Deep Learning for Pump Scheduling

Sophie Demassey, Valentina Sessa, Amirhossein Tavakoli

March 5, 2024

Centre de Mathématiques Appliquées, Mines Paris PSL

the pump scheduling problem

plan the operation of a drinking water distribution network
to minimize the electricity bill of pumping

the pump scheduling problem

plan the operation of a drinking water distribution network
to minimize the electricity bill of pumping

energy efficiency

- **growing** demand in water: up to 50% in the world by 2050
- **energy-intensive**: 4% in the US electricity consumption
- response to a dynamic incentive electricity tariff with **load shifting**

the pump scheduling problem

plan the operation of a drinking water distribution network
to minimize the electricity bill of pumping

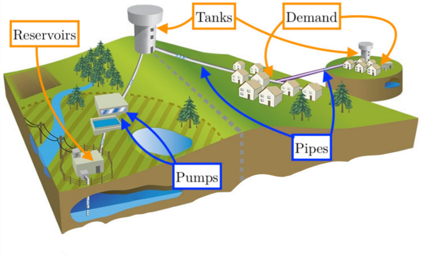
energy efficiency

- **growing** demand in water: up to 50% in the world by 2050
- **energy-intensive**: 4% in the US electricity consumption
- response to a dynamic incentive electricity tariff with **load shifting**

hard optimization

- **discrete control** (on/off) over a discretized time horizon
- **nonlinear** behavior (pressure/flow relation)
- time-coupling constraints: **storage state** (elevated water tanks)

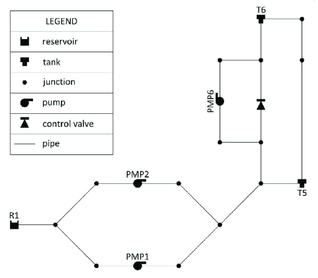
A DRINKING WATER DISTRIBUTION NETWORK



a directed graph G

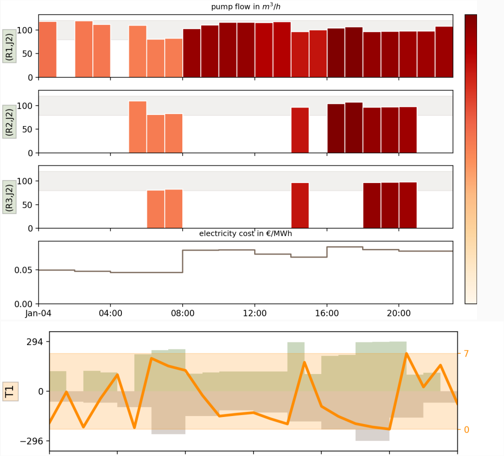
arcs A: pipes, pumps, valves

nodes J: users, tanks, sources



Pump scheduling problem

solved on a daily basis: plan the operation of the pumps over time $t \in \{1, ..., T-1\}$, to satisfy the water demand D_t, at minimum cost given tariff C_t



pump control/operation

on/off switch $x_{ta} \in \{0, 1\}$

flow $q_{ta} \in \mathbb{R}$

electricity tariff

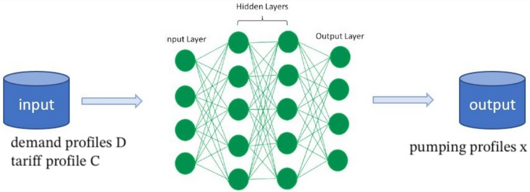
$C_t \in \mathbb{R}_+$

tank state/level

$H_{tj} \in [\underline{H}_j, \overline{H}_j]$

BAU optimization with Machine Learning

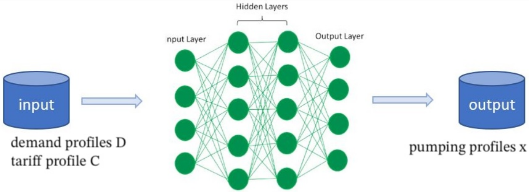
Train a ML model offline on the network historical data
to predict the optimal discrete control profile



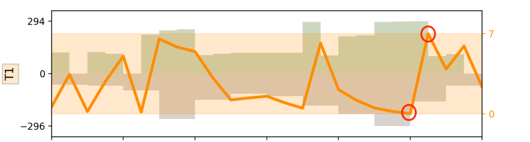
- pros: available history, high seasonality but little variation across years

BAU optimization with Machine Learning

Train a ML model offline on the network historical data
to predict the optimal discrete control profile



