Optimization of Trading Strategies in Continuous Intraday Markets

Anthony Papavasiliou, Gilles Bertrand

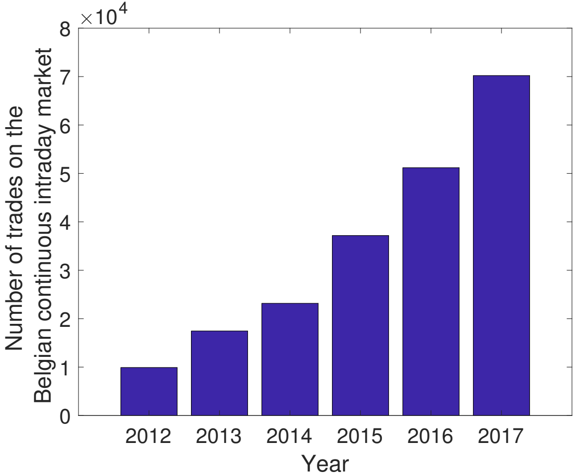
Center for Operations Research and Econometrics
Université catholique de Louvain

January 31, 2019

Outline

Outline

Motivation



$\times 10^4$
8
7
6
5
4
3
2
1
0
Number of trades on the
Belgian continuous intraday market
2012
2013
2014
2015
2016
2017
Year

Description of the Continuous Intraday Market

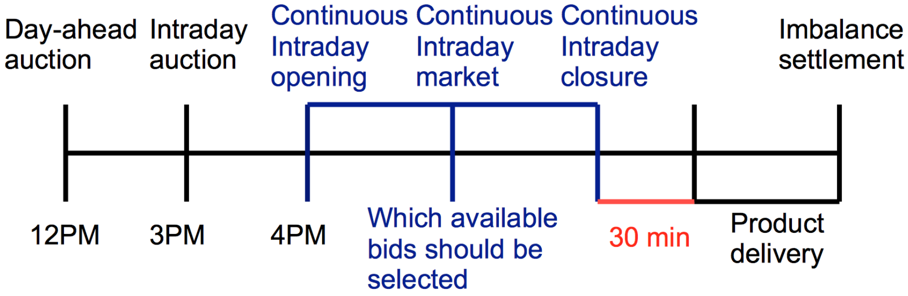


Figure: Short-term German electricity market

Format of Intraday Bids

	Hour	Quarter	Type	Price (€/MWh)	Quantity (MW)
Bid 1	1	h	s	28	10
Bid 2	1	h	b	25	5
Bid 3	1	q1	b	30	8
Bid 4	1	q2	b	25	2.5
Bid 5	1	q3	s	27	0.3
Bid 6	2	h	b	29	0.8
Bid 7	14	q4	s	32	3

- Bids arrive continuously in the intraday platform
- Bids are reserved on first-come-first-serve basis

Literature Review

Intraday price models

- [Kiesel 2015]: Econometric study of the parameters influencing the price evolution
- [Kiesel 2017]: modelling of order arrivals using Hawkes process

Trading by assuming a price model

- [Aid 2015]: solving the trading problem of a thermal generator using stochastic differential equations, assuming some model for the price evolution
- [Braun 2016]: solving the problem of optimizing pumped storage trading if we have access to a price curve for the coming hours

Trading without assuming a price model

- [Skajaa 2015]: heuristic method for covering the position of a wind farm based on imbalance price forecast

Our Goal

We are interested in a *model-free* approach that can handle

- continuous arrival of orders
- multi-stage uncertainty
- management of flexible (e.g. pumped hydro, storage) assets

Outline

Rolling Intrinsic

We consider the *rolling intrinsic* policy as a benchmark [Lohndorf, Wozabal, 2015]

- Applied for intraday trading with pumped hydro
- Receding horizon approach
- Myopic: accept any feasible trade that gives an *instantaneous* profit

Rolling Intrinsic Model

Accept any feasible trade that gives a positive profit

$$(P_t): \quad \max_{q^{s/b}_{i,t}, v_{t,d}} \sum_{d \in D} \sum_{i \in I_d} \left(P^b_i \cdot q^b_{i,t} - P^s_i \cdot q^s_{i,t} \right)$$

$$q^{s/b}_{i,t} \leq Q^{s/b}_{i,t} \qquad \forall i \in I_d, d \in D$$

$$v_{t,d} = v_{t-1,d} + \sum_{b \in D | b \leq d} \sum_{i \in I_b} \left(q^s_{i,t} - q^b_{i,t} \right) \qquad \forall d \in D$$

$$v_{t,d} \leq V \qquad \forall d \in D$$

$$v_{t,d} \geq 0 \qquad \forall d \in D$$

$$q^{s/b}_{i,t} \geq 0 \qquad \forall i \in I_d, d \in D$$

Perfect Foresight

Use perfect foresight model in order to:

- obtain upper bounds for any trading policy
- gain insights from the KKT conditions in order to design our policy

Perfect Foresight Model

The variables are not indexed by t anymore because perfect foresight setting is equivalent to having access to all bids at once

$$\max \sum_{d \in D} \sum_{i \in I_d} \left(P_i^b \cdot q_i^b - P_i^s \cdot q_i^s \right)$$

$$q_i^{s/b} \leq Q_i^{s/b} \qquad \forall i \in I_d, d \in D \quad (\mu_{i,d}^s)$$

$$v_d = v_{d-1} + \sum_{i \in I_d} \left(q_{i,d}^s - q_{i,d}^b \right) \qquad \forall d \in D \quad (\lambda_d)$$

$$v_d \leq V \qquad \forall d \in D \quad (\gamma_d)$$

$$v_d \geq 0 \qquad \forall d \in D \quad (\beta_d)$$

$$q_{i,d}^{s/b} \geq 0 \qquad \forall i \in I, d \in D \quad (\nu_{i,d}^s)$$

KKT Analysis of Perfect Foresight Policy

- If $\lambda_d < P_i^b$, we have $q_i^b = Q_i^b$
- If $\lambda_d > P_i^b$, we have $q_i^b = 0$
- If $\lambda_d < P_i^s$, we have $q_i^s = 0$
- If $\lambda_d > P_i^s$, we have $q_i^s = Q_i^s$

Interpretation of λ_d: **threshold** above which we sell and below which we buy

This suggests that a threshold policy could be a reasonable trading strategy

Outline

Definition of a Markov Decision Process

Markov decision process

A Markov decision process is a tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$, where

- $\mathcal{S}$ is a set of states
- $\mathcal{A}$ is a set of actions
- $\mathcal{P}^a_{s,s'} = \mathbb{P}[S_{t+1} = s' | S_t = s, A_t = a]$ is the probability to arrive in state s' if we follow action a in state s
- $\mathcal{R}$ is a reward function, $\mathcal{R}(s, a) = \mathbb{E}[R_{t+1} | S_t = s, A_t = a]$

Definition of a Markov Decision Process

Objective function

We optimize over a set of policies for the sum of reward if we follow a policy

$$\max_{\pi \in \Pi} \sum_{t=1}^{T} \mathbb{E}\left[R_t(S_t, A^\pi(S_t))\right]$$

Policy Function Approximation

Policy function approximation (PFA)

The idea in PFA is to approximate directly the policy

$$\pi(a|s;\theta) = \mathbb{P}[A_t = a|S_t = s;\theta]$$

Illustration of Threshold Policies: Purely Financial Problem

- We have to decide whether to accept a bid at the intraday price p^{ID}
- We know the intraday price p^{ID}, but the real-time price p^{RT} is uncertain

MDP Formulation of Purely Financial Problem

Purely financial problem as an **MDP**

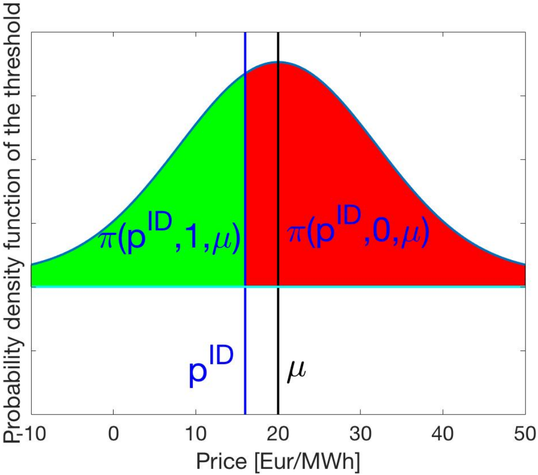
- $S = \{p^{\text{ID}}\}$, the intraday price
- $A = \{a\}$, a binary variable whose value is equal to 1 if we accept the bid, or 0 if we reject it
- $\mathcal{R}(s, a) = \mathbb{E}[p^{\text{ID}} - p^{\text{RT}} | p^{\text{ID}}] \cdot a$

