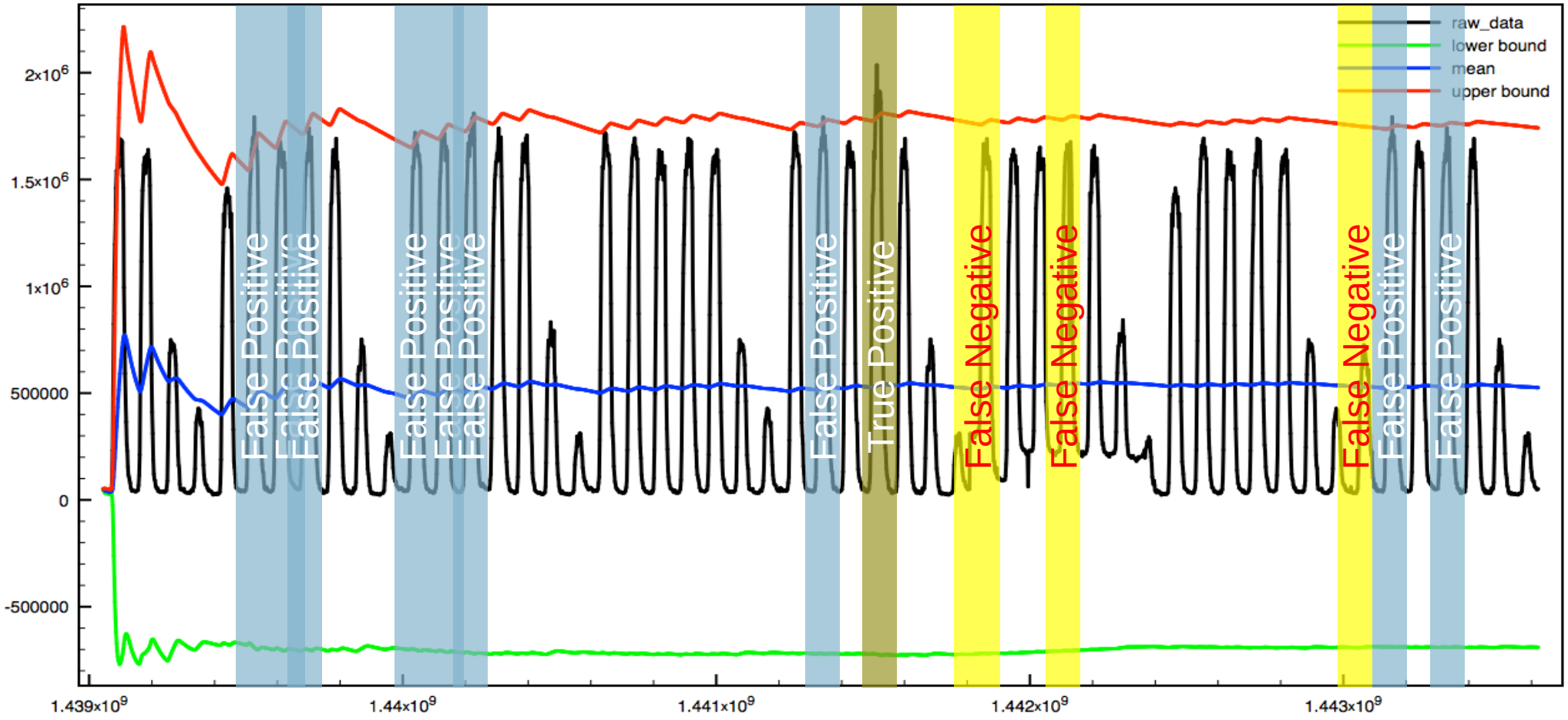
Learning Periodicity – Ineffectiveness of Simple Statistics



