GP-based Electricity Price Forecasting

Alberto Bartoli, Giorgio Davanzo, Andrea De Lorenzo, Eric Medvet
DI[3] — Università di Trieste, Italy

`http://bartoli.inginf.units.it`

April 2011

Scenario: Electric Power Auctions

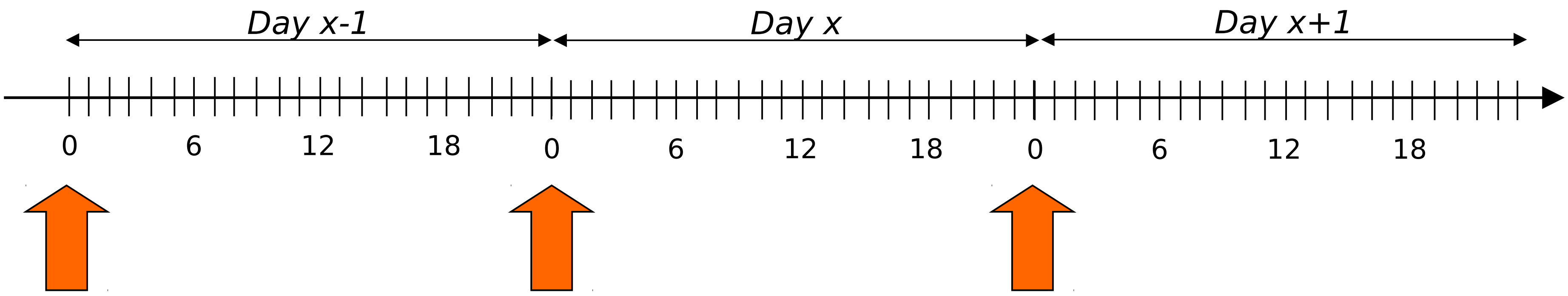
- Electric power market increasingly relying on **auctions**

- **Producers** offer < quantity, selling price >
- **Consumer** offer < quantity, buying price >

- **Central Authority** establishes energy flows and settling price

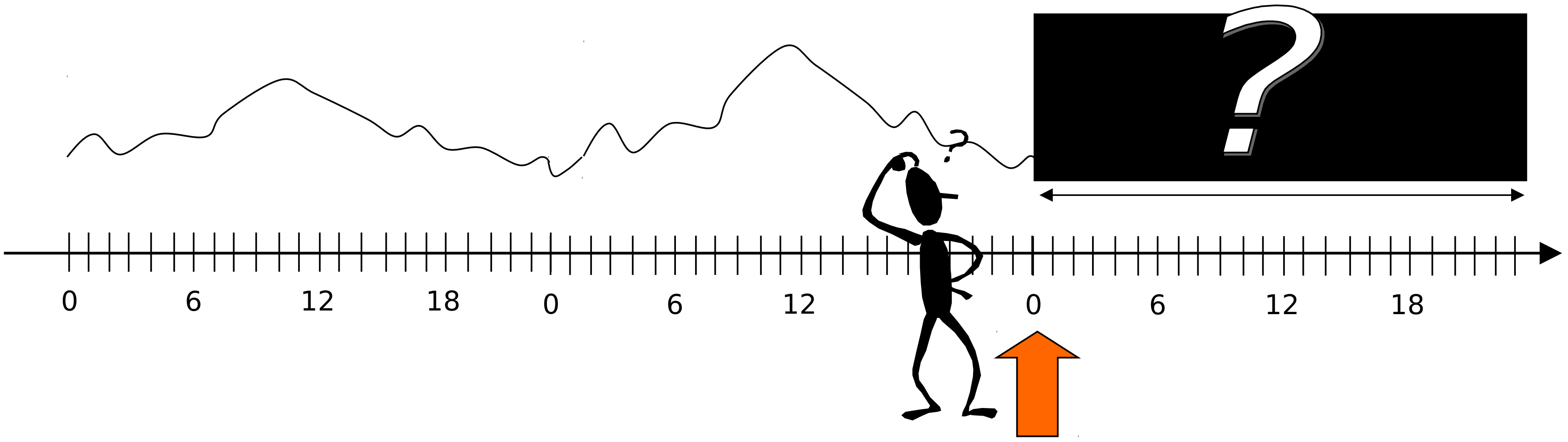
- Two forms:
  - **Day-ahead**
  - Hour-ahead

Scenario:
Day-Ahead Auction



- Each auction involves generating **24** pairs < quantity, **price**>
- One for **each hour** of the **next day**

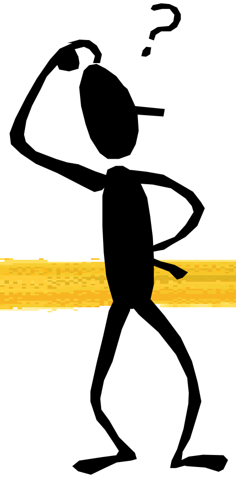
Which Price ?



- Producer: Which selling price ?

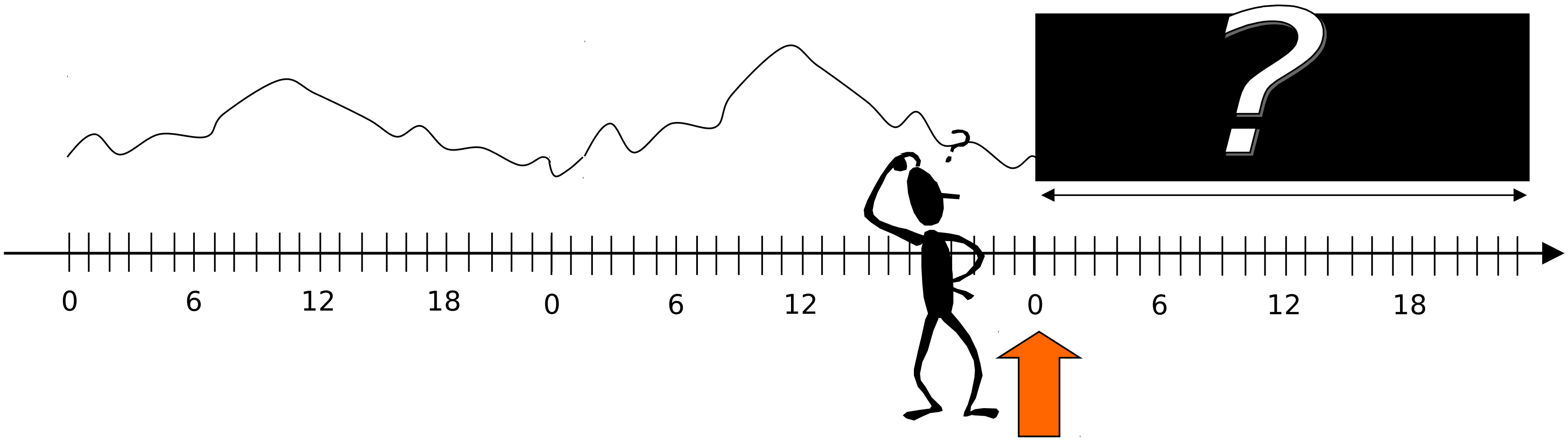
- Consumer: Which buying price ?

