From Energy Systems to Material Science: Optimization as a Common Denominator for a Sustainable Future

Milena Petkovic[1,2], Janina Zittel[2], Thorsten Koch[2], Natasa Dropka[1]
Leibniz Institute for Crystal Growth
Zuse Institute Berlin

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung
MODAL FORSCHUNGSCAMPUS

CO@WORK, 25.09.2024, Berlin, Germany

Zuse Institute Berlin

- An interdisciplinary research institute for **applied mathematics** and **data-intensive HPC**
- Focuses on **modeling**, **simulation** and **optimization** with scientific cooperation partners from academia and industry
- Priority application areas: the life and materials sciences, logistics, infrastructure planning, and OR
- Approx. 230 employees

Research Departments:
Modeling and Simulation of Complex Processes
Visual and Data-Centric Computing
AI in Society, Science, and Technology
Applied Algorithmic Intelligence Methods
Network Optimization
Distributed Algorithms
Supercomputing

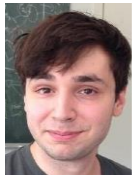
Research Service Units:
IT and Data Services
Digital Data and Information for Society, Science, and Culture (digiS)
NHR Center

Applied Algorithmic Intelligence Methods (A²IM) Department


Dreßke, Carsten


Shinano, Dr. Yuji


Koch, Prof. Dr. Thorsten


Turner, Mark Ruben


Le, Dr. Thi Thai


Vanaret, Dr. Charlie William


Le, Thi Thanh


Yüksel-Ergün, Dr. Inci


Pedersen, Jaap


Zakiyeva, Dr. Nazgul


Petkovic, Dr. Milena

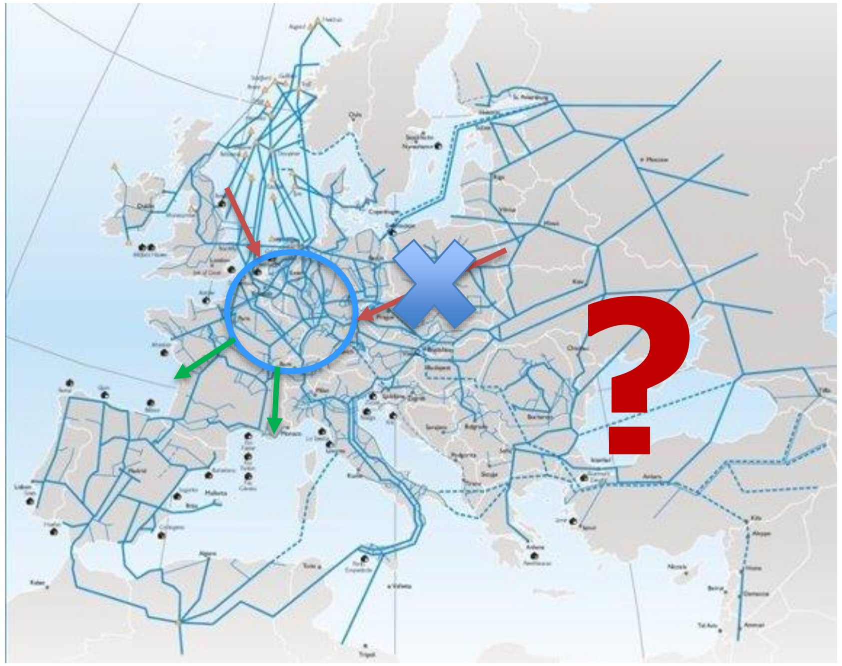

Zittel, Dr. Janina

A²IM

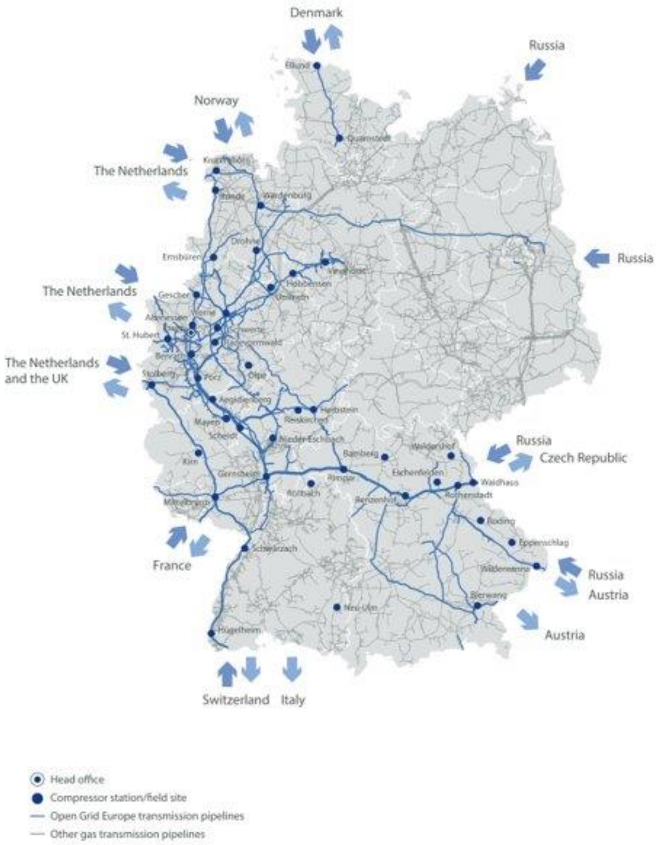
Prescriptive Systems	Predictive Methods
Multi Energy Systems	Sustainable Energy Planning

- The A²IM department is applying advanced AI methods from Mathematical Optimization and Machine Learning to explore new smart algorithmic solutions for real-world problems.
- Our research is concerned, in particular, with better planning, extension, and control of vital and complex infrastructure networks.