prelert®

Behavioral Analytics

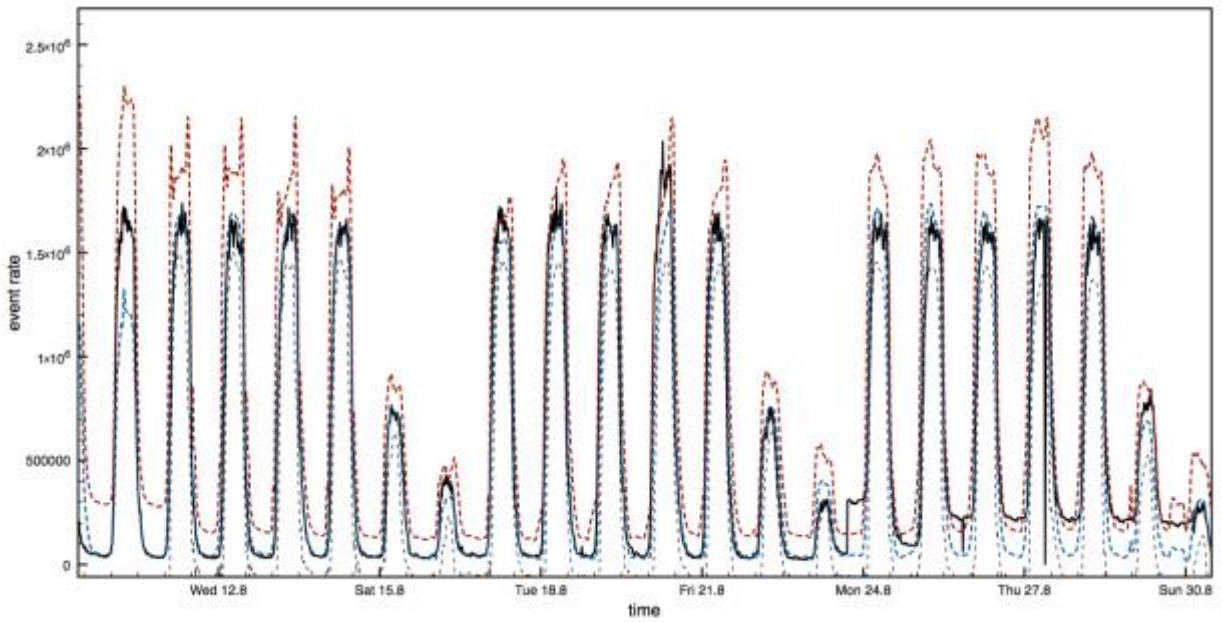
Learning Periodicity – Ineffectiveness of Simple Statistics



prelert®

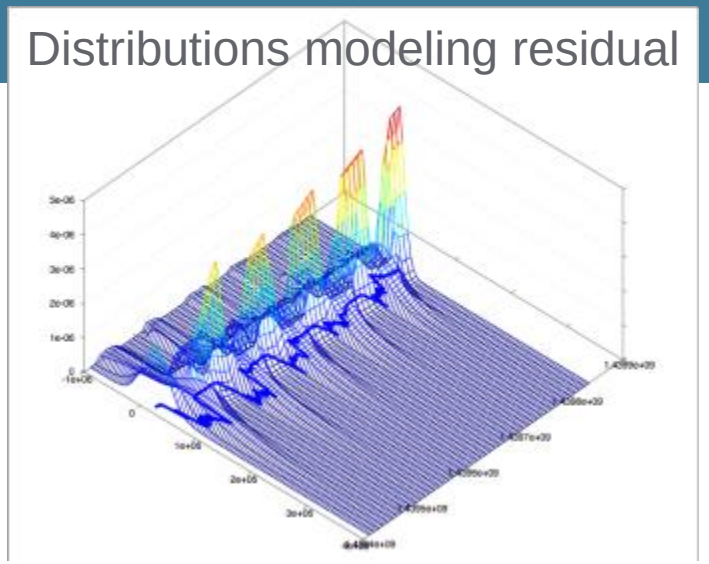
Behavioral Analytics
Learning Periodicity

