Institute for Operations Research
and Computational Finance

University of St.Gallen

SMSA 2015 Statistics in Energy

# Optimization of hydro storage systems and indifference pricing of power contracts

**Florentina Paraschiv (University of St. Gallen, ior/cf)**

**Raimund M. Kovacevic (Vienna University of Technology)**

**Michael Schürle (University of St. Gallen, ior/cf)**

## Outlook

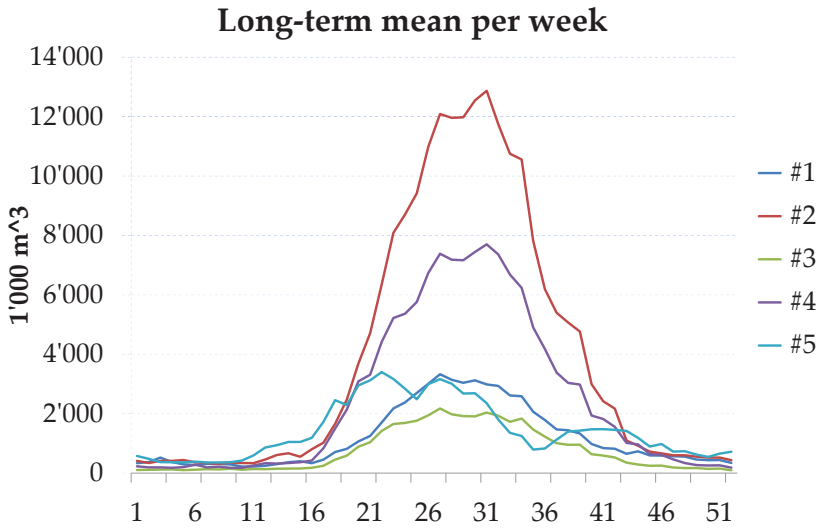
- Producer owns **pumped-storage hydropower plants** (PSHPs). These consist of large (seasonal) connected reservoirs
- We must decide on a production schedule
- We aim at a **mid-term planning model for hydropower production** based on **multistage stochastic programming**
  - Stochastic inflows; stochastic electricity prices
  - Produced electricity is sold at the spot market
- Valuation of non-standard power contracts with indifference pricing (work in progress)
  - Stochastic load profiles
  - Indifference price: entering the contracts with individual consumers or selling only at the spot market

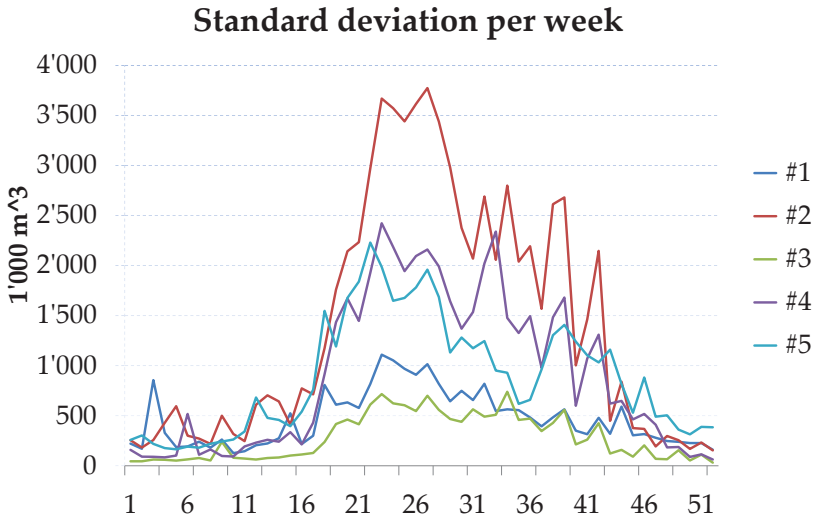
## Inflow data

- Daily inflows are given for **a system with five connected reservoirs** (sample period: 26 years)
- Data are aggregated to weekly inflows
- The inflows show a characteristic seasonal pattern:
  - Winters are dry
  - Inflows are higher in the middle of the year
- Thus, **a seasonal adjustment will be performed**