German Gas Network:
The Heart of European Gas Transport



Critical infrastructure to supply Central, Southern and Western Europe with natural gas



Impact on decision making process:

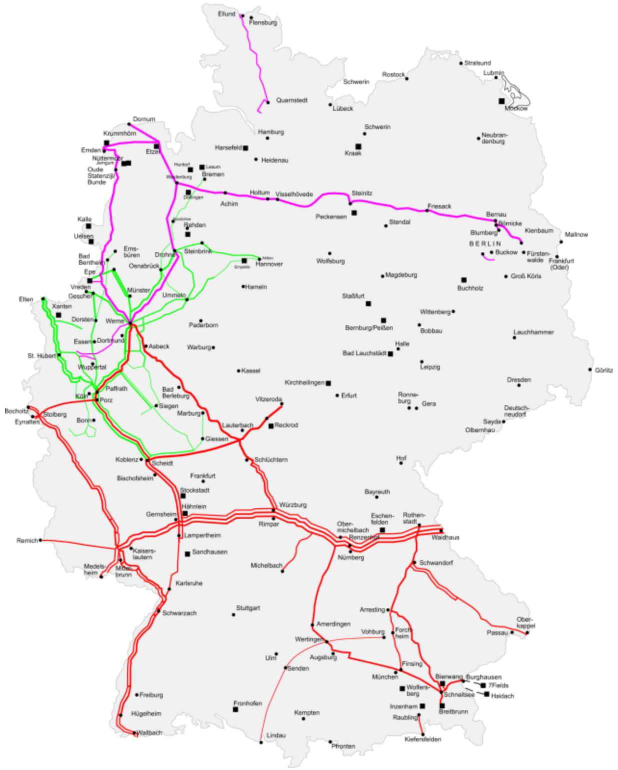
- Experience from the past is obsolete
- Not sufficient historical data for new scenarios

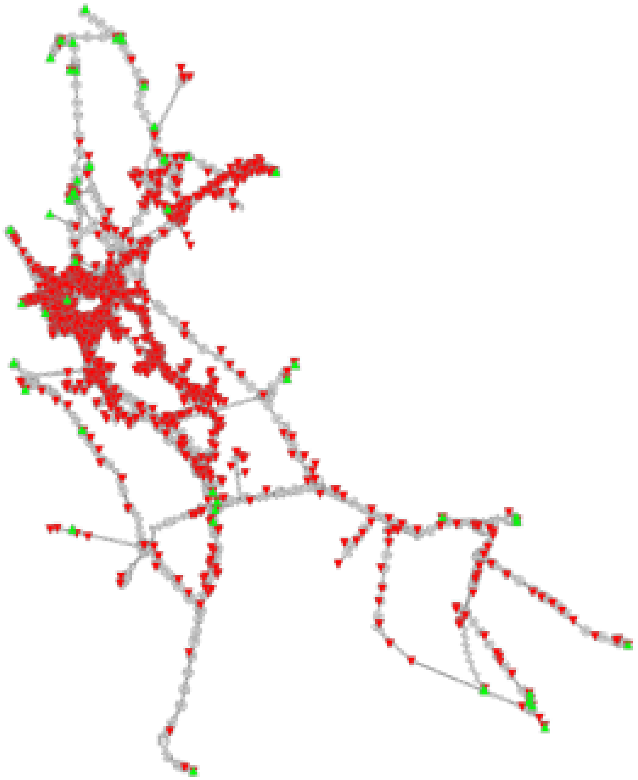
Modal EnergyLab & OGE

ZIB ZUSE INSTITUTE BERLIN

- **Operates the longest network of pipelines in Germany ~ 12,000 km**
- More than 450 customers
- 120 billion kWh of energy transported every month
- ~ 25% of the total energy demand in Germany/Europe is supplied by natural gas
- In order to guarantee a secure supply in the future, further IT systems are needed to support the dispatcher

Full Network	
1,194	Entries+Exits
6,247	Pipes
3,403	Valves
291	Control valves
22	Resistors
100	Compressors





The Unbounded European Gas Market since 2009

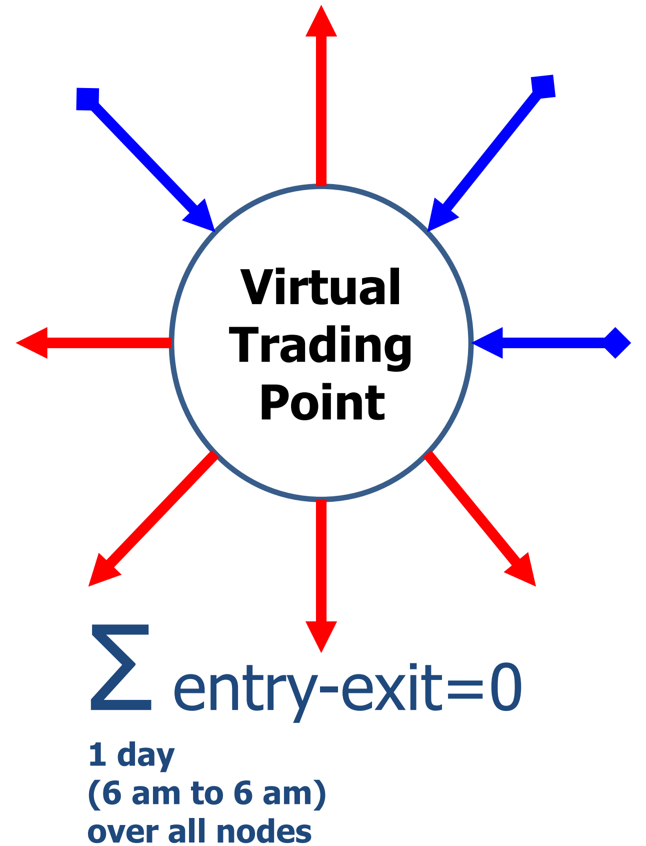
Gas Trading Companies ∩ Transport System Operators = ∅