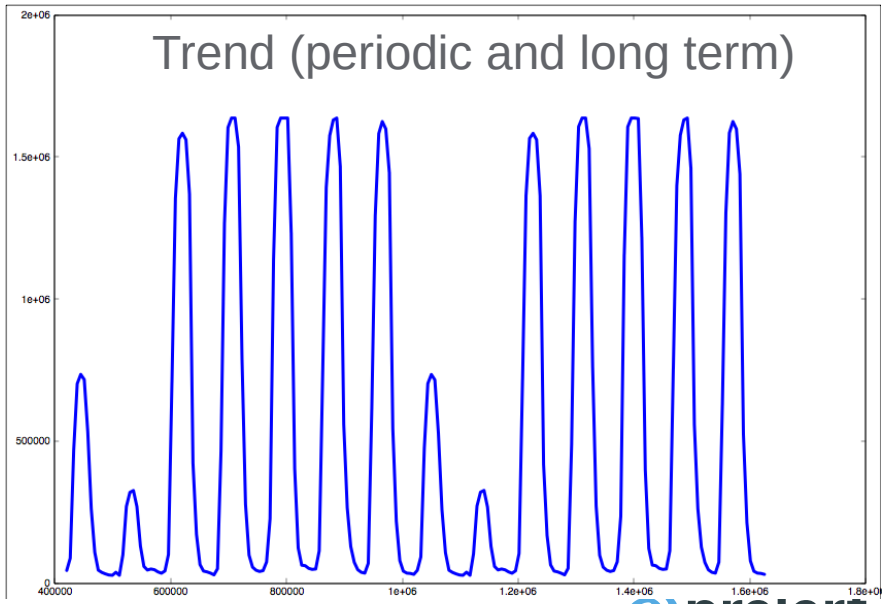
Predictive Model



=

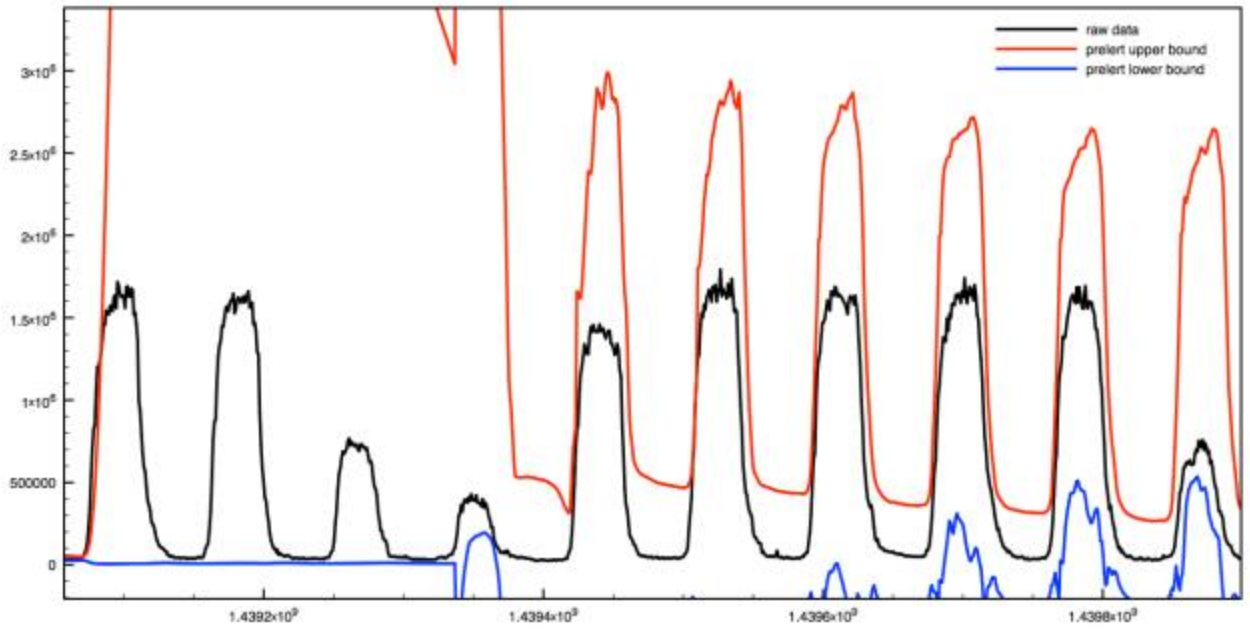
Distributions modeling residual