# Seasonality

- Let $f_i(\tau)$ be the inflow in reservoir $i = 1, \ldots, 5$ at time $\tau = 1, \ldots, T$
- Let $\bar{f}_i(w')$ long-term mean of the inflows, $w' = 1, ..., 52$ and $\hat{\sigma}_i(w')$ the standard deviations of the inflows after seasonal adjustment for all reservoirs over one year
- Allocate each time point to one week $w'$: $\mathcal{T}(w') := \{\tau = 1, ..., T \mid w(\tau) = w'\}$
  - Long-term seasonal adjustment: $f_i^{adj}(\tau) = f_i(\tau) - \bar{f}_i(w(\tau))$
  - Correct for time-varying volatility: $f_i^{std}(\tau) = \frac{f_i^{adj}(\tau)}{\hat{\sigma}_i(w(\tau))}$



## Principal component analysis

- The principal component analysis is performed on the stochastic component of the inflows
- According to the eigenvalue criterion, only the first factor should be considered
- Then 75% of the variation can be explained; with 2(3) factors we explain about 86(93) percent of the variation

Table 1: Principal component analysis

| Component | Total | % of Variance | Cumulative % |
|---|---|---|---|
| 1 | 3.744 | 74.9 | 74.9 |
| 2 | 0.573 | 11.5 | 86.3 |
| 3 | 0.371 | 7.4 | 93.8 |
| 4 | 0.279 | 5.6 | 99.3 |
| 5 | 0.033 | 0.7 | 100.0 |

## Spot model for electricity prices

The formulation of the regime-switching model reads:

$$X_t = \begin{cases} \ell_t^L - \xi_t^-, & \text{if the system is in the lower spike regime,} \\ s_t \cdot \exp(r_t), & \text{if the system is in the base regime or} \\ \ell_t^U + \xi_t^+, & \text{if the system is in the upper spike regime.} \end{cases} \tag{1}$$

$$dr_t = -\kappa_\alpha r_t dt + \sigma_\alpha dZ_t \tag{2}$$

$$l_t^L = s_t / \exp(a_\gamma^L), \qquad l_t^U = s_t \cdot \exp(a_\gamma^U). \tag{3}$$

$$\xi_t^- \sim \text{Wei}(\lambda_\beta^-, k_\beta^-), \qquad \xi_t^+ \sim \text{Wei}(\lambda_\beta^+, k_\beta^+)$$

| | |
|---|---|
| $X_t$ | spot price in hour $t$ |
| $s_t$ | seasonality shape (deterministic component) |
| $r_t$ | stochastic component of the spot price in the base regime |
| $\xi_t^+$ | upward spike |
| $\xi_t^-$ | downward spike |
| $\ell_t^U$ | upper limit of the base regime in hour $t$ |
| $\ell_t^L$ | lower limit of the base regime in hour $t$ |

## Transition probabilities and results

- The probabilities of remaining in the current regime or switching to another one from one hour to the next are modeled by a (time-dependent) transition matrix:

$$\Pi_\delta = \begin{pmatrix} 1-\pi_\delta^{LB} & \pi_\delta^{LB} & 0 \\ \pi_\delta^{BL} & 1-\pi_\delta^{BL}-\pi_\delta^{BU} & \pi_\delta^{BU} \\ 0 & \pi_\delta^{UB} & 1-\pi_\delta^{UB} \end{pmatrix}. \tag{4}$$

  where $\delta := \delta(t)$ is a function that maps $t$ to a set of indices (seasonality)

- Estimation procedure: **two-step maximum likelihood**

## Estimation results

Estimation sample: Hourly EEX spot prices from 1 Jan 2009 to 31 Dec 2013

| Param. | $a_\gamma^L$ | $a_\gamma^U$ | $\kappa_\alpha$ | $\sigma_\alpha$ | $\lambda_\beta^-$ | $k_\beta^-$ | $\lambda_\beta^+$ | $k_\beta^+$ | $E(\xi_t^-)$ | $E(\xi_t^+)$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Overall | 0.697 | 0.875 | 2.848 | 0.566 | 0.127 | 0.737 | 0.09 | 0.798 | 9.493 | 12.681 |
| | | | (0.05) | (0.001) | (0.005) | (0.01) | (0.018) | (0.101) | | |
| Summer | 0.649 | 0.894 | 3.045 | 0.555 | 0.184 | 0.936 | 0.105 | 1.326 | 5.591 | 8.759 |
| | | | (0.075) | (0.002) | (0.008) | (0.015) | (0.042) | (0.899) | | |
| Winter | 0.89 | 1.054 | 2.538 | 0.606 | 0.101 | 0.646 | 0.065 | 1.134 | 13.657 | 14.771 |
| | | | (0.057) | (0.002) | (0.006) | (0.015) | (0.027) | (0.542) | | |