Which price ?



- **Producer: Which selling price ?**
  - High, obviously
  - …But not “too high”
    (otherwise it might not sell all the energy it needs to sell)
  - Must select a price in line with the final price for the next day
- **Consumer: Which buying price ?**
  - Low, obviously
  - …But not “too low”
    (otherwise it might noy buy all the energy it needs to buy)

Price Forecasting



- Price forecasting is **essential** for maximizing revenues
  - For **all actors** and for **any bidding strategy**
  - IEEE Trans on Power Systems (2005-): 23 papers (!)

- Note: Bad forecasts last for the full day

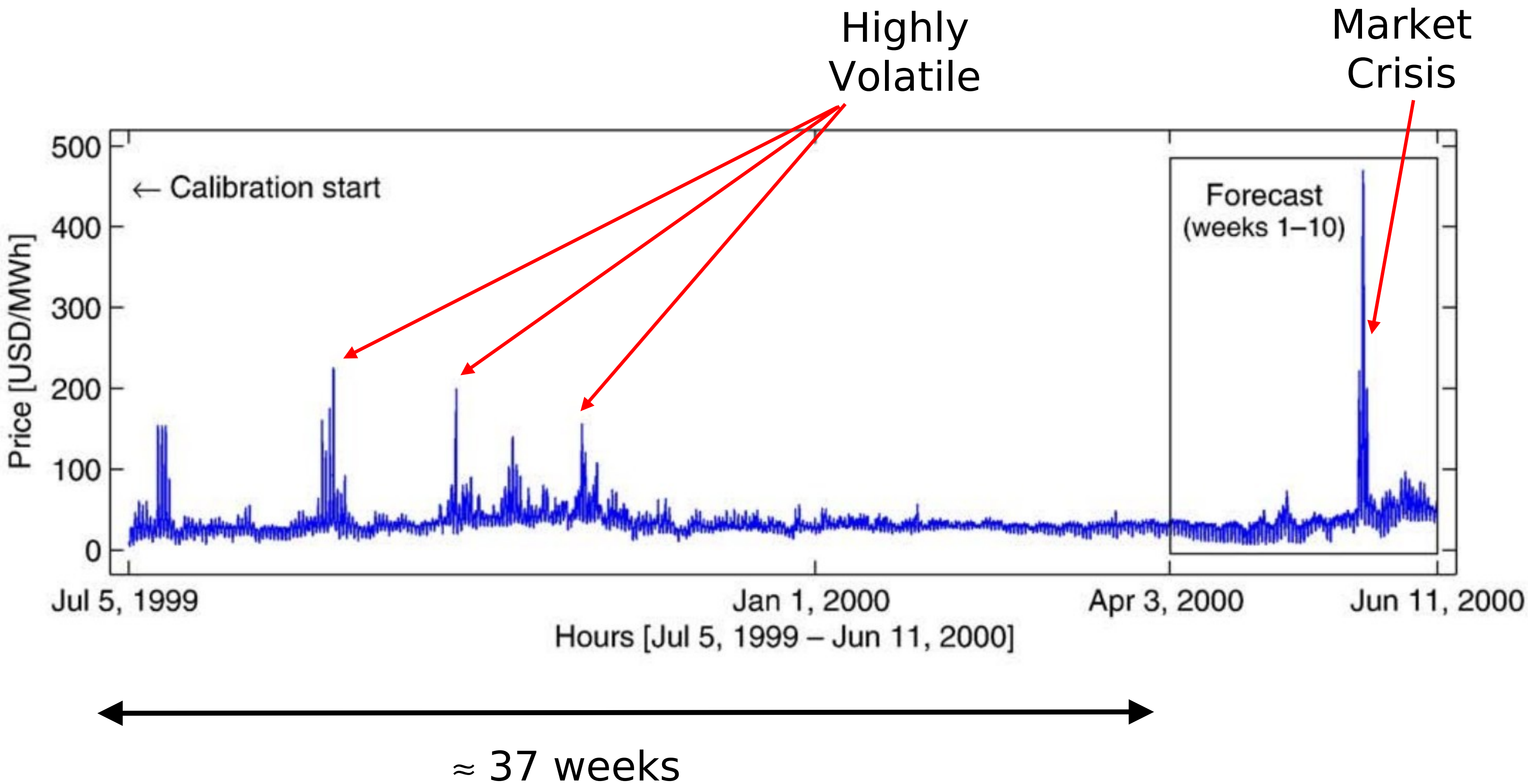
Our contribution

- Day-ahead price forecasting
- **Hybrid estimator**
  - GP-based
  - Neural network-based for **outliers**

- Assessed against a very challenging baseline

- No exogenous variables
  - More practical and simpler to implement
  - E.g., where and when one should measure Temperature ?

Baseline: Dataset (California 1999/2000)



Baseline:
Forecasting models

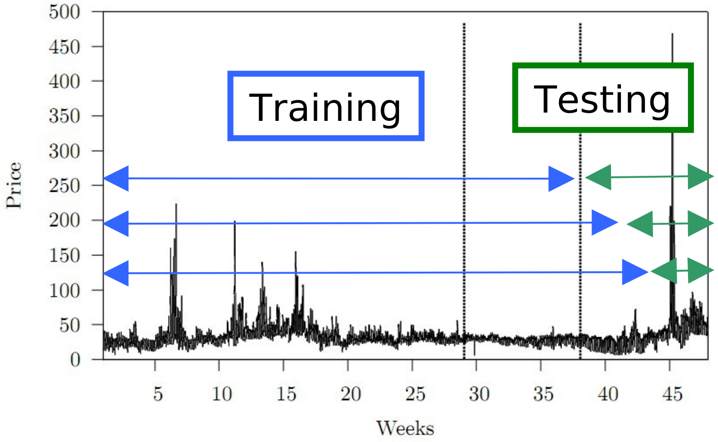
- Highly challenging
  Weron and Miesorek
  International Journal of Forecasting, 2008

- **12** models proposed earlier in the literature
  - 6 with an exogenous variable (load)
  - 6 without any exogenous variables