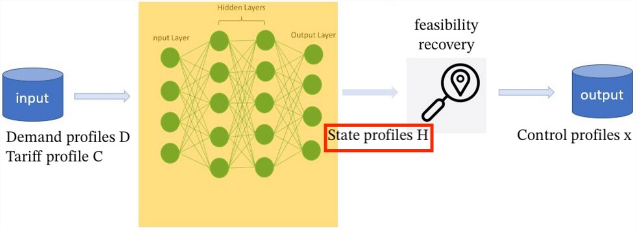
- pros: available history, high seasonality but little variation across years
- cons: feasible decisions x are sparse and scarce in $\{0,1\}^{T\times A}$



- hard to repair an approximate x to meet the storage capacities
- SOA heuristics: tackle storage capacities as soft constraints

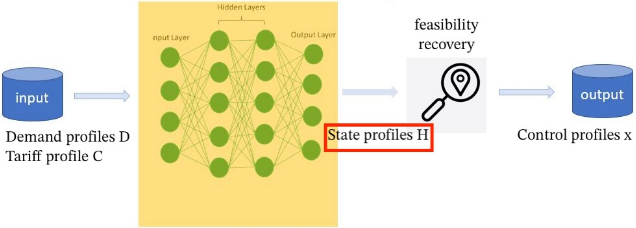
Prop: learn continuous state vs. discrete control

Train a ML model statically on the network historical data to predict the optimal continuous state profiles



Prop: learn continuous state vs. discrete control

Train a ML model statically on the network historical data to predict the optimal continuous state profiles



- regression rather than classification
- local search around a predicted H to restore a feasible X:
  - allows for smoother moves
  - exploits problem structure: time/space decomposition

Mathematical decomposition

MINLP MODEL

$$(\mathcal{P}) : \min_{x,q,H} \sum_{t \in \mathcal{T}} c_t(x_t, q_t) = \sum_{t \in \mathcal{T}} \sum_{a \in \mathcal{A}} (c^0_{ta} x_{ta} + c^1_{ta} q_{ta}) \quad s.t. :$$

$q_t \in \mathcal{E}(H_t, D_t, x_t)$	$\forall$ time t	flow/head equilibrium
$q_{tj} = \sigma_j(H_{(t+1)j} - H_{tj})$	$\forall$ time $t,$ tank j	flow conservation at tanks
$\underline{H}_{tj} \leq H_{tj} \leq \overline{H}_{tj}$	$\forall$ time $t,$ tank j	tank capacities
$x_{ta} \in \{0, 1\}$	$\forall$ time $t,$ arc a	pump status

MINLP MODEL

$$(\mathcal{P}) : \min_{x,q,H} \sum_{t\in\mathcal{T}} c_t(x_t, q_t) = \sum_{t\in\mathcal{T}} \sum_{a\in\mathcal{A}} (c^0_{ta} x_{ta} + c^1_{ta} q_{ta}) \quad s.t. :$$

$q_t \in \mathcal{E}(H_t, D_t, x_t)$	$\forall$ time t	flow/head equilibrium
$q_{tj} = \sigma_j(H_{(t+1)j} - H_{tj})$	$\forall$ time t, tank j	flow conservation at tanks
$\underline{H}_{tj} \leq H_{tj} \leq \overline{H}_{tj}$	$\forall$ time t, tank j	tank capacities
$x_{ta} \in \{0, 1\}$	$\forall$ time t, arc a	pump status

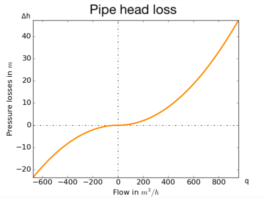
time decomposition: relax/penalize/dualize the flow conservation constraints

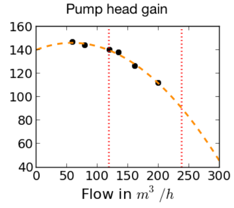
Static Equilibrium Problem $\mathcal{E}(H_t, D_t, x_t)$

At each time t, flow/head equilibrium $(q_t, h_t) \in \mathcal{E}(H_t, D_t, x_t)$ iff

$h_{tj} = H_{tj}$	$\forall$ tank j	tank head
$q_{tj} = D_{tj}$	$\forall$ user j	flow conservation
$x_{ta} = 0 \implies q_{ta} = 0$	$\forall$ arc a	inactive arc
$x_{ta} = 1 \implies h_{ta} = \phi_a(q_{ta})$	$\forall$ arc a	flow/head loss

where ϕ_a is a quadratic antisymmetric fit





Static Equilibrium Problem $\mathcal{E}(H_t, D_t, x_t)$

At each time t, flow/head equilibrium $(q_t, h_t) \in \mathcal{E}(H_t, D_t, x_t)$ iff