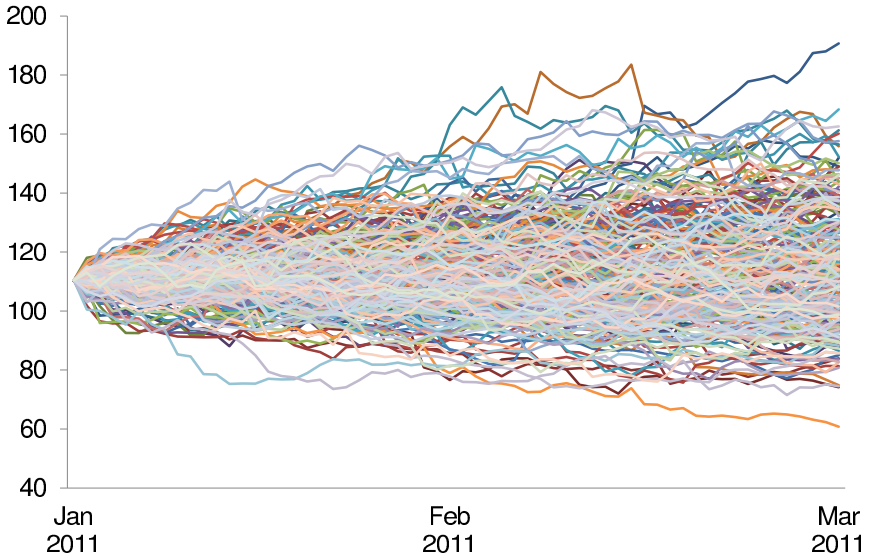
Table 2: Estimation results#1 (standard errors in parenthesis). $E(\xi_t^{+/-})$ shows the expected value of upwards/downwards spike magnitudes

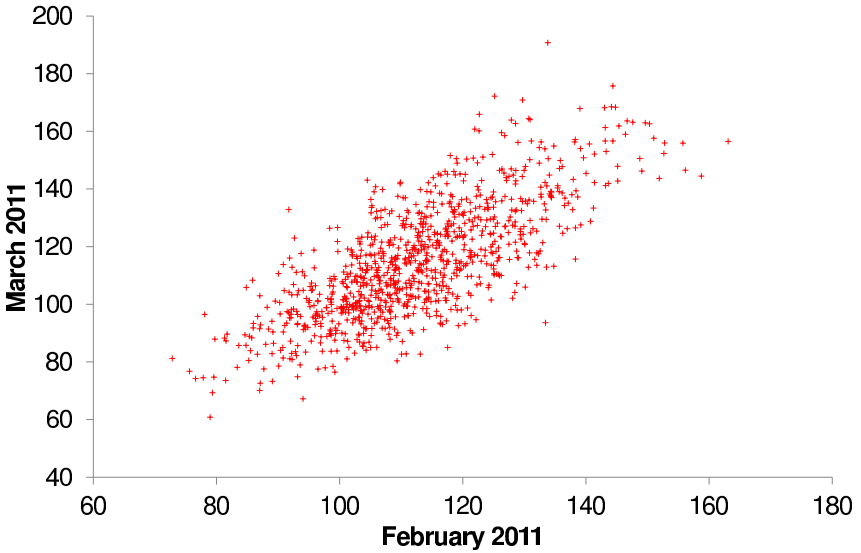
| Param. | $\pi_\delta^{BL/LB}$ | $\pi_\delta^{BU/UB}$ | $\#base$ | $\#lower$ | $\#upper$ | loglike |
|---|---|---|---|---|---|---|
| Overall | 1.053% | 0.078% | 41937 | 1837 | 45 | 27156 |
| | 24.061% | 73.333% | | | | |
| Summer Base | 1.631% | 0.015% | 21130 | 826 | 4 | 27338 |
| | 28.800% | 66.666% | | | | |
| Summer Peak | 0.193% | 0.013% | | | | |
| | 22.368% | 100% | | | | |
| Winter Base | 1.309% | 0.007% | 21064 | 788 | 7 | |
| | 24.386% | 100% | | | | |
| Winter Peak | 0.078% | 0.064% | | | | |
| | 12.631% | 80% | | | | |

Table 3: Estimation results#2 (standard errors in parenthesis)