- For each model, **24 different calibrations**—one for each hour

- Everything is **recalibrated every da**

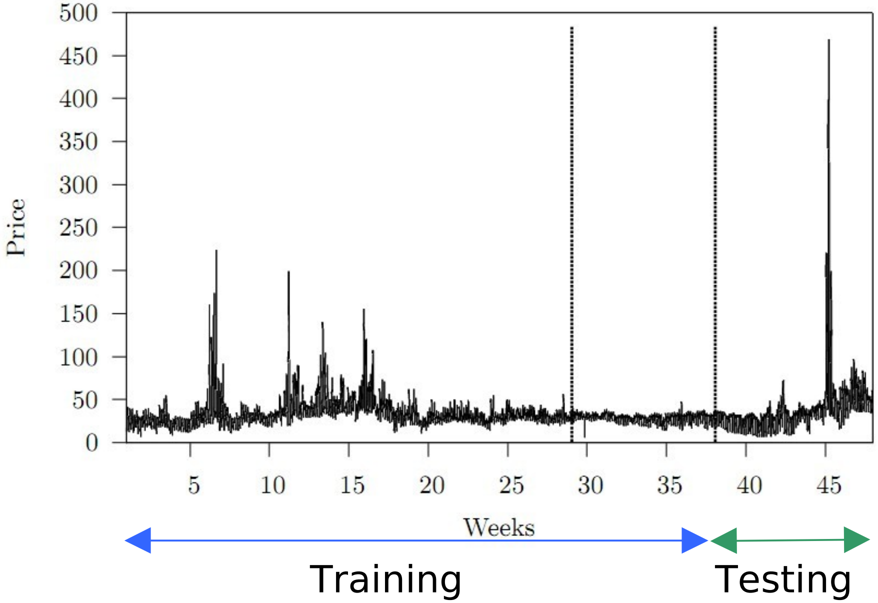
- Training data **increase every day**



Our results

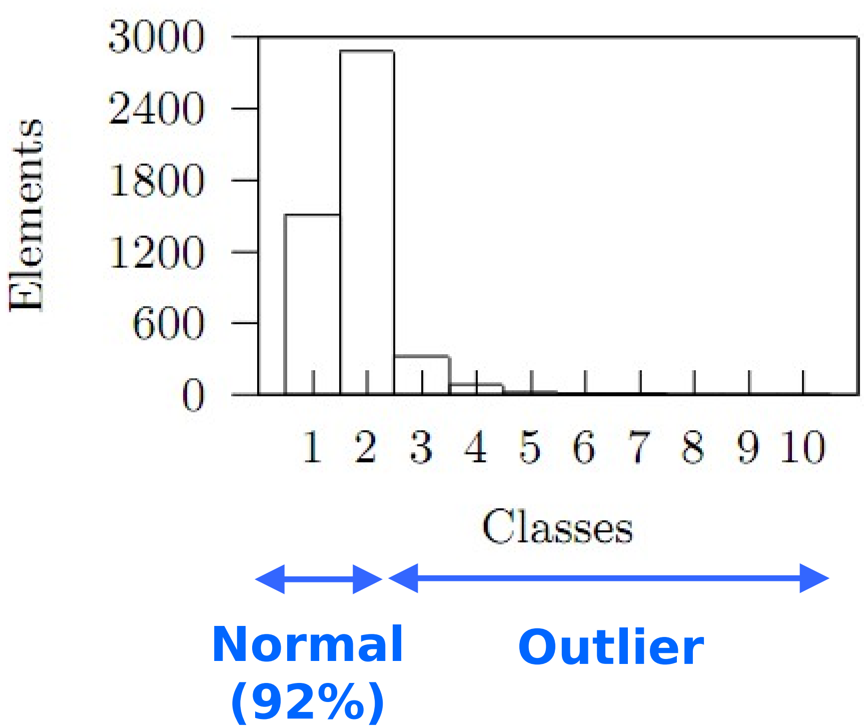
- More **accurate** than:
  - **All** 12 models
  - **Ideal** (not implementable) best-of-week estimator