$h_{tj} = H_{tj}$	$\forall$ tank j	tank head
$q_{tj} = D_{tj}$	$\forall$ user j	flow conservation
$x_{ta} = 0 \implies q_{ta} = 0$	$\forall$ arc a	inactive arc
$x_{ta} = 1 \implies h_{ta} = \phi_a(q_{ta})$	$\forall$ arc a	flow/head loss

where ϕ_a is a quadratic antisymmetric fit

- nonconvex system; unique solution easy to compute for **given** state H_t and control x_t

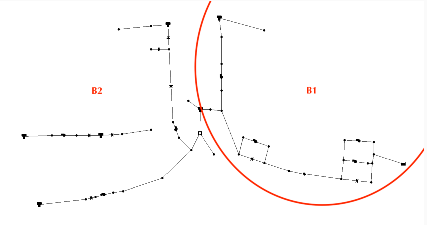
Static Equilibrium Problem $\mathcal{E}(H_t, D_t, x_t)$

At each time t, flow/head equilibrium $(q_t, h_t) \in \mathcal{E}(H_t, D_t, x_t)$ iff

$h_{tj} = H_{tj}$	$\forall$ tank j	tank head
$q_{tj} = D_{tj}$	$\forall$ user j	flow conservation
$x_{ta} = 0 \implies q_{ta} = 0$	$\forall$ arc a	inactive arc
$x_{ta} = 1 \implies h_{ta} = \phi_a(q_{ta})$	$\forall$ arc a	flow/head loss

where ϕ_a is a quadratic antisymmetric fit

- nonconvex system; unique solution easy to compute for **given** state H_t and control x_t
- space decomposition along the tanks; few pumps in each component:



Recover feasibility: from learned H to a feasible X

Tank levels H are coupling elements of the model:

Fixing the tank levels:

1. **Temporal decomposition**: separates the model in independent static equilibrium subproblems:

$$q_t \in \mathcal{E}(H_t, D_t, x_t) \quad \forall \text{ time } t$$

Recover feasibility: from learned H to a feasible X

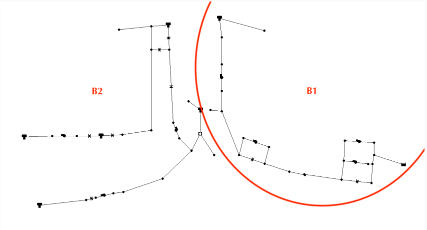
Tank levels H are coupling elements of the model:

Fixing the tank levels:

1. **Temporal decomposition**: separates the model in independent static equilibrium subproblems:

$$q_t \in \mathcal{E}(H_t, D_t, x_t) \quad \forall \text{ time } t$$

2. **Graph decomposition**: separates the static equilibrium subproblems along the tanks



Recover feasibility 1: extended IP (approximate)

Original model

$$\min_{x,q,H} \sum_t c_t(x_t, q_t)$$

$$\begin{aligned}
& q_t \in \mathcal{E}(H_t, D_t, x_t) && \forall t \\
& q_{tj} = \sigma_j(H_{(t+1)j} - H_{tj}) && \forall t, j \\
& \underline{H}_{tj} \leq H_{tj} \leq \overline{H}_{tj} && \forall t, j \\
& x_{ta} \in \{0, 1\} && \forall t, a
\end{aligned}$$

Extended IP [INOC 2019]

$$\min_{y,H} \sum_t \sum_s C_{ts} y_{ts}$$

$$\begin{aligned}
& \sum_s y_{ts} = 1 && \forall t \\
& \sum_s Q_{tsj} y_{ts} = \sigma_j(H_{(t+1)j} - H_{tj}) && \forall t, j \\
& \underline{H}_{tj} \leq H_{tj} \leq \overline{H}_{tj} && \forall t, j \\
& y_{ts} \in \{0, 1\} && \forall s \in \mathcal{S}_t
\end{aligned}$$

given learned $\tilde{H}$:

- solve $\mathcal{E}(\tilde{H}_t, D_t, x_t)$ for each configuration $s := x_t \in \{0,1\}^A$
- compute cost C_{ts} and tank inflows Q_{ts}
- keep $s \in \mathcal{S}_t$ if $Q_{ts} \approx \sigma(\tilde{H}_{(t+1)} - \tilde{H}_t)$
- $|\mathcal{S}_t|$ is limited: symmetry breaking, space decomposition

Recover feasibility 2: variable-splitting (heuristic)