# Generating scenarios

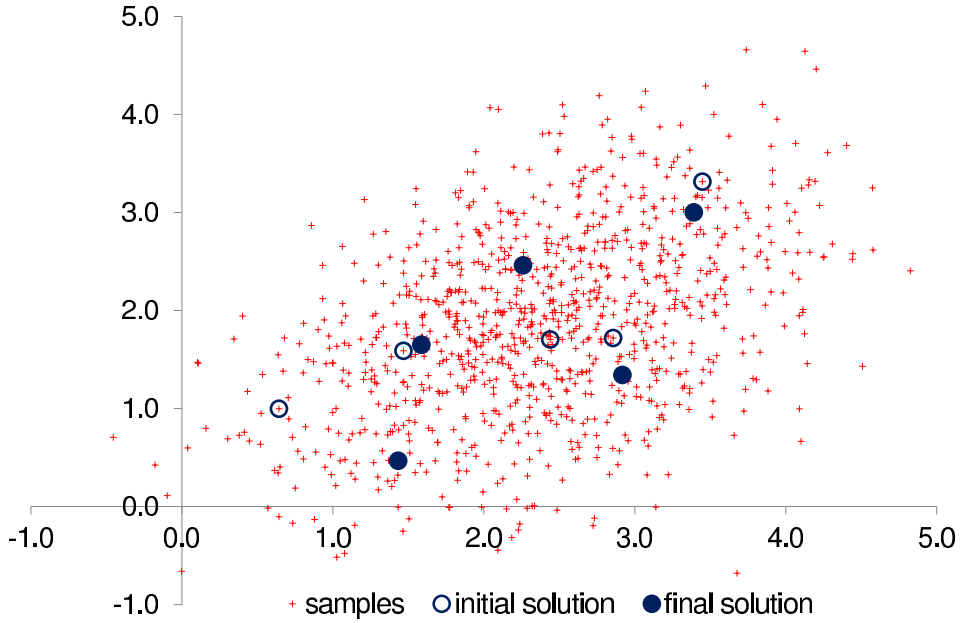
- We generate 10'000 scenarios in weekly steps for:
  - The stochastic component of the spot price process for the base regime
  - Two principal components of the inflow model
- Find a tree that is:
  - Small enough as a numerically tractable approximation
  - Large enough to capture important features of the problem
- The scenario tree should be as close as possible to the observed stochastic process



# Excursus: Minimizing the distance for distributions

- Consider a random variable $G$ that is either continuous or discrete (with a very large number of mass points)
- We want to approximate $G$ by a simpler random variable $\tilde{G}$
- The distance between $G$ and $\tilde{G}$ can be measured by the Wasserstein distance $d(G, \tilde{G})$
- It is known that the Wasserstein distance is related to a transportation problem
- The problem of minimizing the distance $d(G, \tilde{G})$ is solved by assigning data points to a few "clusters" (which represent the approximate distribution $\tilde{G}$)