- **One** single model for the full day
- **Never** recalibrated

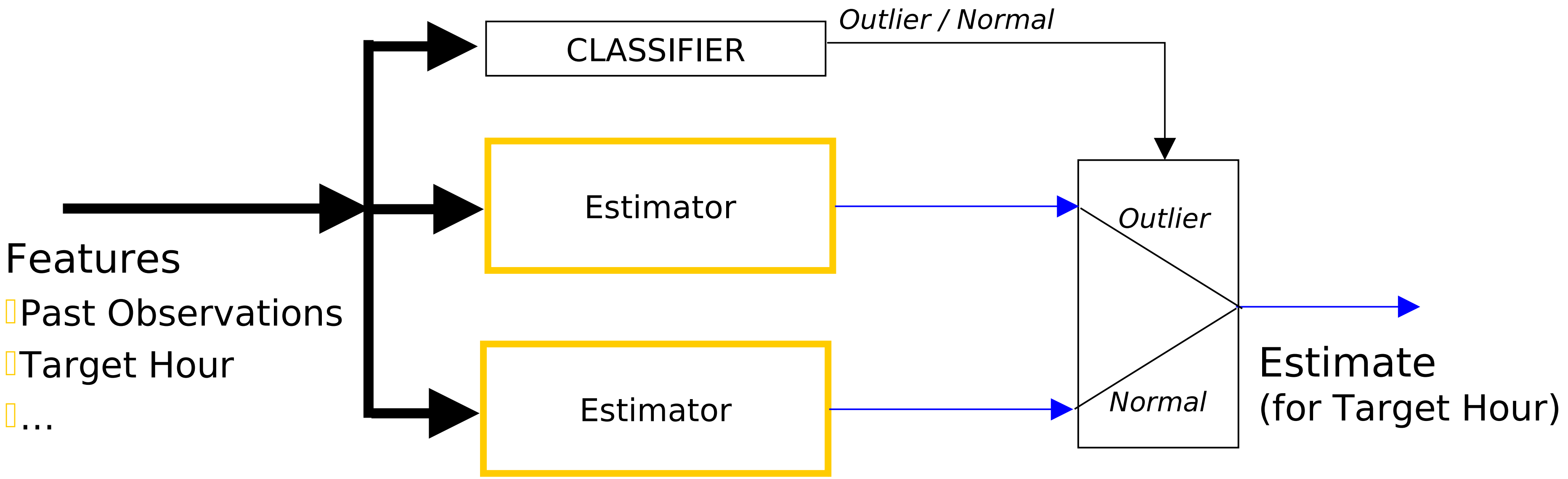


Our approach: Outliers

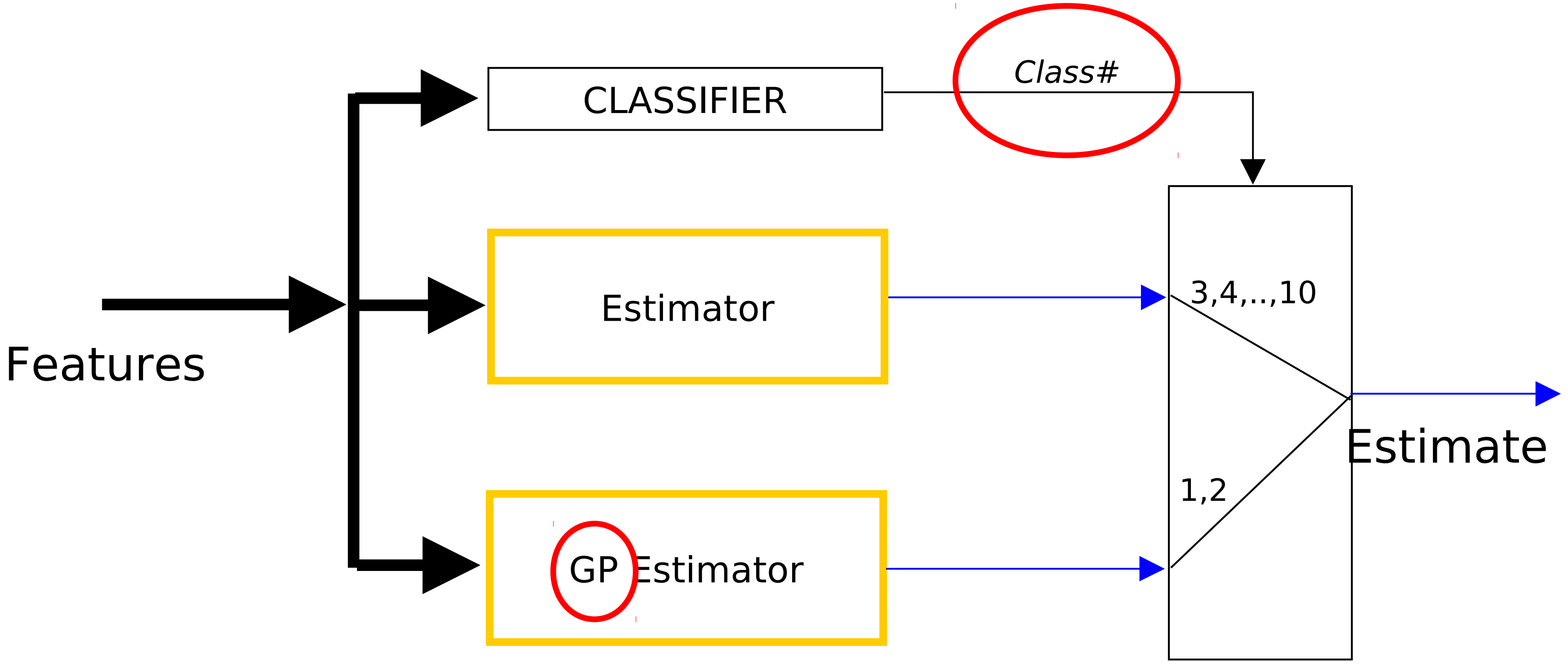
- Outlier (our definition):
  - 10 equally-sized price intervals (“classes”)
  - Classes including at least 90% of the observations are **normal**
  - Remaining classes are **outliers**



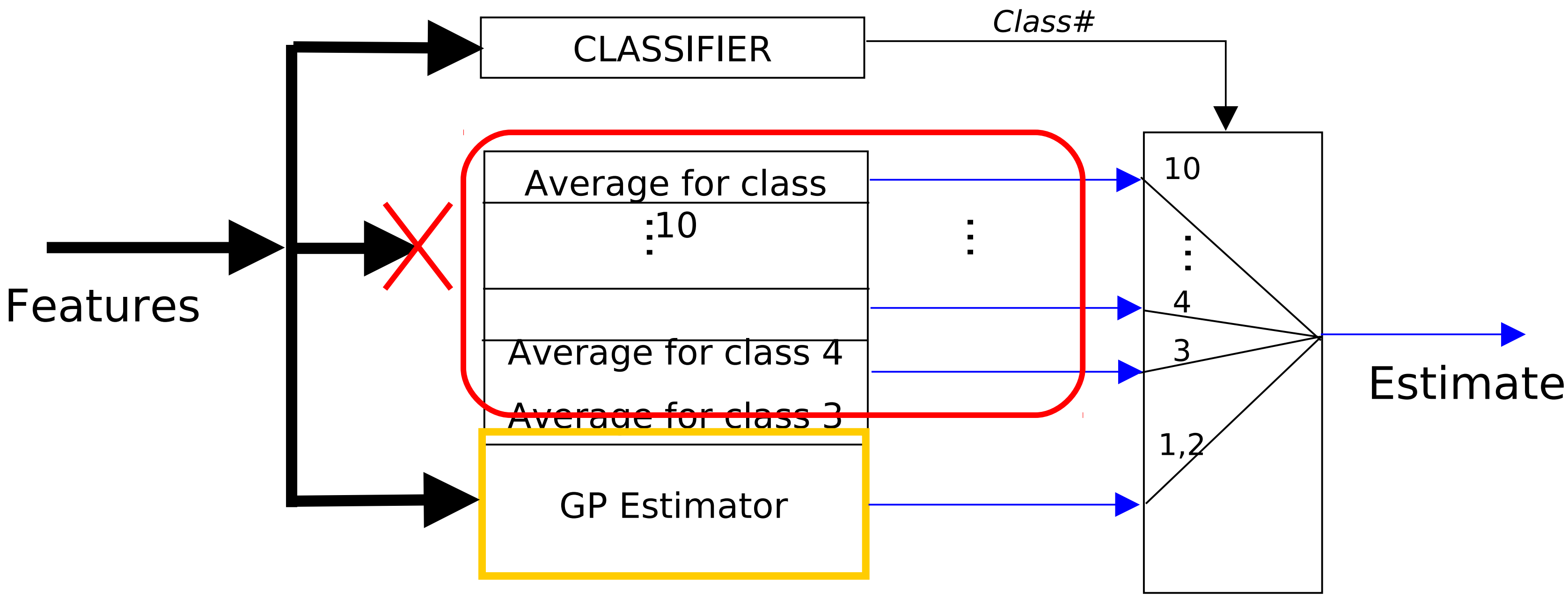
Our approach: Basic Idea



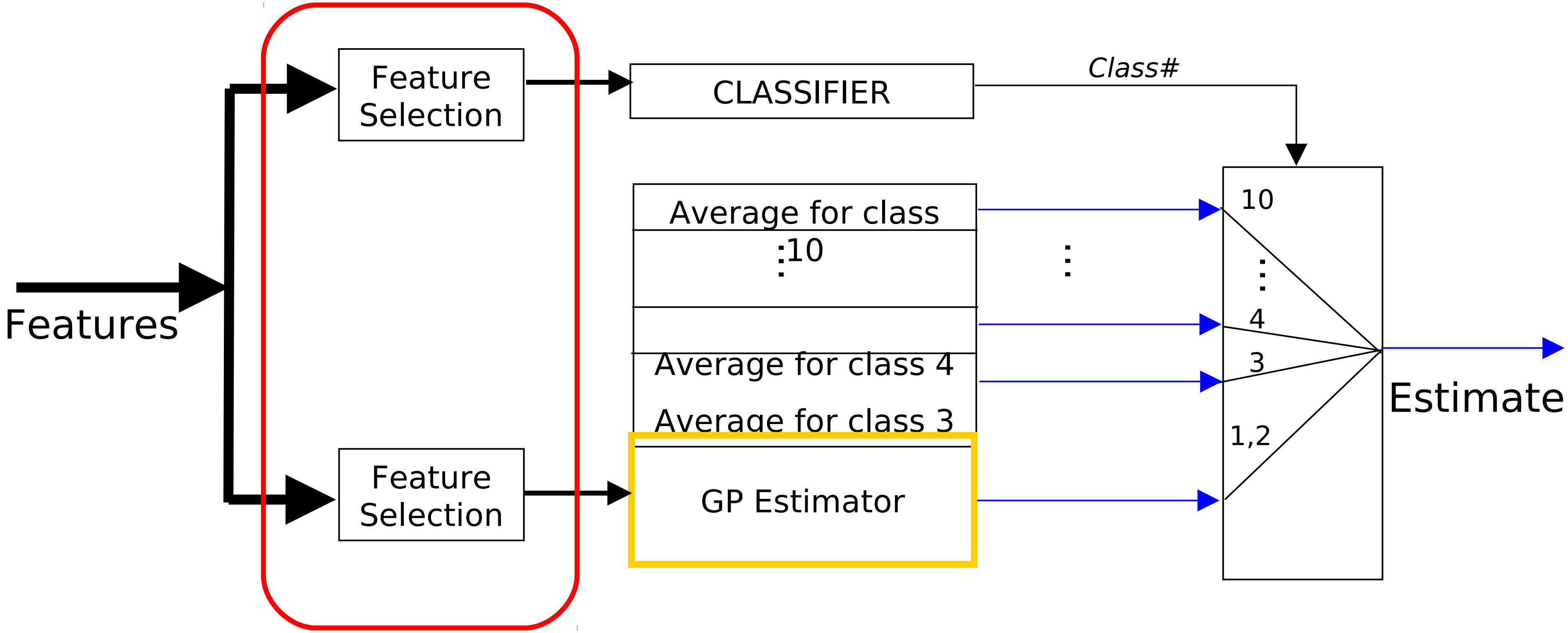
Our approach: Details (I)