Original model

$$\min_{x,q,H} \sum_t c_t(x_t, q_t)$$

$$q_t \in \mathcal{E}(H_t, D_t, x_t) \qquad \forall t$$

$$d_{tj} = 0 \qquad \forall t, j$$

$$\underline{H}_{tj} \leq H_{tj} \leq \overline{H}_{tj} \qquad \forall t, j$$

$$x_{ta} \in \{0, 1\} \qquad \forall t, a$$

with $d_{tj} = q_{tj} - \sigma_j(H_{(t+1)j} - H_{tj})$

Alternating Direction Method: start with $H = \tilde{H}$

1. solve $\mathcal{P}(H)$ get (x, q)
2. solve $\mathcal{P}(x, q)$ get H
3. stop if $\|d_t\| < \epsilon$ or goto 1 and possibly update ρ

Variable-splitting [ISCO 2024]

$$\mathcal{P}(H) : \min_{x,q} \sum_t c_t(x_t, q_t) + \rho_t d_t$$

$$q_t \in \mathcal{E}(H_t, D_t, x_t) \qquad \forall t$$

$$x_{ta} \in \{0, 1\} \qquad \forall t, a$$

$$\downarrow \qquad \uparrow$$

$$\mathcal{P}(x, q) : \min_{H} \rho_t d_t$$

$$q_t \in \mathcal{E}(H_t, D_t, x_t) \qquad \forall t$$

$$\underline{H}_{tj} \leq H_{tj} \leq \overline{H}_{tj} \qquad \forall t, j$$

Deep Learning

Deep Learning architecture

$$\mathcal{H} : (D, C) \rightarrow H$$

- Both input (D, C) and output H resemble temporal sequential data

Deep Learning architecture

$$\mathcal{H} : (D, C) \rightarrow H$$

- Both input (D, C) and output H resemble temporal sequential data
- Naive inception architecture: several parallel convolutional with various kernel sizes to capture local trends in the input data

Deep Learning architecture

$$\mathcal{H} : (D, C) \rightarrow H$$

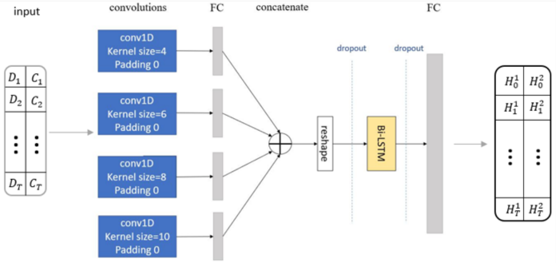
- Both input (D, C) and output H resemble temporal sequential data
- Naive inception architecture: several parallel convolutional with various kernel sizes to capture local trends in the input data
- + an LSTM unit after concatenation to capture temporal dependencies

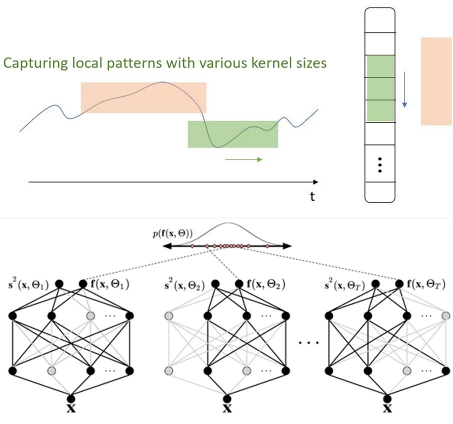
Deep Learning architecture

$$\mathcal{H} : (D, C) \to H^1, H^2, H^3, \dots, H^{50}$$

- Both input (D, C) and output H resemble temporal sequential data
- Naive inception architecture: several parallel convolutional with various kernel sizes to capture local trends in the input data
- + an LSTM unit after concatenation to capture temporal dependencies
- + Monte Carlo dropout to generate multiple outputs H^k to implement diversification in local search with multi-start

Deep Learning architecture





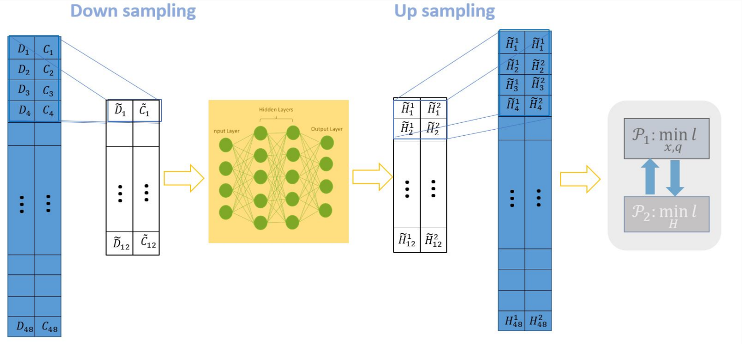
Monte-Carlo drop out

Scaling: if no training data are available

- Training set: daily data (D, C) with associated optimal H
- Computing an optimal H for each input data (D, C) is not viable for fine time-discretization, e.g., $T = 24$ or $T = 48$