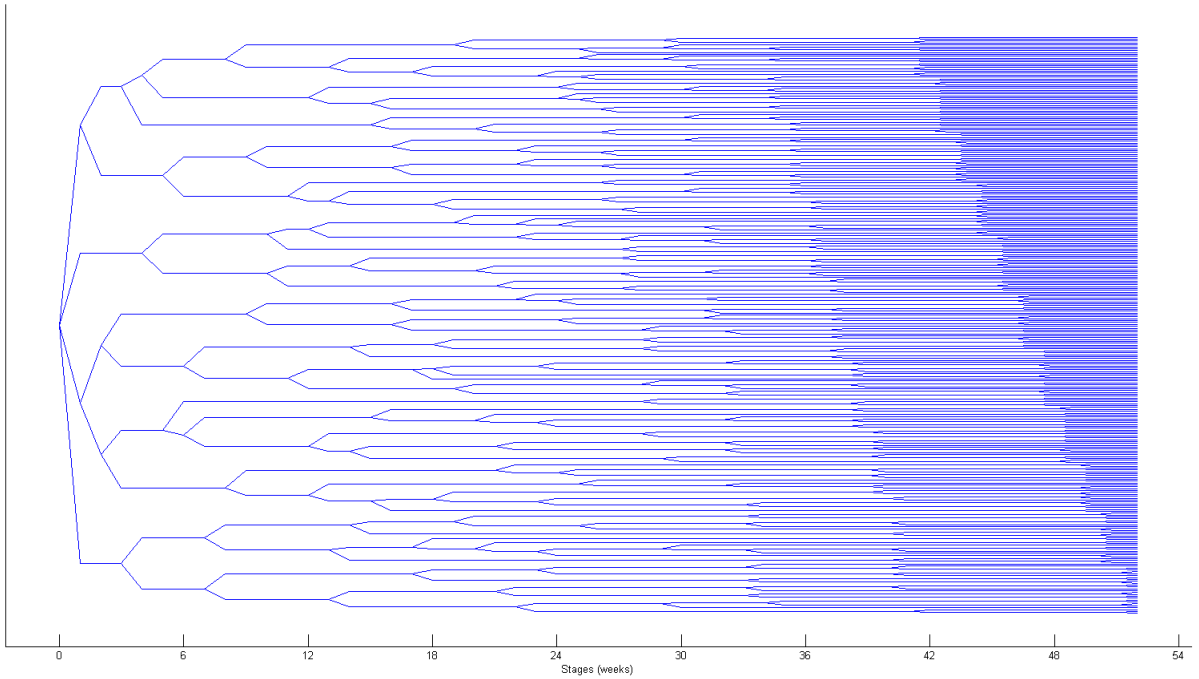



$$\mu = \begin{pmatrix} 2.37 \\ 1.83 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} 0.71 & 0.33 \\ 0.33 & 0.91 \end{pmatrix}$$

# Extension to nested distances

- Pflug/Pichler (2012) introduced and analyzed a generalization of the well known Wasserstein distance
- Kovacevic/Pichler (2012) propose an algorithm for improving the distance between the trees
- This *nested distance* takes the information from the filtration into account (rather than comparing only scenario paths)
- Based on this concept, the tree resulting from the first step is further improved by adjusting the probabilities and values



## Time aggregation: price level approach

- On the one hand, the valuation of non-standard power contracts requires an hourly time resolution
- On the other hand, the computational effort would explode since scenario trees grow exponentially with the number of stages
- Thus, an **aggregation technique** is required!
- We apply the price level approach (a.k.a. occupation times)

## Optimization model prerequisites

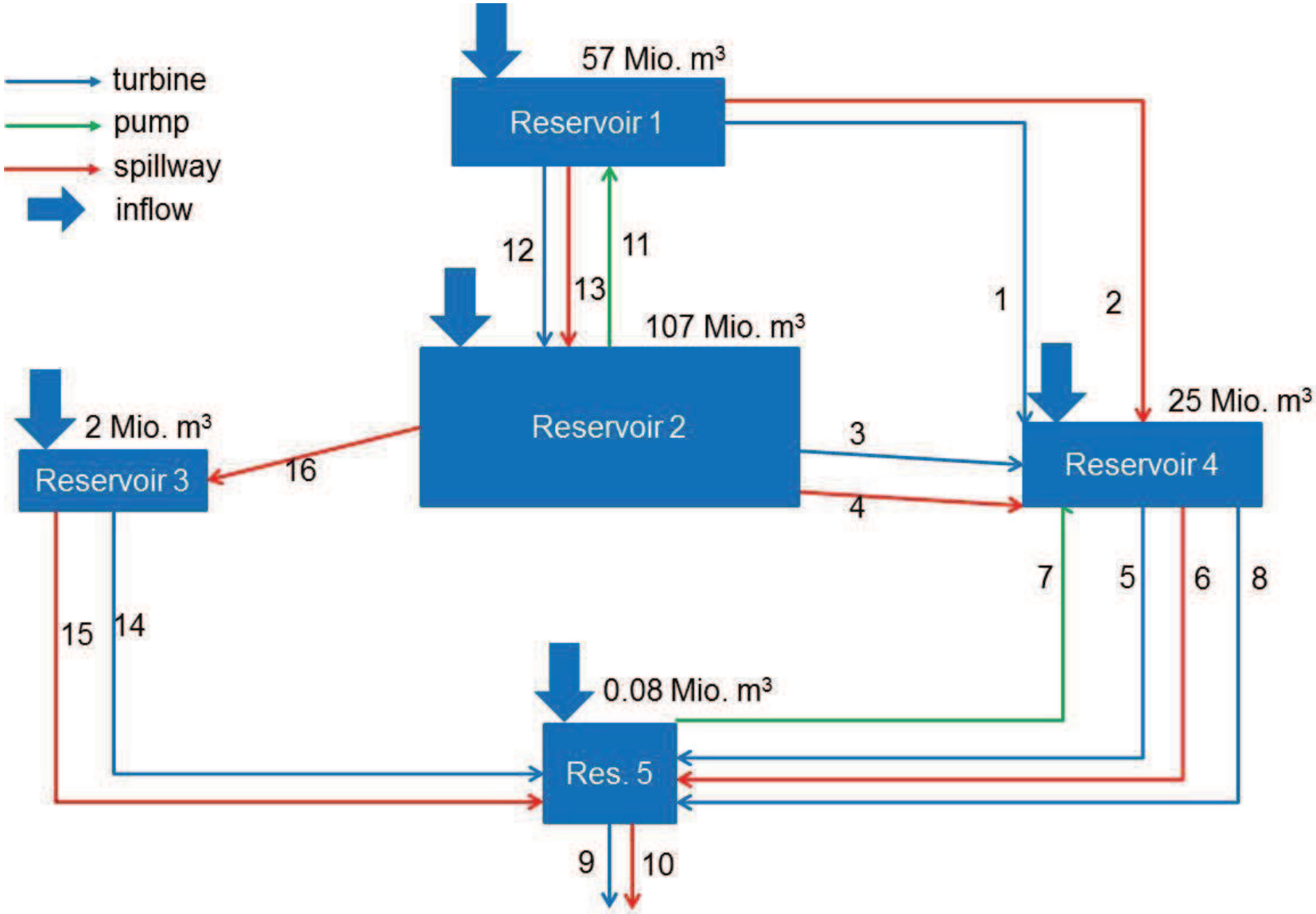
- Notation:

| | |
|---|---|
| $\mathcal{I} = \{1, \ldots, I\}$ | set of reservoirs |
| $\mathcal{J} = \{1, \ldots, J\}$ | set of connections between reservoirs |
| $\mathcal{T} \subset \mathcal{J}$ | set of turbines |
| $\mathcal{P} \subset J$ | set of pumps |
| $\mathcal{W} \subset J$ | set of spillways |
| $\mathcal{L} = \{1, ...L\}$ | price level index set |

- The system topology is described by an incidence matrix A such that

$$A_{ji} = \begin{cases} 1 & \text{if water flows into reservoir } i \text{ over connection } j, \\ -1 & \text{if water flows out of reservoir } i \text{ over connection } j, \\ 0 & \text{if connection } j \text{ is not related to reservoir } i. \end{cases}$$

# Reservoirs layout



## System parameters

- **System parameters (deterministic):**

| | |
|---|---|
| $\bar{s}_i$ | maximum storage volume (minimum is zero) in $m^3$ |
| $g_i$ | minimum storage level (end of planning horizon in $m^3$ |
| $\underline{q}_j, \bar{q}_j$ | minimum/maximum water flow in $m^3/h$ |
| $\eta_j$ | production efficiency in $MWh/m^3$ |
| $\underline{p}_j, \bar{p}_j$ | lower/upper bounds on produced power in $MWh$ |

- **Remarks:**
  - Turbines have positive efficiency (conversion from water to electricity)
  - Pumps have negative efficiency, negative minimum power and maximum power of zero
  - Spillways have zero efficiency and nonnegative flow capacity

- **System parameters (stochastic):**

| | |
|---|---|
| $S_t$ | electricity spot price at time $t$ in $EUR/MWh$ |
| $F_{it}$ | inflow in reservoir $i$ (precipitation, snowmelt etc.) in $m^3$ |

## Tree notation

- The (joint) stochastic process of the uncertain factors is represented by a finite tree with node set $\mathcal{N} = \{0, 1, ..., N\}$
- $\mathcal{N}_t$ is the set of nodes at level $t = 0, ..., T$
- The tree structure represents the filtration of the process and can be defined by stating the unique predecessor $n^-$ for each node $n$
- In the tree-based model with price levels notation changes:

  $S_{n\ell}$ spot price at level $\ell$ in node $n$

  $F_{in}$ inflow in reservoir $i$ in node $n$ in $m^3$

- Further definitions:

  $\Delta_n$ length of planning period beginning in node $n$

  $\pi_n$ probability of node $n$

  $\pi_{n\ell}$ probability of price level $\ell$ in node $n$ (index set: $\mathcal{L}$)

## Optimization model

- For each price level $\ell$ at each node $n$ the producer decides how much water $q_{jn\ell}$ flows over connection $j$ (in $m^3/h$)
- This corresponds to the produced power $p_{jn\ell}$ (in $MW$) and energy $e_{jn\ell}$ (in $MWh$) over the period:

$$p_{jn\ell} = \eta_j q_{jn\ell}, \tag{5}$$

$$e_{jn\ell} = p_{jn\ell}\pi_{n\ell}\Delta_n \tag{6}$$

- Technical constraints for produced (consumed) power:

$$\underline{p}_j \leq p_{jn\ell} \leq \bar{p_j} \tag{7}$$

- The stored water $s_{in}$ in each reservoir at the beginning of the period is restricted by the reservoir capacity:

$$0 \leq s_{in} \leq \bar{s}_i \tag{8}$$

- It starts from an initial value $s_{i0}$ and develops over time as:

$$s_{in} = s_{in-} + \sum_{j\in\mathcal{J}} A_{ji} \sum_{\ell\in\mathcal{L}} \pi_{n\ell} q_{jn-\ell}\Delta_n + F_{in} \tag{9}$$