REGULATION (EC) No 715/2009 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

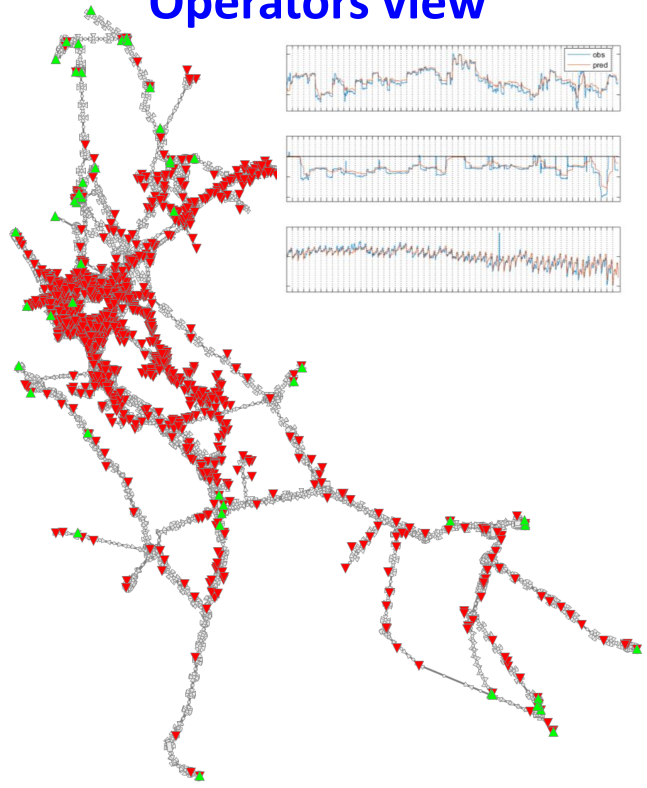
Traders buy and sell gas | transmission system operators transport it.

- Capacity products are typically either **firm** = sure deliver **or flexible** = best effort
- The traders give transport orders to the TSO within the limits of the acquired capacity.
- The TSO then has to fulfill the order accordingly.

Trader view



Transmission System Operators view

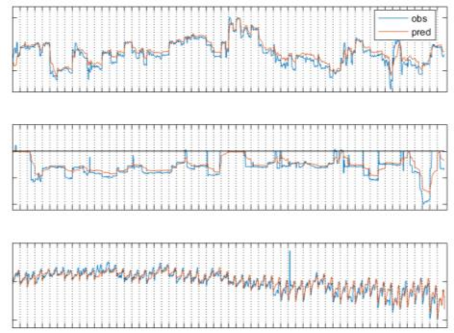


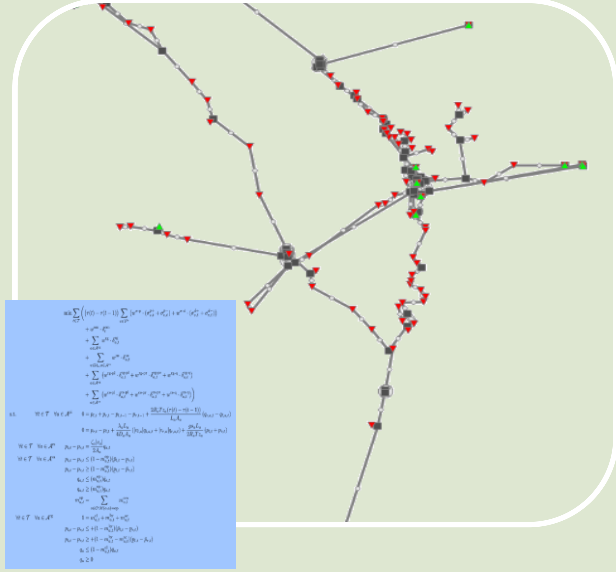
KOMPASS/ NAVI

Forecast

Flow Model

Station Model





Station Mittelbrunn

Station Stolberg

Station Gernsheim

Station Scheidt

...

Recomandations

The KOMPASS near real-time Decision Support System

DATA
- Control Database aggregates new data from 3 external systems every 15 min.

Initial State
- Based on the data a valid initial state is constructed.

Initial solution
- Several heuristics try to devise possible solutions to the overall flow situation.

MILP
- Mixed-Integer Programing (MILP) based flow model
- warm starting using the heuristic solutions, computes an optimized overall strategy.

Demand curves
- Result of the flow model defines the demand curves for the individual stations.

Station model
- For each station in parallel:
  - Heuristics quickly devise initial solutions.
  - The MILP station model computes optimized actions for the stations.

Recomandations.
- Actions are combined and postprocessed into recommendations.