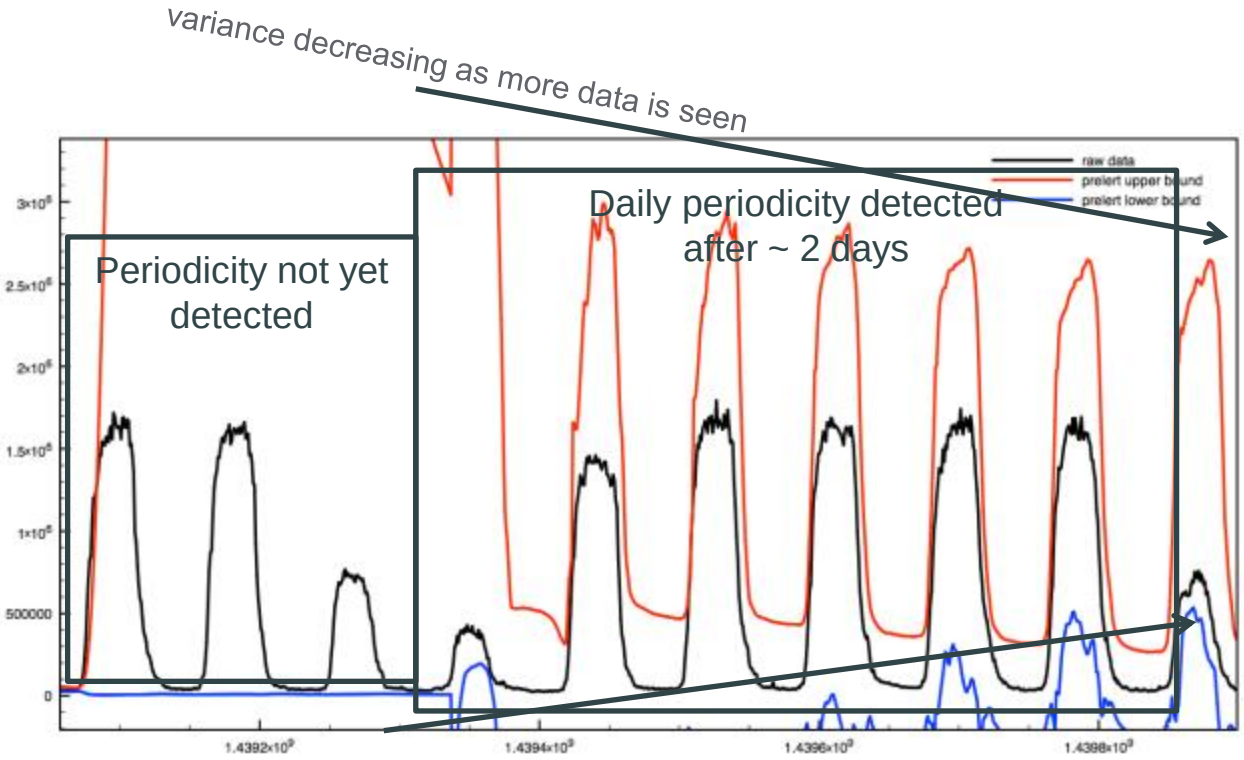




prelert

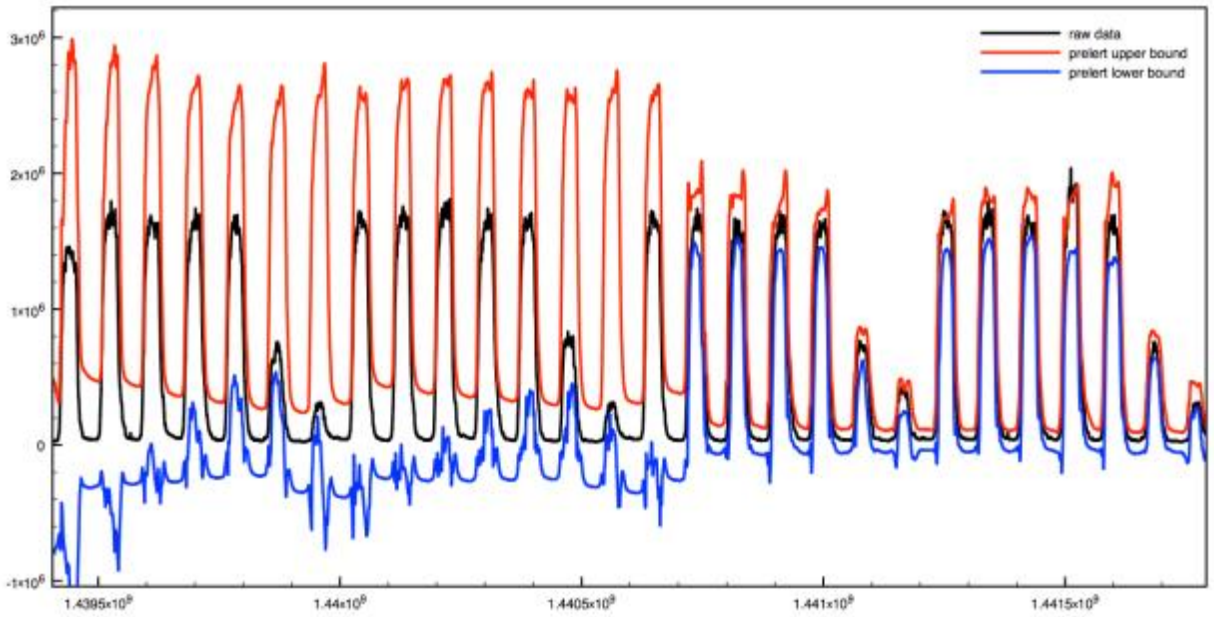
Behavioral Analytics
Automatically Learning Periodicity