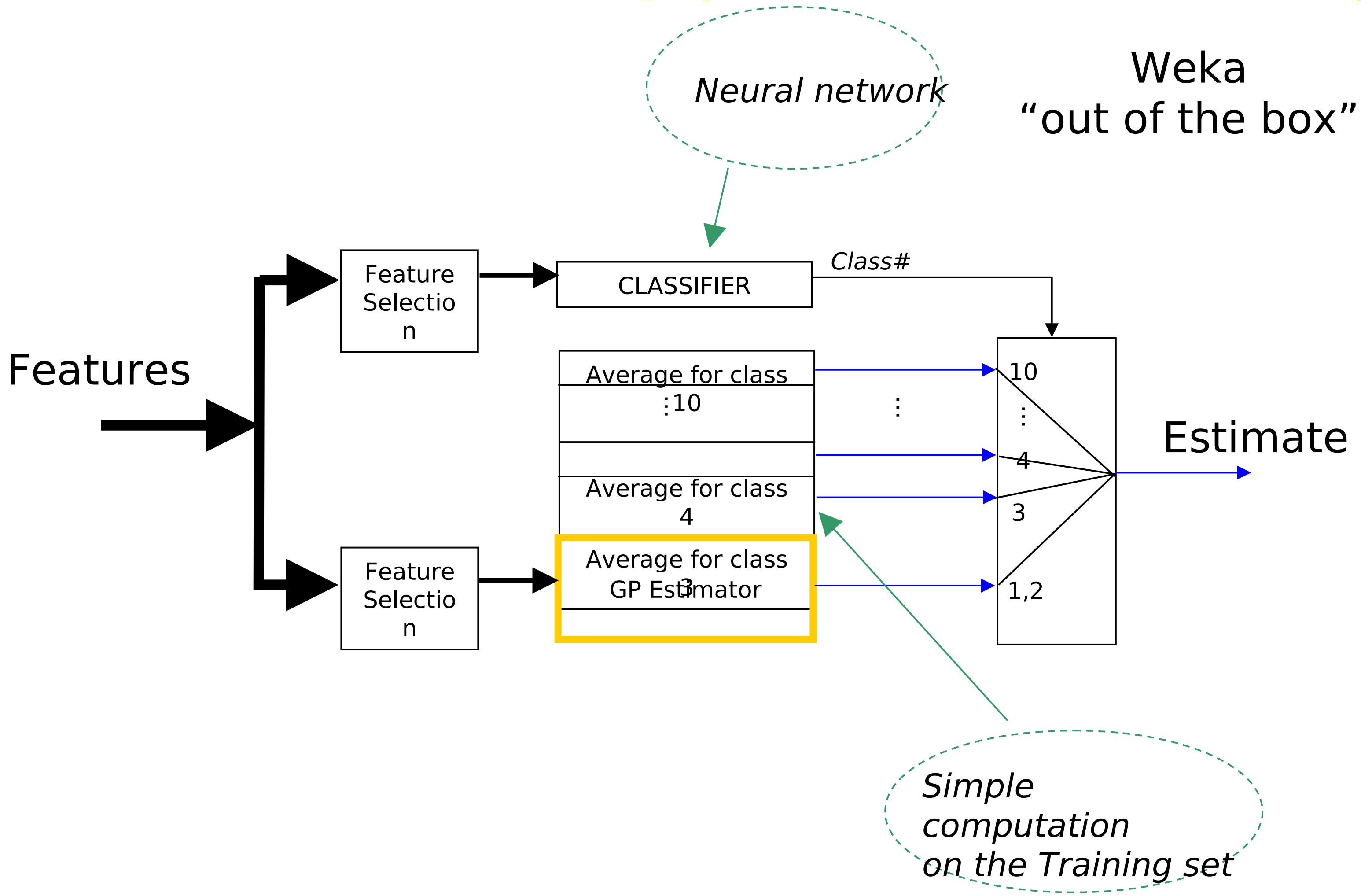
Our approach: Details (II)



Our approach (finally…)