Mixed-Integer Program for Optimal Network Control

Objective function

$$
\begin{aligned}
\min \sum_{t\in\mathcal{T}} \Bigg(&(\tau(t)-\tau(t-1)) \sum_{v\in\mathcal{V}^{\mathrm{b}}} \left(w^{\sigma\text{-p}}\cdot(\sigma^{p+}_{v,t}+\sigma^{p-}_{v,t}) + w^{\sigma\text{-d}}\cdot(\sigma^{d+}_{v,t}+\sigma^{d-}_{v,t})\right) \\
&+ w^{\mathrm{om}}\cdot\delta^{\mathrm{om}}_{t} \\
&+ \sum_{a\in\mathcal{A}^{\mathrm{rg}}} w^{\mathrm{rg}}\cdot\delta^{\mathrm{rg}}_{a,t} \\
&+ \sum_{u\in\mathcal{U}_a,a\in\mathcal{A}^{\mathrm{cs}}} w^{\mathrm{us}}\cdot\delta^{\mathrm{us}}_{u,t} \\
&+ \sum_{a\in\mathcal{A}^{\mathrm{rg}}} \left(w^{\text{rg-pl}}\cdot\delta^{\text{rg-pl}}_{a,t} + w^{\text{rg-pr}}\cdot\delta^{\text{rg-pr}}_{a,t} + w^{\text{rg-q}}\cdot\delta^{\text{rg-q}}_{a,t}\right) \\
&+ \sum_{a\in\mathcal{A}^{\mathrm{cs}}} \left(w^{\text{cs-pl}}\cdot\delta^{\text{cs-pl}}_{a,t} + w^{\text{cs-pr}}\cdot\delta^{\text{cs-pr}}_{a,t} + w^{\text{cs-q}}\cdot\delta^{\text{cs-q}}_{a,t}\right)\Bigg)
\end{aligned}
$$

Momentum equation

s.t. $\forall t\in\mathcal{T} \quad \forall a\in\mathcal{A}^{\mathrm{pi}}$

$$0 = p_{l,t}+p_{r,t}-p_{l,t-1}-p_{r,t-1}+\frac{2R_sTz_a(\tau(t)-\tau(t-1))}{L_aA_a}(q_{r,a,t}-q_{l,a,t})$$

Pressure loss

$$0 = p_{r,t}-p_{l,t}+\frac{\lambda_aL_a}{4D_aA_a}(|v_{l,a}|q_{l,a,t}+|v_{r,a}|q_{r,a,t})+\frac{gs_aL_a}{2R_sTz_a}(p_{l,t}+p_{r,t})$$

Resistor constraint

$\forall t\in\mathcal{T} \quad \forall a\in\mathcal{A}^{\mathrm{rs}}$

$$p_{l,t}-p_{r,t} = \frac{\zeta_a|v_a|}{2A_a}q_{a,t}$$

Valves

$\forall t\in\mathcal{T} \quad \forall a\in\mathcal{A}^{\mathrm{va}}$

$$
\begin{aligned}
p_{l,t}-p_{r,t} &\le (1-m^{\mathrm{op}}_{a,t})(\bar{p}_{l,t}-\underline{p}_{r,t}) \\
p_{l,t}-p_{r,t} &\ge (1-m^{\mathrm{op}}_{a,t})(\underline{p}_{l,t}-\bar{p}_{r,t}) \\
q_{a,t} &\le (m^{\mathrm{op}}_{a,t})\bar{q}_{a,t} \\
q_{a,t} &\ge (m^{\mathrm{op}}_{a,t})\underline{q}_{a,t} \\
m^{\mathrm{op}}_{a,t} &= \sum_{o\in\mathcal{O}:M(o,a)=\mathrm{op}} m^{\mathrm{om}}_{o,t}
\end{aligned}
$$

Regulators

$\forall t\in\mathcal{T} \quad \forall a\in\mathcal{A}^{\mathrm{rg}}$

$$
\begin{aligned}
1 &= m^{\mathrm{cl}}_{a,t}+m^{\mathrm{by}}_{a,t}+m^{\mathrm{ac}}_{a,t} \\
p_{l,t}-p_{r,t} &\le +(1-m^{\mathrm{by}}_{a,t})(\bar{p}_{l,t}-\underline{p}_{r,t}) \\
p_{l,t}-p_{r,t} &\ge +(1-m^{\mathrm{by}}_{a,t}-m^{\mathrm{ac}}_{a,t})(\underline{p}_{l,t}-\bar{p}_{r,t}) \\
q_a &\le (1-m^{\mathrm{cl}}_{a,t})\bar{q}_{a,t} \\
q_a &\ge 0
\end{aligned}
$$

Compressor station

$\forall t\in\mathcal{T} \quad \forall a\in\mathcal{A}^{\mathrm{cs}}$

$$
\begin{aligned}
1 &= \sum_{c\in\mathcal{C}_a} m^{\mathrm{cf}}_{c,a,t}+m^{\mathrm{by}}_{a,t}+m^{\mathrm{cl}}_{a,t} \\
p_{l,t} &= p^{\mathrm{by}}_{a,t}+p^{\text{l-cl}}_{a,t}+\sum_{c\in\mathcal{C}_a} p^{\text{l-cf}}_{c,a,t} \\
p_{r,t} &= p^{\mathrm{by}}_{a,t}+p^{\text{r-cl}}_{a,t}+\sum_{c\in\mathcal{C}_a} p^{\text{r-cf}}_{c,a,t} \\
q_{a,t} &= q^{\mathrm{by}}_{a,t}+\sum_{c\in\mathcal{C}_a} q^{\mathrm{cf}}_{c,a,t}
\end{aligned}
$$