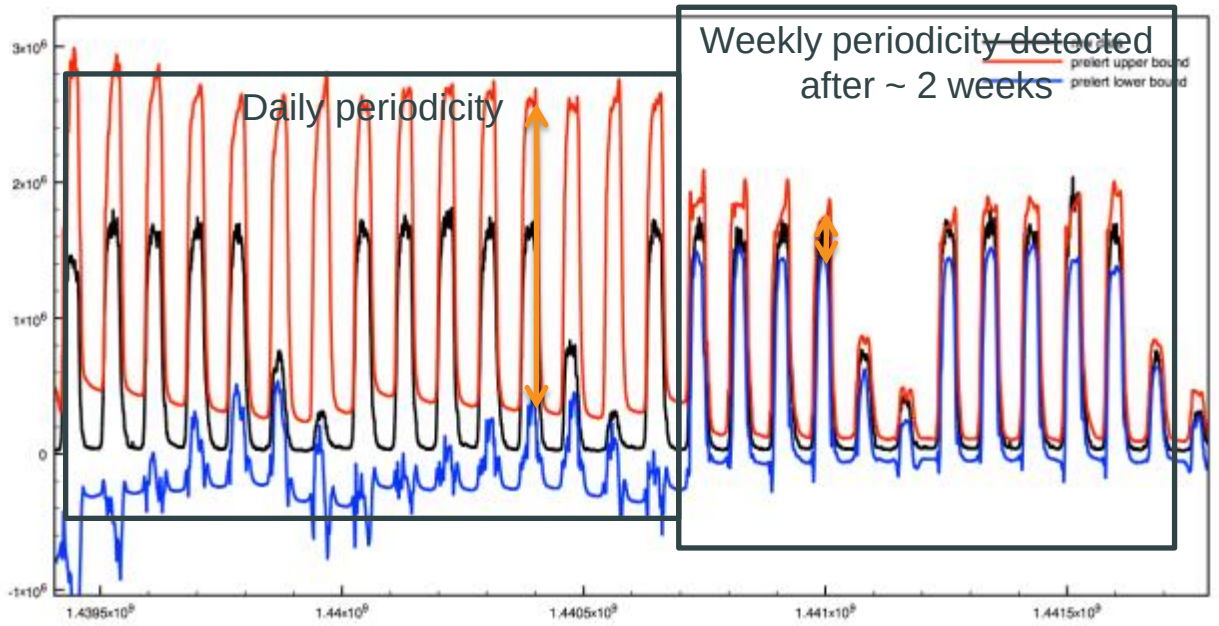




prelert®

Behavioral Analytics
Automatically Learning Periodicity





prelert®

Behavioral Analytics