Policy function approximation for the purely financial problem

We use a stochastic threshold policy with parameters $\theta = (\mu, \sigma)$

$$\pi(p^{\text{ID}}, 0; \theta) = 1 - F_\theta(p^{\text{ID}})$$
$$\pi(p^{\text{ID}}, 1; \theta) = F_\theta(p^{\text{ID}})$$

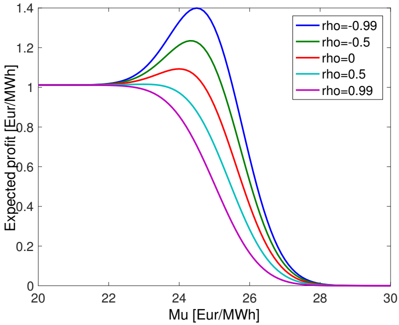
Graphical Illustration of a Stochastic Threshold



Payoff for Bivariate Normal Distribution

Payoff as a function of $\theta = (\mu, 0^+)$:

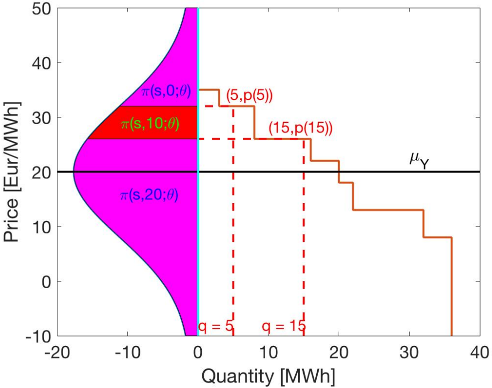
$$J(\mu) = \mathbb{E}\left[p^{\text{ID}} - p^{\text{RT}} | p^{\text{ID}} \geq \mu\right] \cdot (1 - F_{p^{\text{ID}}}(\mu))$$



Assuming that $(p^{\text{ID}}, p^{\text{RT}})$ are *bivariate normal*, $J(\mu)$ can be computed analytically and is a **non-concave** function of μ

Outline

Graphical Representation of Threshold Policy for Pumped Hydro Problem



Graphical Representation of Threshold Policy

We use a *threshold policy*, which is a distribution over actions:

- The bell curve indicates the probability density function of the sell threshold
- The two purple segments and the red segment of the bell curve indicate the probability of each of the three actions:
  - Sell 0 MWh
  - Sell 10 MWh
  - Sell 20 MWh
- The green decreasing function corresponds to the buy bids that are available in the order book for a given trading hour

We are interested in finding an **optimal** threshold

REINFORCE Algorithm

Algorithm

REINFORCE algorithm for finite horizon:

- Initialize θ_0
- for each episode $\{s_1, a_1, r_2, \cdots, s_{T-1}, a_{T-1}, r_T\} \sim \pi_\theta$

  for t $= 1 : T - 1$ do

  $$\theta_{k+1} = \theta_k + \alpha \nabla_\theta \log(\pi(s_t, a_t; \theta)) g_t$$

  end for

  end for

Remark

g_t is the profit from t to the end T of the episode

$\nabla_\theta \log$

These gradients can be expressed in closed form

Generalization of the Threshold Policy

Let $f\colon R^n \to R$ be a differentiable function s.t. $\theta = f(\alpha)$. We can compute the derivative with respect to α by using the chain rule:

$$\begin{aligned}\frac{\partial\pi(s;\theta)}{\partial\alpha} &= \frac{\partial\pi(s;\theta)}{\partial\theta}\frac{\partial\theta}{\partial\alpha} \\ &= \frac{\partial\pi(s;\theta)}{\partial\theta}\frac{\partial f}{\alpha}\end{aligned}$$

This allows us to influence the threshold by observing relevant factors

Expected Behaviour of a Threshold Policy

1. Ensure that the stored volume respects reservoir limits
2. Adapt with respect to the intraday auction price
3. Adapt with respect to the delivery time
4. Adapt with respect to the evolution of intraday prices
5. Adapt with respect to the remaining time

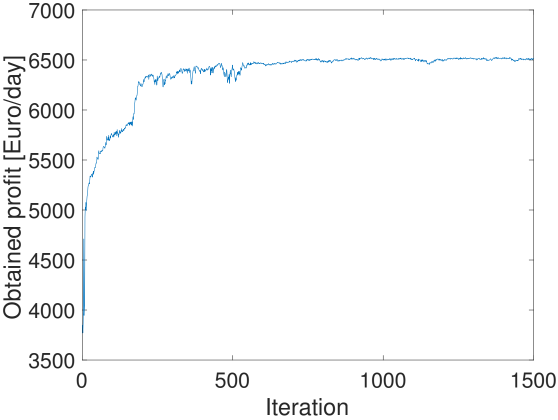
Outline

Case Study

- Data source: 2 years of data of the German CIM, procured from EPEX
- Training data: 200 days of 2015
- Testing data: 165 last days of 2015

Training the Policy Function

This graph shows the evolution of the *profit* with respect to the *iteration*. An *iteration* corresponds to 5 repetitions of our 200 days of learning.