$$
\begin{aligned}
\underline{p}^{\text{l-cf}}_{c,a,t}m^{\mathrm{cf}}_{c,a,t} \le p^{\text{l-cf}}_{c,a,t} &\le \bar{p}^{\text{l-cf}}_{c,a,t}m^{\mathrm{cf}}_{c,a,t} \quad \forall c\in\mathcal{C}_a \\
\underline{p}^{\text{r-cf}}_{c,a,t}m^{\mathrm{cf}}_{c,a,t} \le p^{\text{r-cf}}_{c,a,t} &\le \bar{p}^{\text{r-cf}}_{c,a,t}m^{\mathrm{cf}}_{c,a,t} \quad \forall c\in\mathcal{C}_a \\
\underline{q}^{\mathrm{cf}}_{c,a,t}m^{\mathrm{cf}}_{c,a,t} \le q^{\mathrm{cf}}_{c,a,t} &\le \bar{q}^{\mathrm{cf}}_{c,a,t}m^{\mathrm{cf}}_{c,a,t} \quad \forall c\in\mathcal{C}_a \\
\underline{p}^{\mathrm{by}}_{a,t}m^{\mathrm{by}}_{a,t} \le p^{\mathrm{by}}_{a,t} &\le \bar{p}^{\mathrm{by}}_{a,t}m^{\mathrm{by}}_{a,t} \\
\underline{q}^{\mathrm{by}}_{a,t}m^{\mathrm{by}}_{a,t} \le q^{\mathrm{by}}_{a,t} &\le \bar{q}^{\mathrm{by}}_{a,t}m^{\mathrm{by}}_{a,t} \\
\underline{p}^{\text{l-cl}}_{a,t}m^{\mathrm{cl}}_{a,t} \le p^{\text{l-cl}}_{a,t} &\le \bar{p}^{\text{l-cl}}_{a,t}m^{\mathrm{cl}}_{a,t} \\
\underline{p}^{\text{r-cl}}_{a,t}m^{\mathrm{cl}}_{a,t} \le p^{\text{r-cl}}_{a,t} &\le \bar{p}^{\text{r-cl}}_{a,t}m^{\mathrm{cl}}_{a,t}
\end{aligned}
$$

$$w\cdot p^{\text{l-cf}}_{c,a,t}+x\cdot p^{\text{r-cf}}_{c,a,t}+y\cdot q^{\mathrm{cf}}_{c,a,t}+zm^{\mathrm{cf}}_{c,a,t}\le 0 \quad \forall(w,x,y,z)\in\mathcal{H}_c \quad \forall c\in\mathcal{C}_a$$

Station selection

$$
\begin{aligned}
m^{\mathrm{by}}_{a,t} &= \sum_{o\in\mathcal{O}:M(o,a)=\mathrm{by}} m^{\mathrm{om}}_{o,t} \\
m^{\mathrm{cl}}_{a,t} &= \sum_{o\in\mathcal{O}:M(o,a)=\mathrm{cl}} m^{\mathrm{om}}_{o,t} \\
m^{\mathrm{cf}}_{c,a,t} &= \sum_{o\in\mathcal{O}:M(o,a)=c} m^{\mathrm{om}}_{o,t} \quad \forall c\in\mathcal{C}_a
\end{aligned}
$$

$\forall t\in\mathcal{T}$

$$
\begin{aligned}
1 &= \sum_{o\in\mathcal{O}} m^{\mathrm{om}}_{o,t} \\
1 &= \sum_{f\in\mathcal{F}} m^{\mathrm{fd}}_{f,t}
\end{aligned}
$$

Flow conservation and demand

$\forall t\in\mathcal{T} \quad \forall v\in\mathcal{V}^{\mathrm{b}}$

$$
\begin{aligned}
0 &= \sum_{(l,v)=a\in\mathcal{A}^{\mathrm{pi}}} q_{v,a,t} - \sum_{(v,r)=a\in\mathcal{A}^{\mathrm{pi}}} q_{v,a,t} \\
&\quad + \sum_{(l,v)=a\in\mathcal{A}\setminus\mathcal{A}^{\mathrm{pi}}} q_{a,t} - \sum_{(v,r)=a\in\mathcal{A}\setminus\mathcal{A}^{\mathrm{pi}}} q_{a,t} + d_{v,t} \\
d_{v,t} &\ge (1-\sum_{f=(f^+,f^-)\in\mathcal{F}:v\notin f^-} m^{\mathrm{fd}}_{f,t})\underline{d}_{v,t} \\
d_{v,t} &\le (1-\sum_{f=(f^+,f^-)\in\mathcal{F}:v\notin f^+} m^{\mathrm{fd}}_{f,t})\bar{d}_{v,t} \\
\hat{p}_{v,t} &= p_{v,t}-\sigma^{p+}_{v,t}+\sigma^{p-}_{v,t}
\end{aligned}
$$

$\forall t\in\mathcal{T} \quad \forall v\in\mathcal{V}^{0}$

$$
\begin{aligned}
0 &= \sum_{(l,v)=a\in\mathcal{A}^{\mathrm{pi}}} q_{v,a,t} - \sum_{(v,r)=a\in\mathcal{A}^{\mathrm{pi}}} q_{v,a,t} \\
&\quad + \sum_{(l,v)=a\in\mathcal{A}\setminus\mathcal{A}^{\mathrm{pi}}} q_{a,t} - \sum_{(v,r)=a\in\mathcal{A}\setminus\mathcal{A}^{\mathrm{pi}}} q_{a,t}
\end{aligned}
$$

$\forall t\in\mathcal{T} \quad \forall o\in\mathcal{O}$

$$m^{\mathrm{om}}_{o,t} \le \sum_{(o,f)\in\mathcal{OF}} m^{\mathrm{fd}}_{f,t}$$

Exit pressure

$\forall t\in\mathcal{T} \quad \forall v\in\mathcal{V}^{\text{b-ex}}$

$$p_{v,t} \le \bar{p}^{\mathrm{exit}}_{v} + (1-\sum_{f=(f^+,f^-)\in\mathcal{F}:v\in f^-} m^{\mathrm{fd}}_{f,t})(\bar{p}_{v,t}-\bar{p}^{\mathrm{exit}}_{v})$$

Flow direction

$\forall t\in\mathcal{T} \quad \forall(f,\mathcal{V}^{w_1},\mathcal{V}^{w_2})\in\mathcal{W}$