Automatically Learning Periodicity – Comparison to Holt-Winters

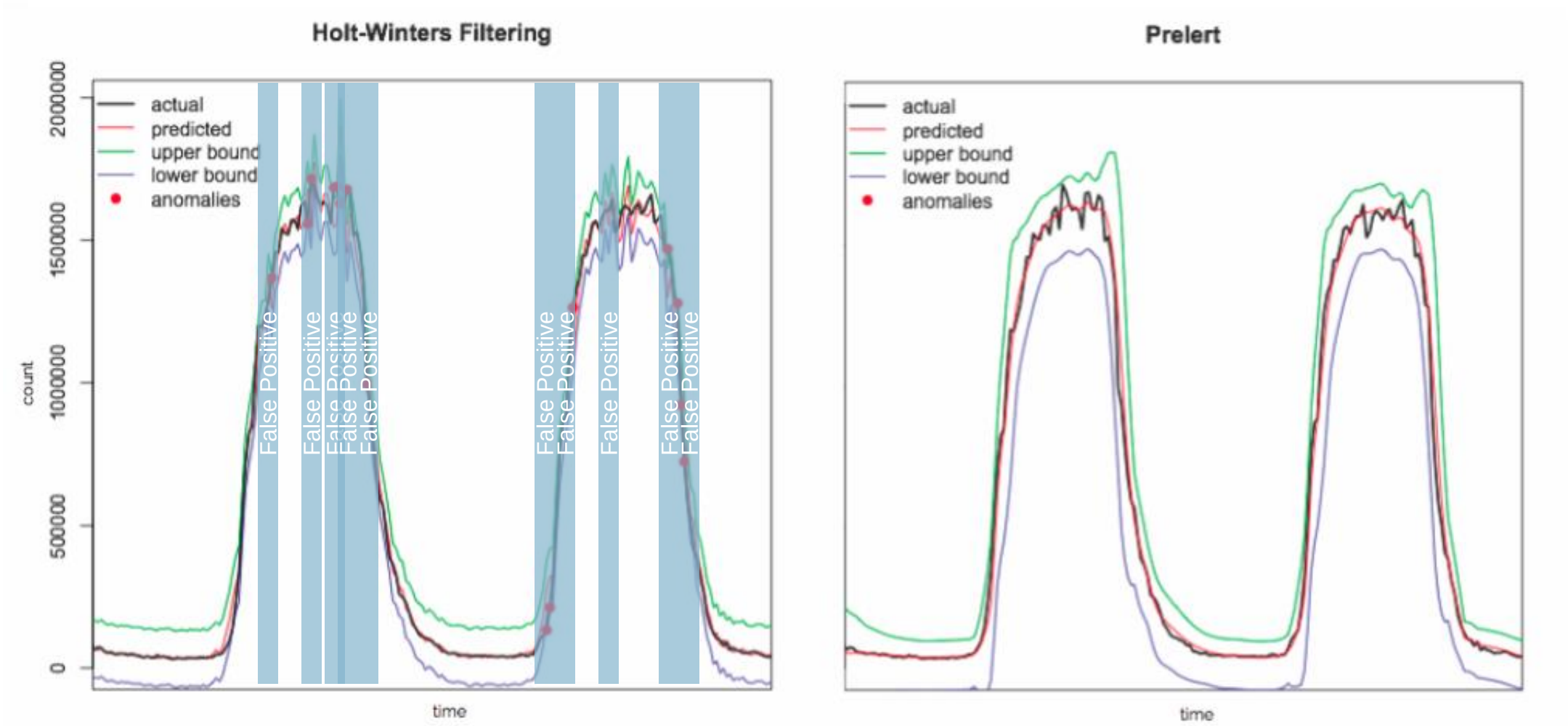


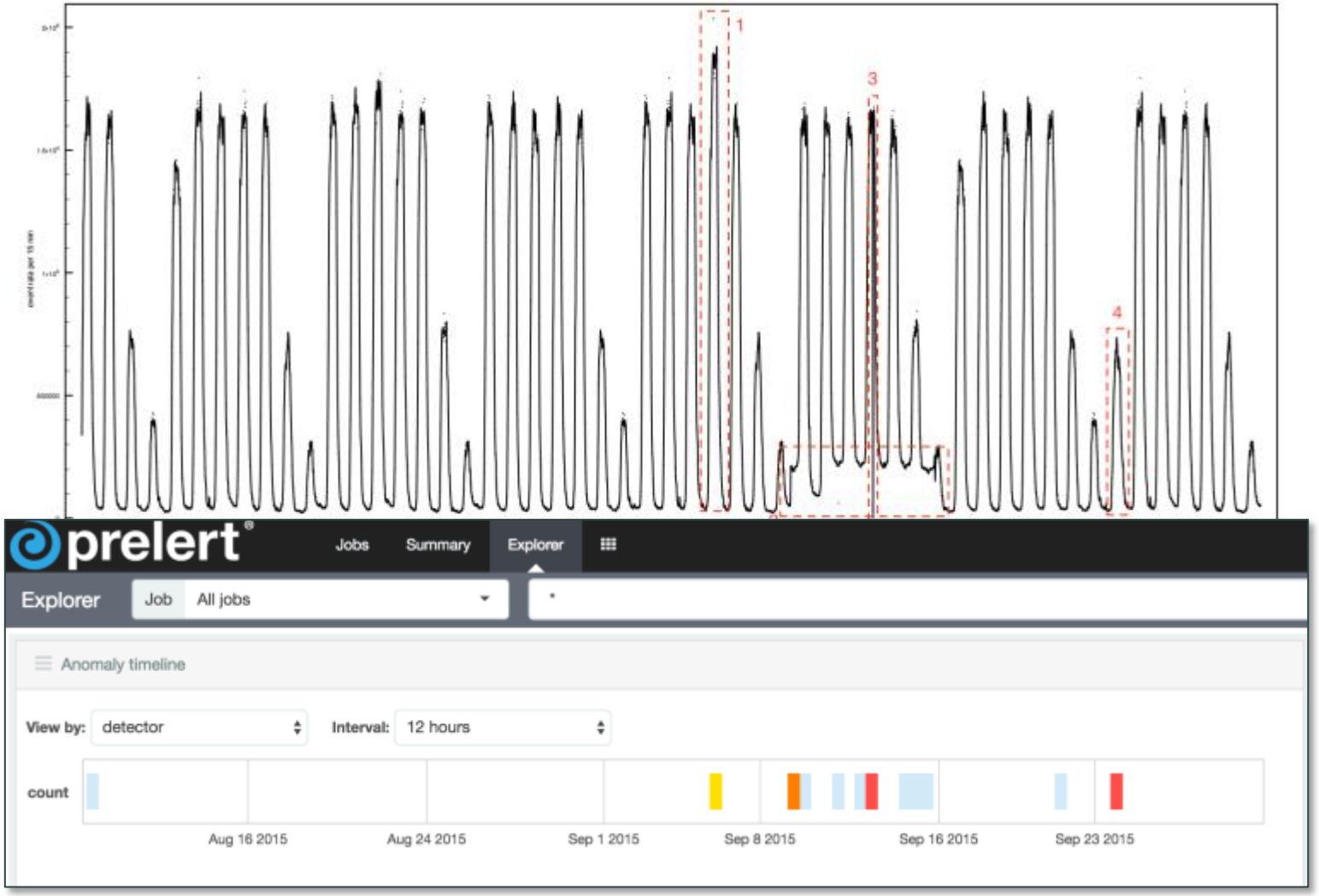
Figure 7: Holt-Winters Filtering result, left, illustrates over-fitting and false positives. Prelert Anomaly Detective result, right, illustrates accurate fit and zero false positives