SCALING: IF NO TRAINING DATA ARE AVAILABLE

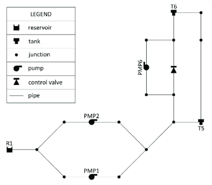
- Training set: daily data (D, C) with associated optimal H
- Computing an optimal H for each input data (D, C) is not viable for fine time-discretization, e.g., $T = 24$ or $T = 48$
- Scaling: train the DL model using coarse-grained resolution data (D, C, H), e.g. $T = 12$
- resize/resample input and output by linear interpolation



Experiments

Experimental set

- data generation: 6 years of daily instances (D, C) drawn from realistic highly seasonal data adapted to the *Van Zyl* network

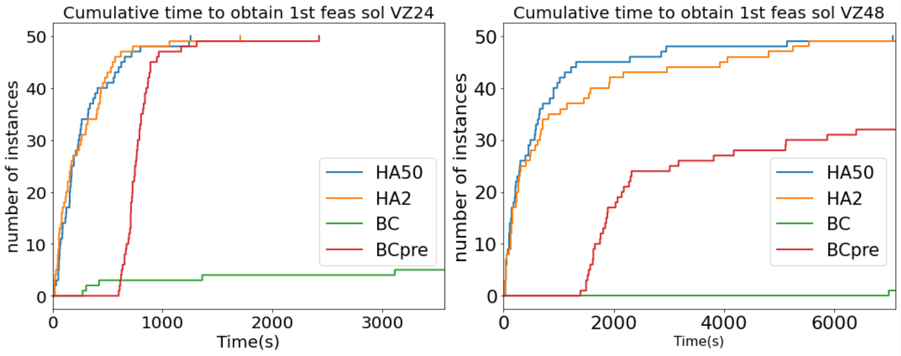


- data collection: solve with coarse-time $(T = 12)$ by a specialized branch-and-check algorithm **BC** [Opt&Eng 2021] with advanced preprocessing **BC+Pre** [ICAE 2022]
- test set: 50 instances with $T = 12, 24, 48$
- compare the **first feasible solutions** computed with DL+ADM for a fixed penalty value $\rho = 50$ or $\rho = 2$ (**HA50, HA2**) with **BC** and **BC+Pre**

GAP TO THE BEST LB [BCPRE] AND AVG TIME

		#solved	Mean%	Min%	Max%	time (s)
VZ12	**HA50**	49	6.6	0.0	21.2	254
1800s	**HA2**	44	4.6	0.0	11.3	305
	BC	48	5.4	1.6	12.5	121
	BC+Pre	**50**	**4.3**	0.4	12.4	**124**
VZ24	**HA50**	**50**	9.5	3.3	23.4	**285**
3600s	**HA2**	**50**	8.4	3.4	16.3	**279**
	BC	5	11.1	7.2	12.6	1097
	BC+Pre	**50**	**7.5**	2.4	39.6	809
VZ48	**HA50**	**50**	9.8	3.8	21.0	**776**
7200s	**HA2**	**49**	10.3	4.4	19.7	1014
	BC	1	-	-	-	-
	BC+Pre	32	6.4	3.4	8.9	2517

Number of solutions wrt time



Conclusion and perspective

- a combination of complementary data and mathematical models to reach feasible high-quality solutions in a short time
- models are independent, other combinations exist
- local search in the state H-space vs control x-space: exploiting time and space decomposition
- a natural mapping $H \mapsto x$ exists in many control application
- future work: convergence to optimality

References

- [ISCO 2024] Demassey S., Sessa V., Tavakoli A. Deep learning and alternating direction method for discrete control with storage. In International Symposium on Combinatorial Optimization 2024.
- [ICAE 2022] Tavakoli A., Sessa V., Demassey S. Strengthening mathematical models for pump scheduling in water distribution. In 14th International Conference on Applied Energy 2022.
- [Opt&Eng 2021] Bonvin G., Demassey S., Lodi A. Pump scheduling in drinking water distribution networks with an LP/NLP-based branch and bound. Optimization and Engineering 2021.
- [INOC 2019] Bonvin G., Demassey S. Extended linear formulation of the pump scheduling problem in water distribution networks. In International Network Optimization Conference 2019.

- papers available at `https://sofdem.github.io/`
- code available at `https://github.com/sofdem/gopslpnlpbb`