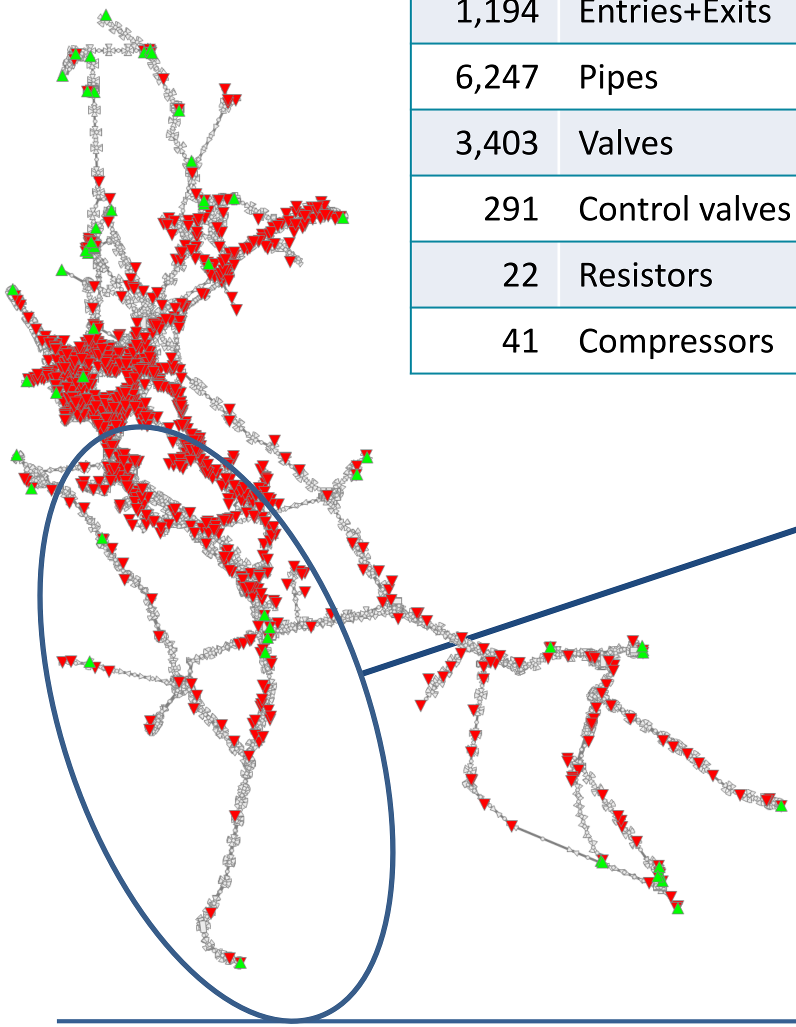
$$0 \le (1-m^{\mathrm{fd}}_{f,t})C_1 - \sum_{v\in\mathcal{V}^{w_1}} \mathrm{sgn}(f,v)d_{v,t} + \sum_{v\in\mathcal{V}^{w_2}} \mathrm{sgn}(f,v)d_{v,t}$$

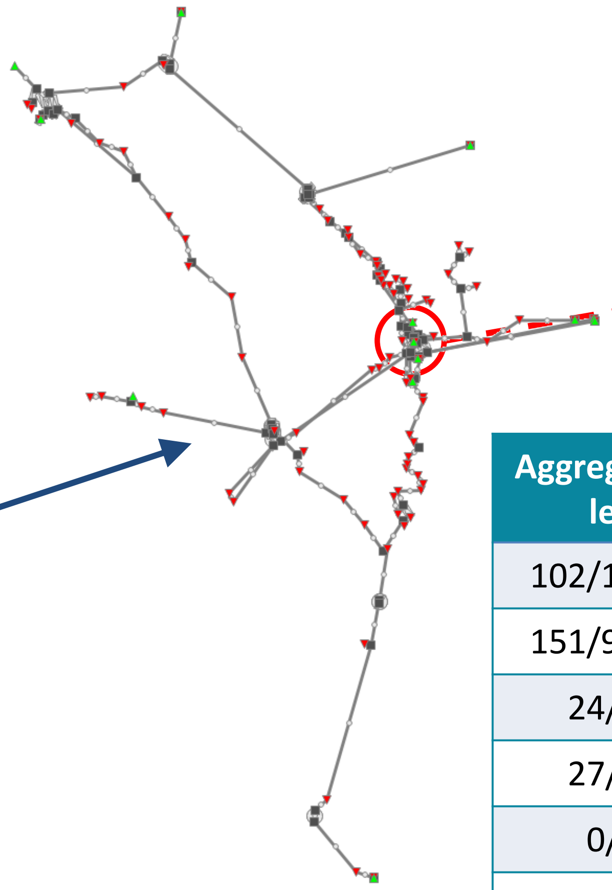
$\forall t\in\mathcal{T} \quad \forall g\in\mathcal{FG}$

$$\hat{d}_{g,t} = \sum_{v\in g}\left(d_{v,t}-\sigma^{d+}_{v,t}+\sigma^{d-}_{v,t}\right)$$

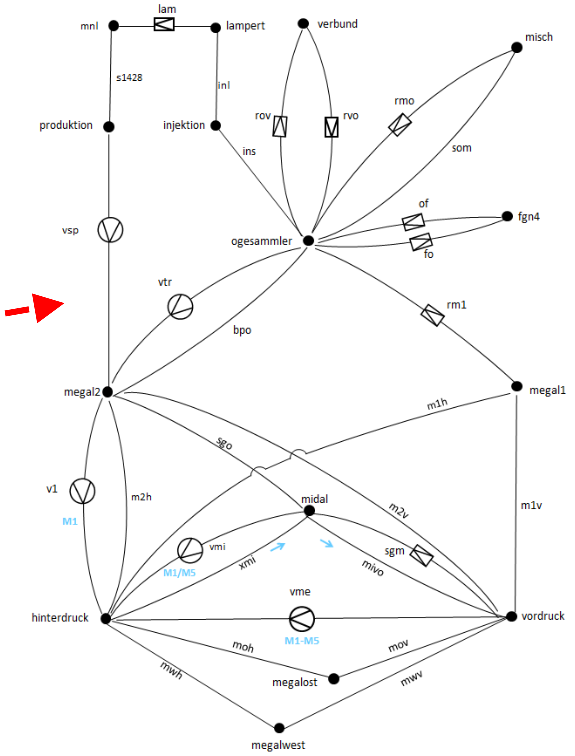
Determine Transient Gas Flows with Network Optimization