Competing Policies

We have compared the results of four different methods:

- **Rolling intrinsic 4pm:** rolling intrinsic method launched at 4pm
- **Rolling intrinsic 11pm:** rolling intrinsic method launched at 11pm
- **Threshold:** our proposed threshold policy
- **Perfect foresight**

Comparison of Policies

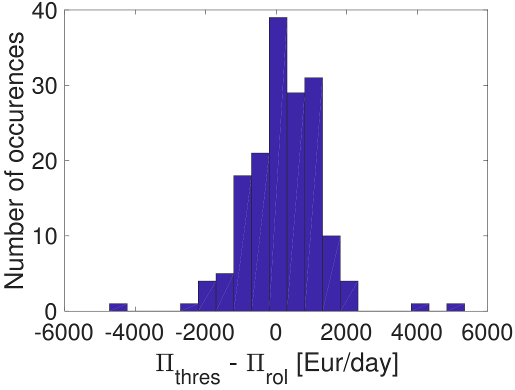
Method	Profit mean [€/day]	Profit standard deviation [€/day]
Rolling intrinsic 4pm	4042	1968
Rolling intrinsic 11pm	**4871**	2034
Threshold	**5076**	2484
Perfect foresight	10321	4416

In the next slides, we will compare the two best performing policies:

- rolling intrinsic starting at 11pm
- threshold policy

Distribution of Profits

- One *occurrence* corresponds to one day of trading
- The profit is accumulated gradually and is not coming from one spike



Significance of Profit Difference

We conduct a *p*-value test with the two following hypotheses

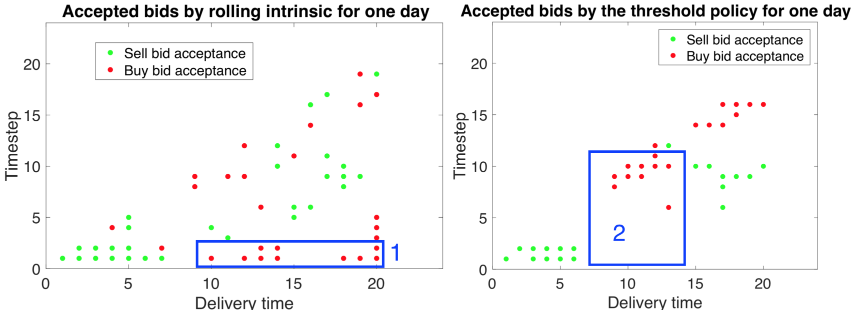
- **Null hypothesis**: $\mathbb{E}[\Pi_{\text{thres}}] = \mathbb{E}[\Pi_{\text{rol}}]$
- **Alternative hypothesis**: $\mathbb{E}[\Pi_{\text{thres}}] > \mathbb{E}[\Pi_{\text{rol}}]$

We find that the probability of obtaining the observed profit differences with $\mathbb{E}[\Pi_{\text{thres}}] = \mathbb{E}[\Pi_{\text{rol}}]$ is equal to 0.7%

Different Attitude towards Risk

There is a trade-off between

1. arbitraging against earlier bids with less interesting prices (rolling intrinsic, risk-free)
2. waiting for more interesting prices later in the day (threshold policy, more risky)



Conclusions and Perspectives

- Observations
  - The profit of rolling intrinsic varies significantly with the time that trading commences
  - Our method outperforms rolling intrinsic with statistical significance
- Future research
  - We are trading at hourly frequency, we would like to solve the problem at *higher trading frequency*
  - Step size analysis
  - Accelerated learning through parallel computing

Thank you

Contact:

Anthony Papavasiliou, anthony.papavasiliou@uclouvain.be
http://perso.uclouvain.be/anthony.papavasiliou/

Gilles Bertrand, gilles.bertrand@uclouvain.be
https://sites.google.com/site/gillesbertrandresearch/home