Classifier, Outlier estimator

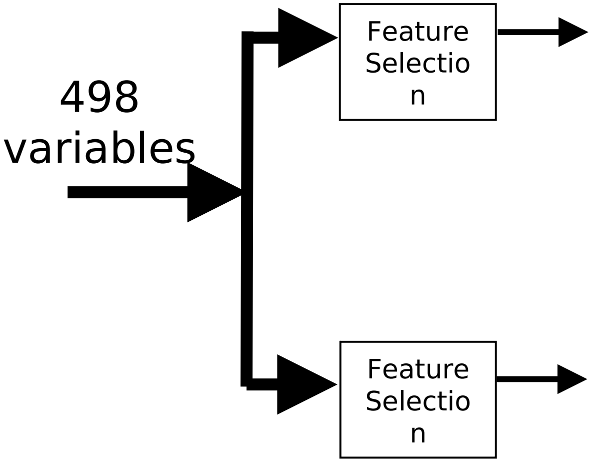


Features

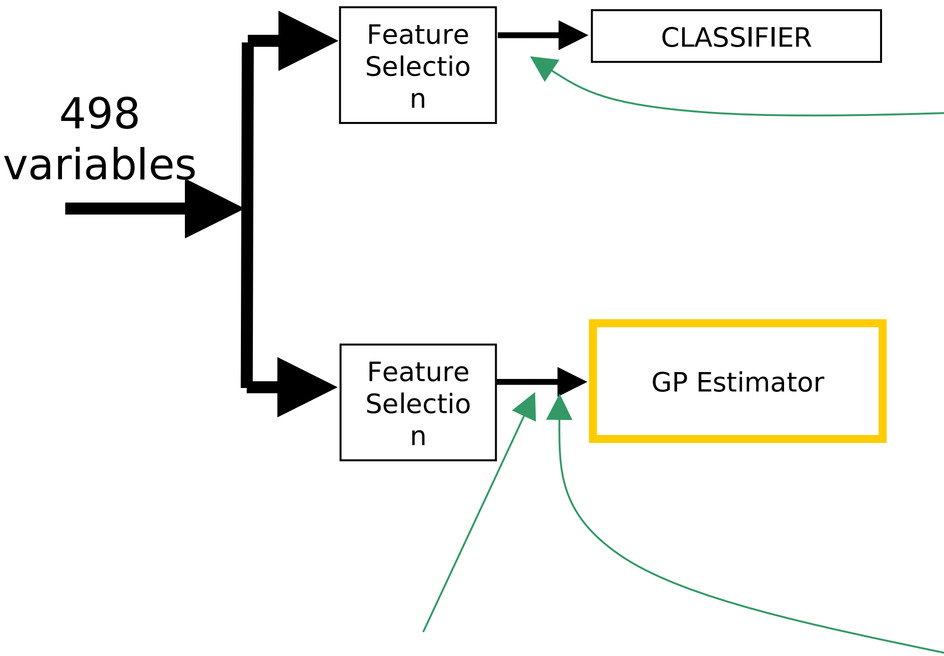
Input: 498 variables

- For each hour in the past week:
  - Observed **price**
  - A **night** flag
    (up to 5 AM)
  - A **holyday** flag
- For each day in the past week:
  - **Maximum** and **minimum** observed price
    (sort of feature "extraction")
- An enumerated variable representing the target hour
- An enumerated variable representing the target weekday

- One week in the past
- No Exogenous variables



Feature selection



- 95 Variables
- Genetic search based on decision trees (Weka out-of-the-box)

Method 1: Those of the best performing baseline method (without exogenous variables)

- Observed price at -24, -48, -168
- Holyday and night flag for target hour
- Minimum price in the previous day

Method 2 “Mutual information”

- Observed price at -24, -168
- Holyday and night flag target hour
- Holyday flag for -24, -168
- Maximum and minimum price in the previous day

Feature selection: Mutual Information

- For each feature Xi:
  - Compute mutual information mi = F(Xi, price)
  - For each feature Xj:
    - Compute mutual information mij = F(Xi, Xj)

- Repeat until “enough features”:
  - Select XH with highest mi
  - For each remaining feature Xj
    - “Adjust” mj as mj := mj - mHj

- We chose to stop with 8 features

GP Estimator (I)

- Functions Set: +, -, *, /
- Terminal Set: 0.1, 1, 10, selected features

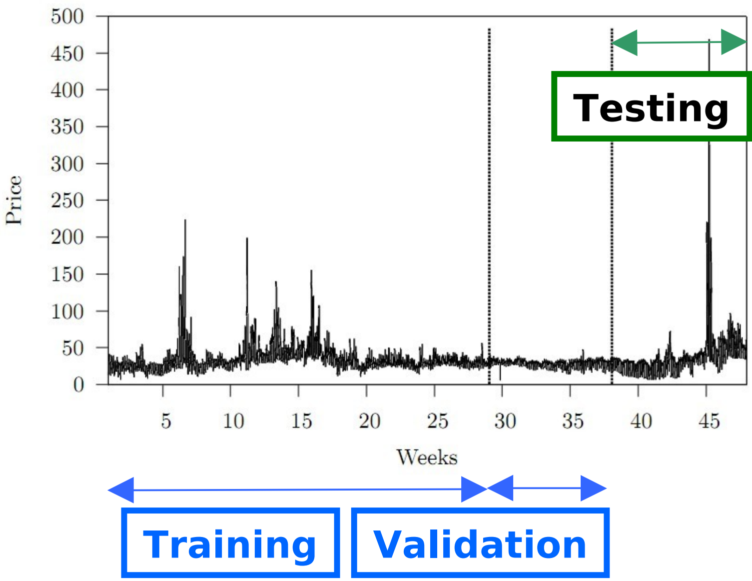
- 500 individuals, 1200 generations
  (full set of GP parameters in the paper)

- Fitness: WMAE
  - Weekly-weighted Mean Absolute Err

$$\mathrm{WMAE} = \frac{\sum_{h=1}^{168} |P_h - \hat{P}_h|}{\sum_{h=1}^{168} P_h}$$

  - Performance index of the baseline

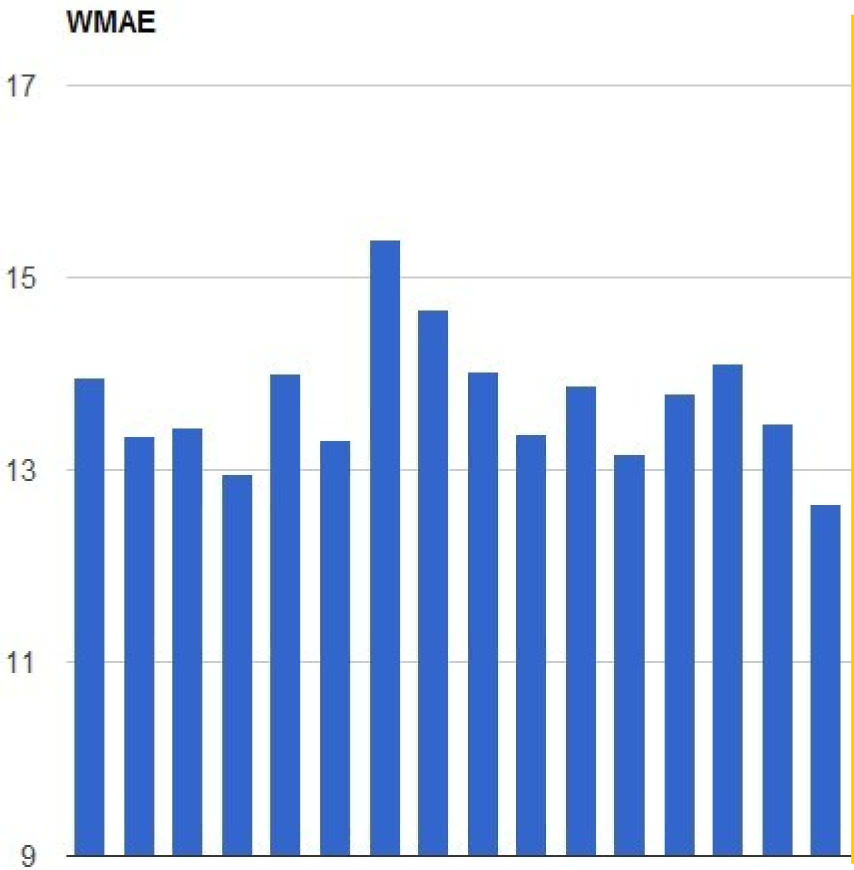
GP Estimator (II)



- 128 independent runs:
  - Evolve population of 500 individuals on the **Training set**
  - Select best individual