- At time $T$, the stored water volume has to meet the minimum level:

$$s_{in} \geq \bar{g}_i \quad (n \in \mathcal{N}_T) \tag{10}$$

## Optimization model

- We start with an initial cash amount $c_0$; for later time points $n > 0$ cash evolves according to the budget equation:

$$c_n = R_{n^-} c_{n^-} + \sum_{j \in \mathcal{T} \cup \mathcal{P}} \sum_{\ell \in \mathcal{L}} e_{jn^-\ell} S_{n^-\ell} \tag{11}$$

  where $R_n$ is the interest rate in node $n$ over the period $\Delta_n$ (equal for lending and borrowing)

- **The objective is defined as a mix of expectation and average value at risk** in view of the available cash at the end of the planning horizon in T:

$$\max \lambda E(\text{“}cash\ in\ T\text{”}) + (1 - \lambda) AVaR(\text{“}cash\ in\ T\text{”}) \tag{12}$$

## Optimization model overview

- The complete model in the tree-based notation is:

$$\max \lambda \sum_{n \in \mathcal{N}_T} \pi_n c_n + (1-\lambda) z \tag{13}$$

$$\begin{aligned}
& s.t. && z \leq y_0 - \frac{1}{\alpha} \sum_{n \in \mathcal{N}_T} \pi_n y_n \\
& n \in \mathcal{N}_T : && c_n - y_0 + y_n \geq 0 \\
& n \in \mathcal{N}_T : && y_n \geq 0 \\
& j \in \mathcal{J}; \ell \in \mathcal{L}; n \in \mathcal{N} : && p_{jn\ell} = \eta q_{jn\ell} \\
& j \in \mathcal{J}; \ell \in \mathcal{L}; n \in \mathcal{N} : && e_{jn\ell} = p_{jn\ell} \pi_{n\ell} \Delta_n \\
& j \in \mathcal{J}; \ell \in \mathcal{L}; n \in \mathcal{N} : && \underline{p}_j \leq p_{jn\ell} \leq \bar{p_j} \\
& i \in \mathcal{I}; n \in \mathcal{N} \setminus \{0\} : && s_{in} = s_{in^-} + \sum_{j \in \mathcal{J}} A_{ji} \sum_{\ell \in \mathcal{L}} \pi_{n\ell} q_{jn^-\ell} \Delta_n + F_{in} \\
& i \in \mathcal{I}; n \in \mathcal{N} \setminus \{0\} : && 0 \leq s_{in} \leq \bar{s}_i \\
& i \in \mathcal{I}; n \in \mathcal{N}_T : && s_{in} \geq \bar{g}_i \\
& n \in \mathcal{N} \setminus \{0\} : && c_n = R_{n^-} c_{n^-} + \sum_{j \in \mathcal{T} \cup \mathcal{P}} \sum_{\ell \in \mathcal{L}} e_{jn^-\ell} S_{n^-\ell}
\end{aligned}$$

## Indifference pricing

- Consider $M$ delivery contracts $m \in \mathcal{M} = \{1, ..., M\}$ of individual consumers
  - Different demand levels $D_{mn\ell}$ for different price levels
  - Electricity is delivered at fixed prices $K_m$ $(EUR/MWh)$
- **Indifference pricing:** Determine the smallest possible price such that the producer meets the obligations from the contracts; new mix of expectation and AVaR $\gamma$ should not be worse than without the contract

$$\min_{K,q,p,e,c,s,h} \sum_{m \in \mathcal{M}} K_m$$

s.t. (4) – (10)