http://info.prelert.com/prelert-analytics-vs-holt-winters-filtering

prelert®

Behavioral Analytics

Automatically Learning Periodicity - Results

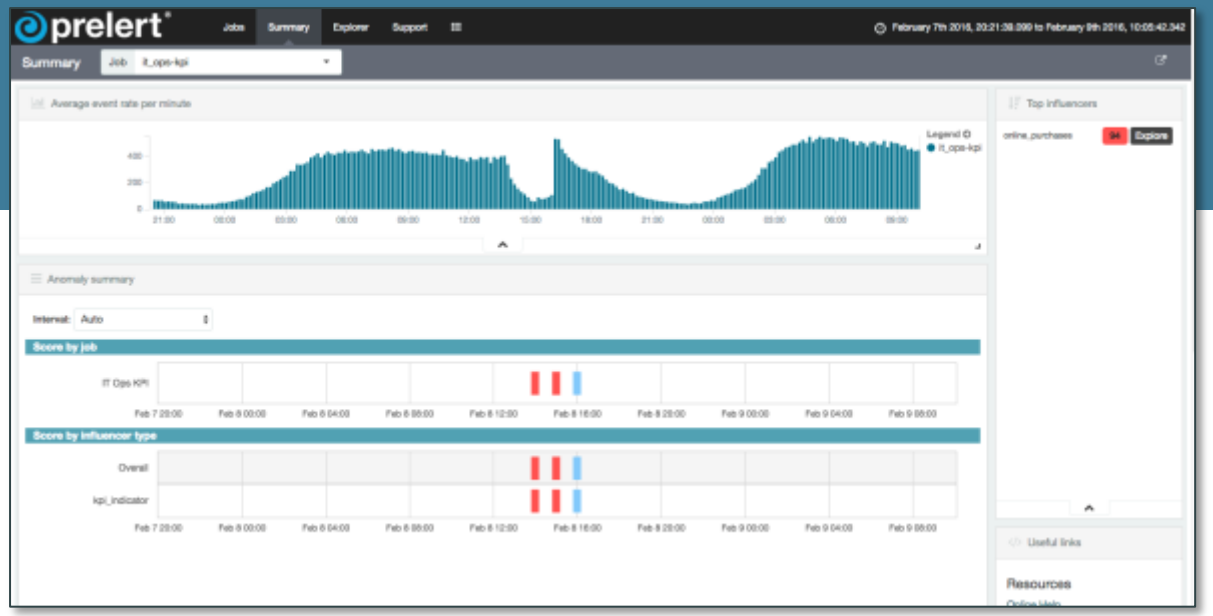


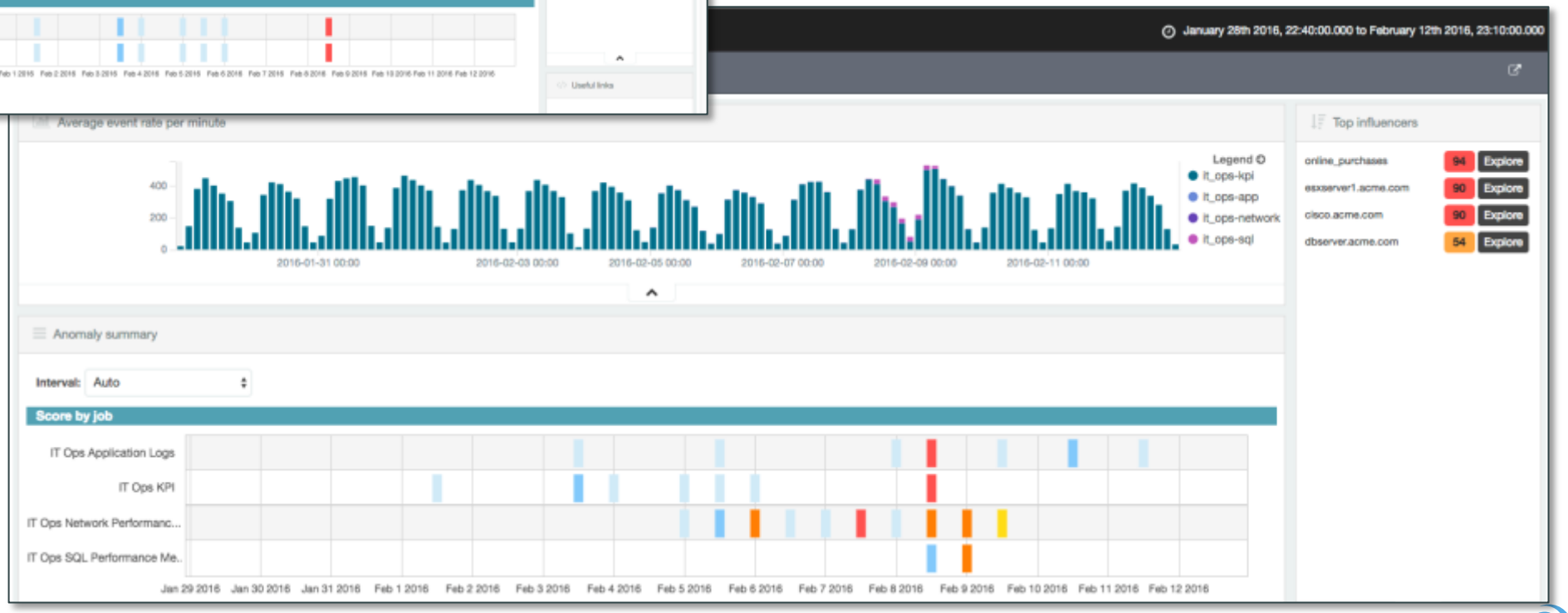
prelert

Anomaly Detection for Elasticsearch – DEMO!

prelert®

Demo Screens





prelert

Thank You!

http://info.prelert.com/products/behavioral-analytics-for-the-elastic-stack

prelert®