Full Network	
1,194	Entries+Exits
6,247	Pipes
3,403	Valves
291	Control valves
22	Resistors
41	Compressors

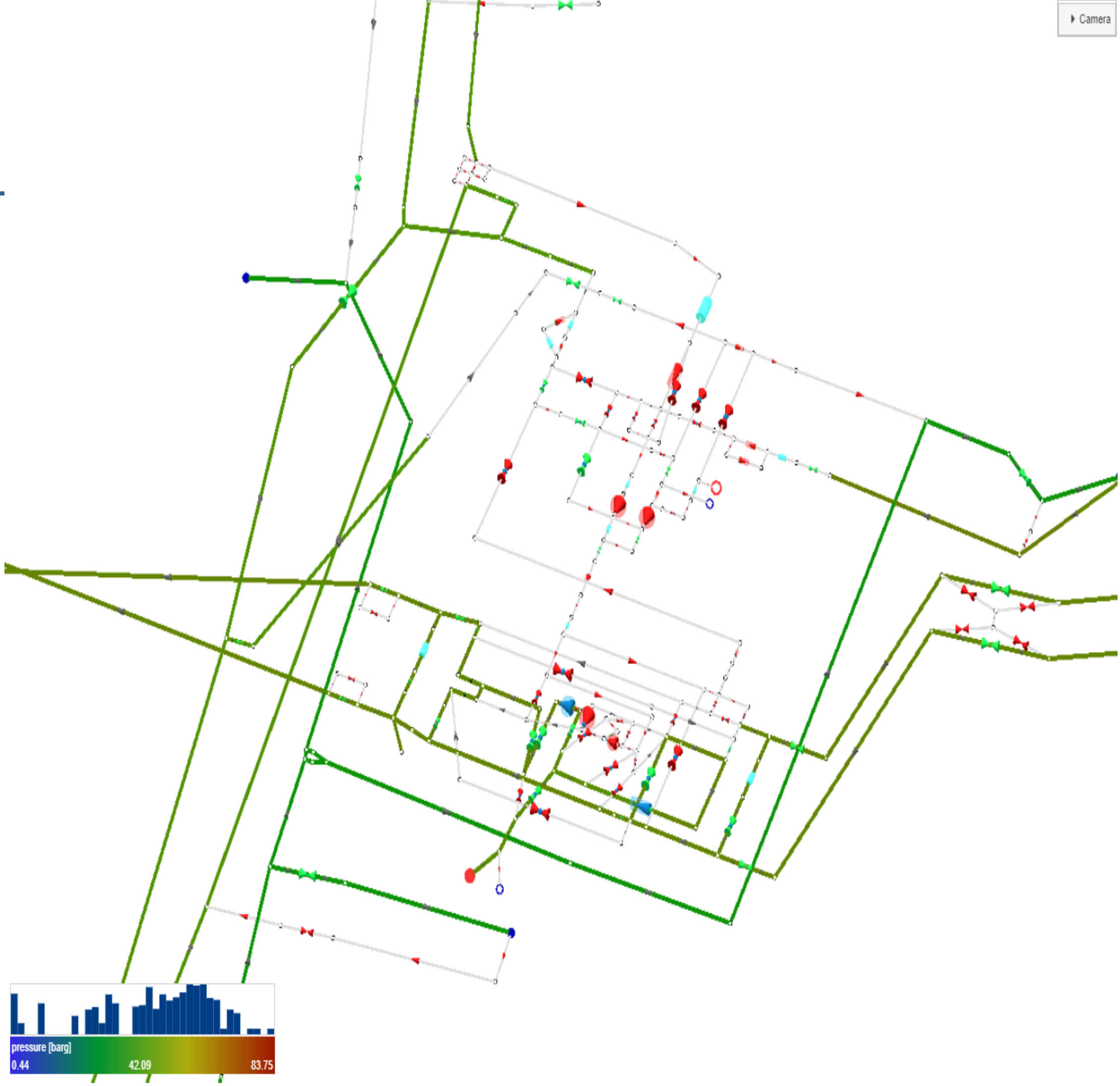




Aggregated subnet with high level station model	
102/124	Entries+Exits
151/964	Pipes
24/81	Valves
27/50	Control valves
0/11	Resistors
16/16	Compressors