AVaR-constraints

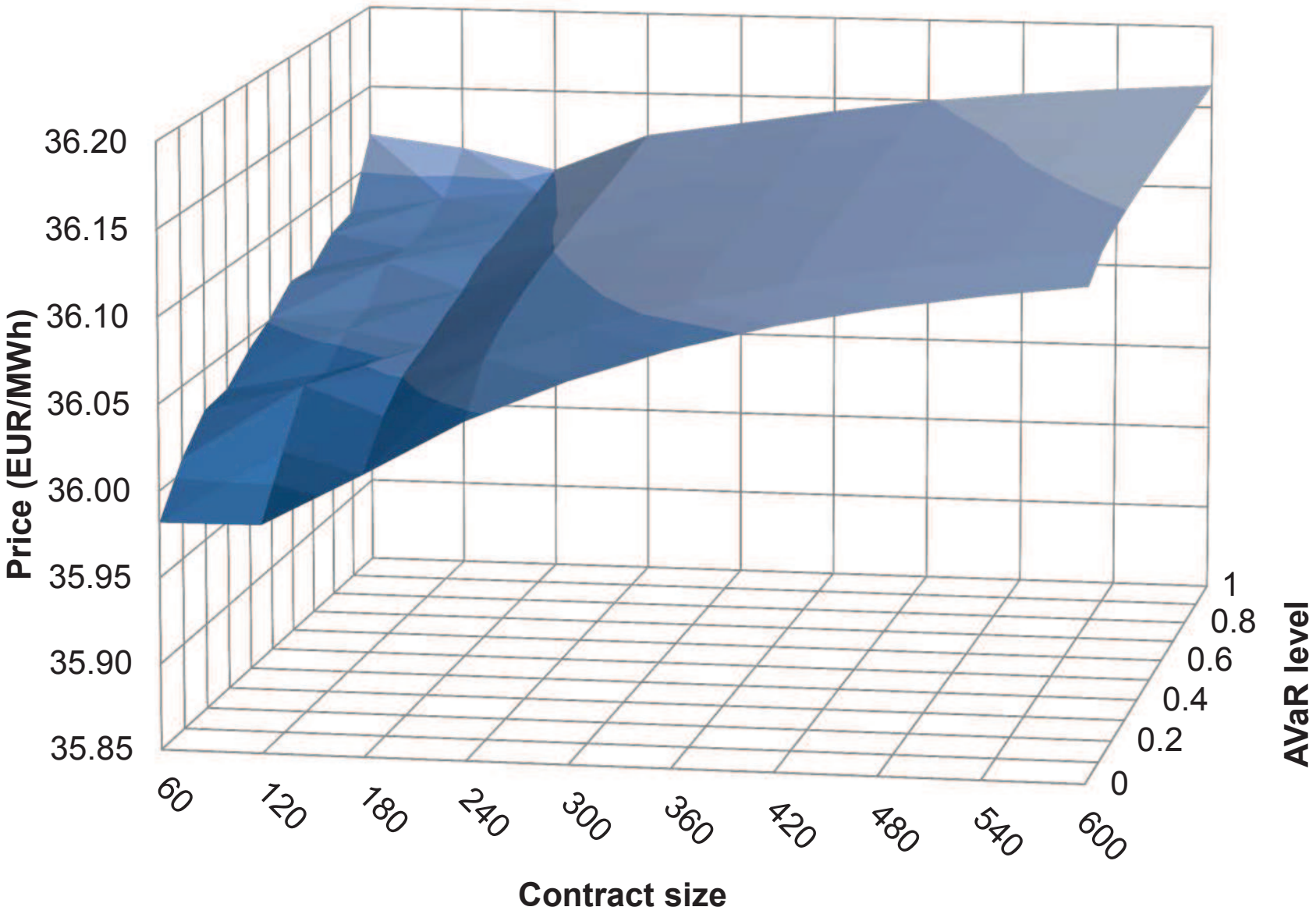
$$\gamma \geq \gamma*$$

$$\sum_{j \in \mathcal{T} \cup \mathcal{P}} e_{jn\ell} = h_{n\ell} + \Delta_n \pi_{n\ell} \sum_{m \in \mathcal{M}} D_{mn\ell}$$

$$c_n = R_{n-} c_{n-} + \sum_{\ell \in \mathcal{L}} S_{n\ell} h_{n\ell} + \Delta_n \sum_{\ell \in \mathcal{L}} \pi_{n\ell} \sum_{m \in \mathcal{M}} K_m D_{mn\ell}$$

$$s_{in} \geq s^*_{in} \quad (n \in \mathcal{N}_T)$$

# Indifference pricing for a constant load, one contract



## Summary

- Optimization of PSHP operations by multistage stochastic programming with aggregation based on price levels
- Model aims at large reservoirs (seasonal storages)
  - Regime-switching model for electricity spot prices
  - Factor model for inflows, taking into account seasonality
- Valuation of power contracts with indifference pricing
- Open:
  - Modelling stochastic load profiles
  - Extension to several contracts (portfolio effects)
  - Comparison with other approaches
  - Scenario generation based on "nested distance" concept
- Further directions of research:
  - Alternative aggregation techniques
  - Inclusion of intraday market