- Assess final population (128 individuals) on the **Validation set**

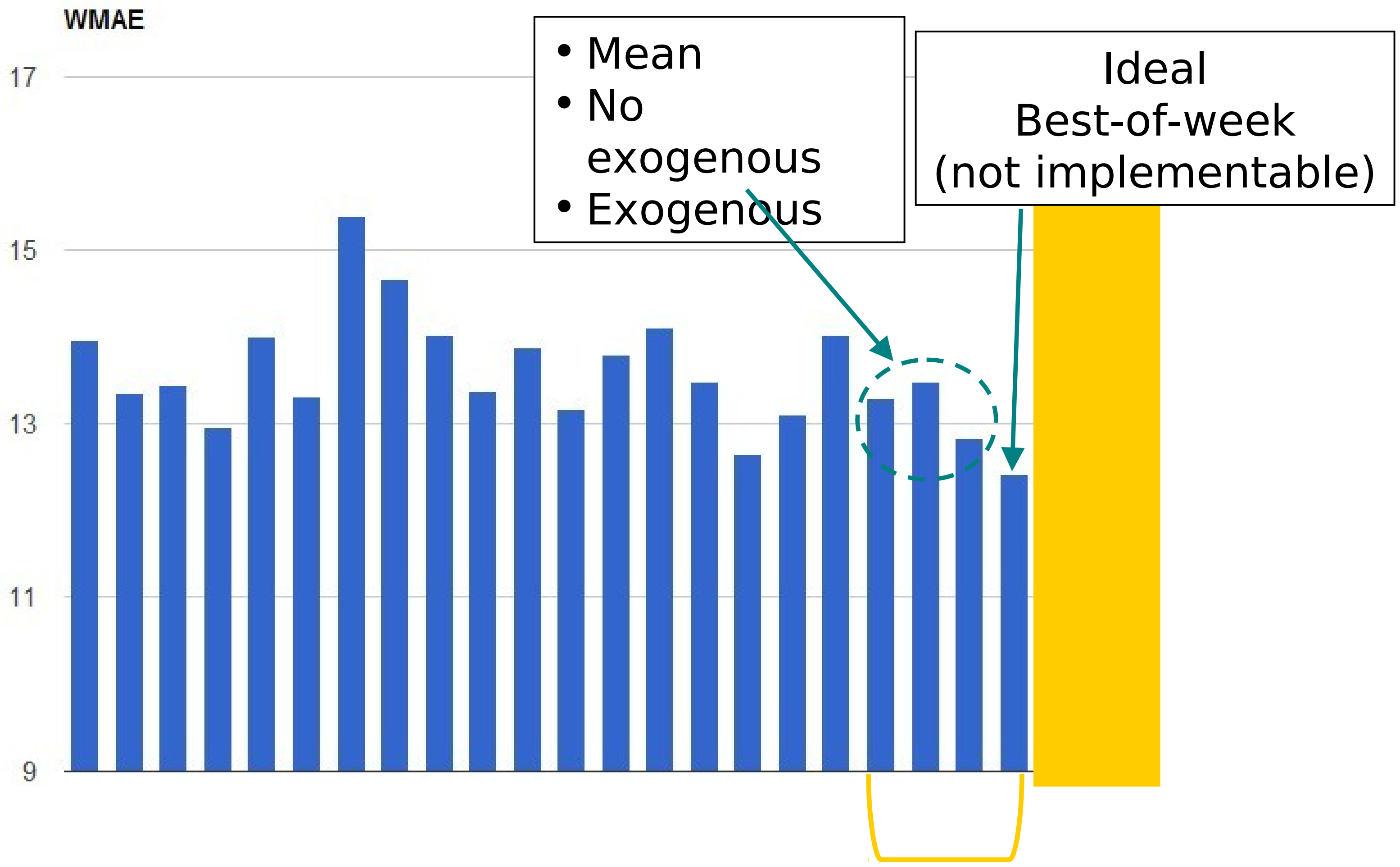
Results: Baseline (I)



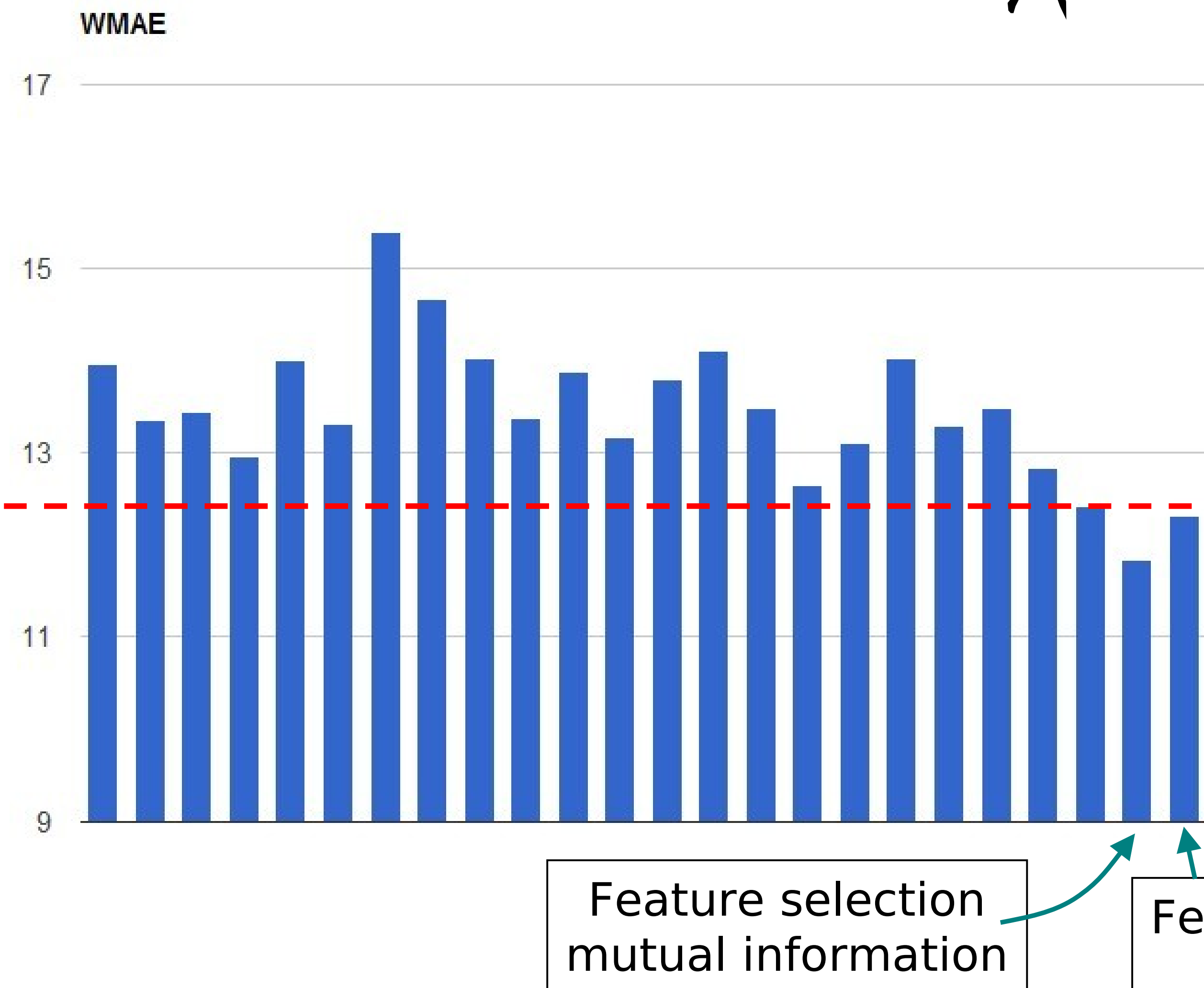
1. Basic autoregressive
2. …with spikes preprocessed
3. Threshold autoregressive
4. Mean-reverting jump diffusion
5. AR calibrated with Hsieh-Manski estimator
6. AR calibrated with ML estimator

Each with and without exogenous variable
(load forecast)

Results: Baseline (II)



Our Results



Mean	13.79
Mean pure-price	14.11
Mean load only	13.47
Ideal	12.64
Hybrid-mutual info	11.84
Hybrid-baseline	12.32

Our Results: more details

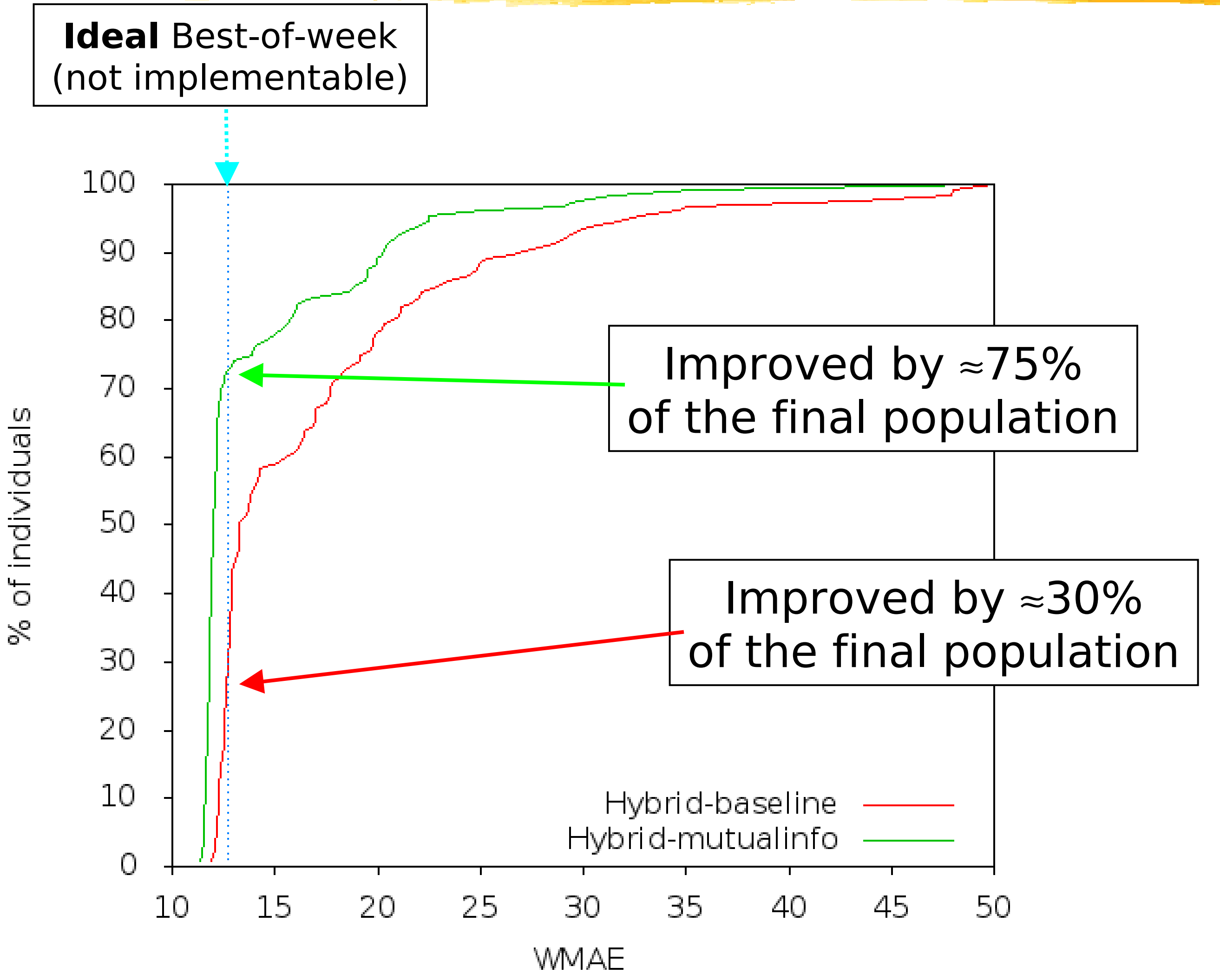
Method	WMAE (%)
GP-mutualInfo	20.70
GP-baseline	16.17
Classifier-base	16.03
Hybrid-mutualInfo	11.84
Hybrid-baseline	12.32
Mean	13.79
Mean pure-price only	14.11
Mean with load only	13.47
Ideal	12.64

- GP Estimator alone
  Different feature selection methods
- Outlier estimator alone
  (better than GP alone…)

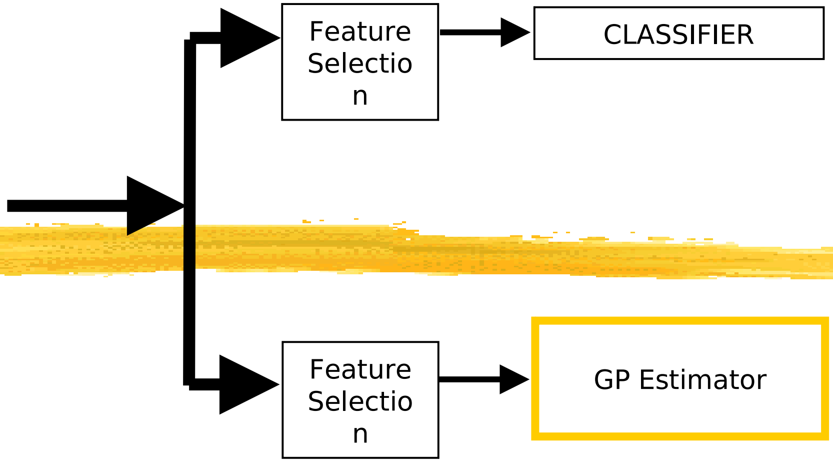
We need two separate estimators

Ensemble is much better than each of them

Not a single, lucky individual



Execution time



- Training of **GP Estimator**
  - 34 hours
  - 4 identical machines: quad-core Xeon with 2 GB RAM

- Training of **Classifier**
  - 1 hour
  - notebook: one-core, 2 GB RAM

- **Feature selection**
  - A few minutes
  - notebook: one-core, 2 GB RAM

Concluding remarks

- Solution to an important practical problem
  - Compares favorably with the state-of-the-art
- Application domain in which GP may compete with traditional approaches
- Simple yet effective way to cope with outliers
- In progress:
  - Other important datasets
  - Retuning policies for long periods