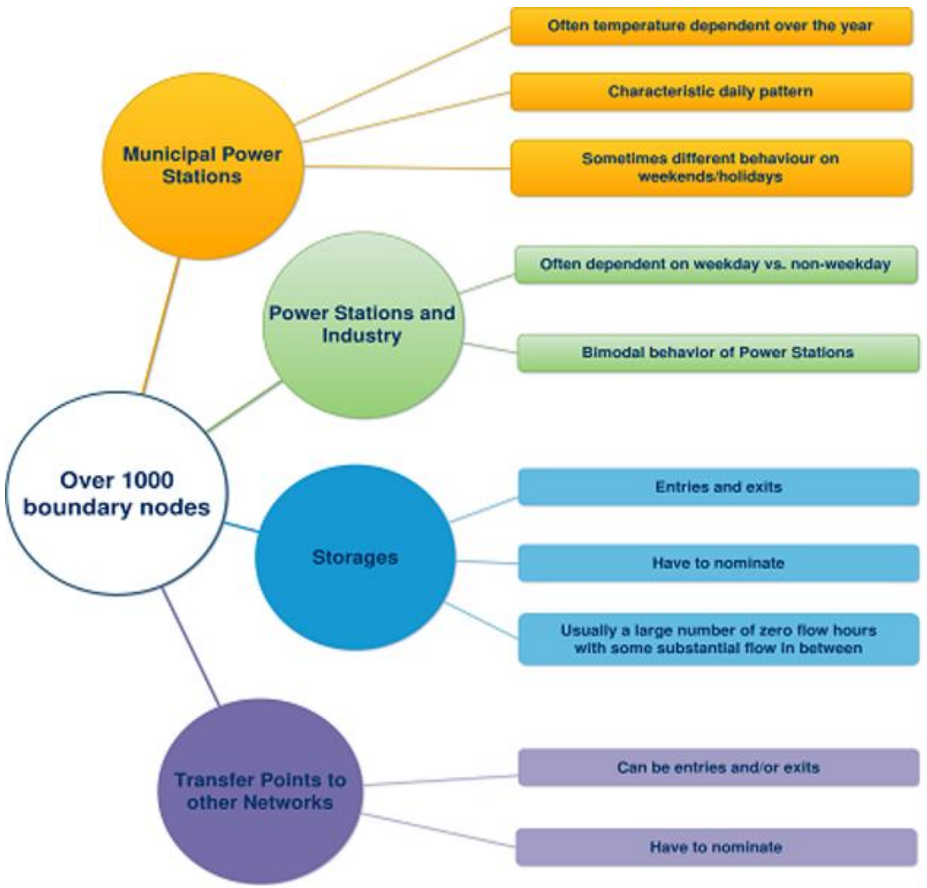


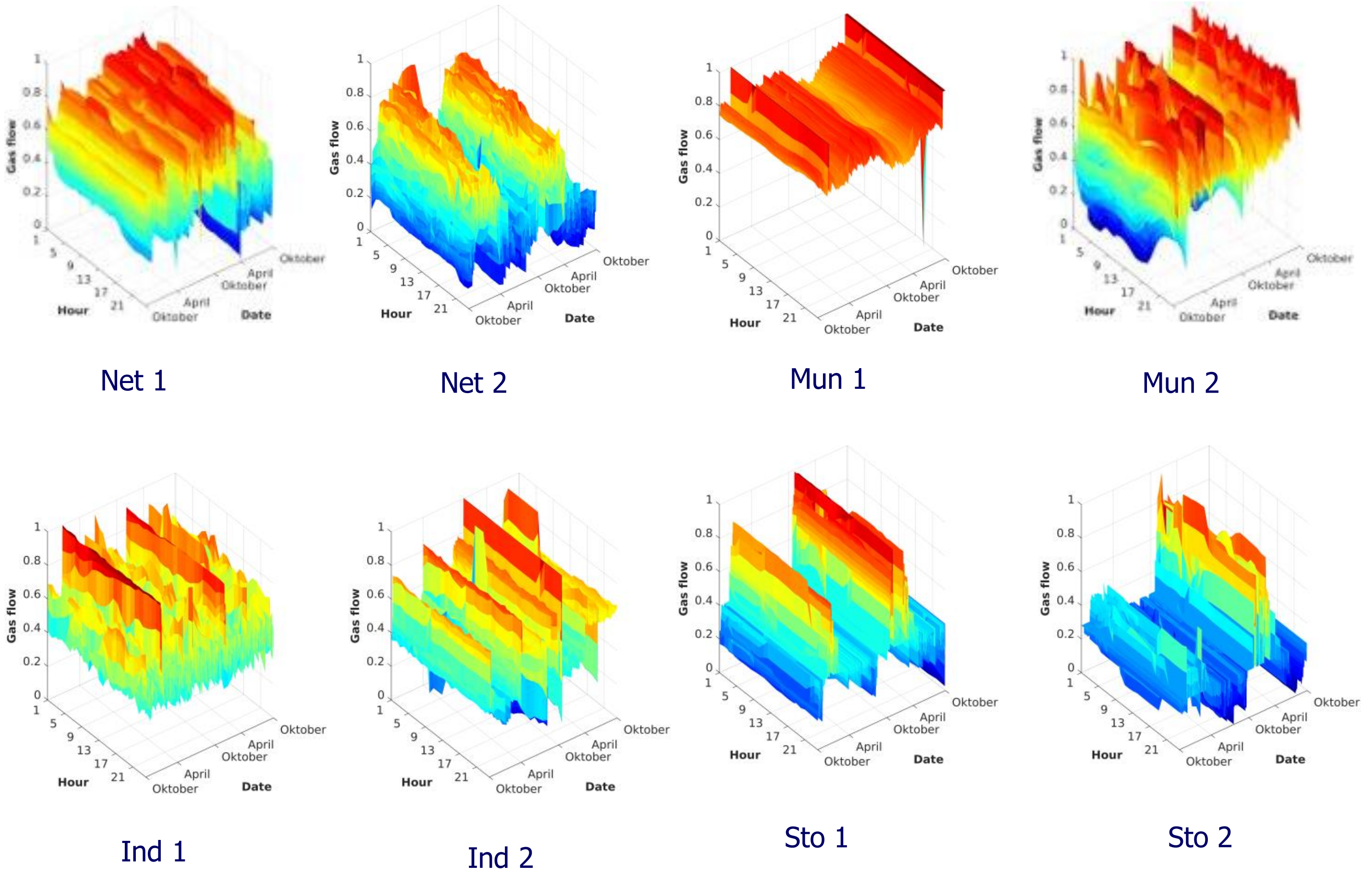
The Combinatorics of Gernsheim



- 30,000,000,000,000,000 mathematically possible combinations of valve and compressor states
- 200,000 feasible operation modes identified based on practitioners knowledge
- 1,285 relevant operation modes extracted using analytical evaluation of historical data

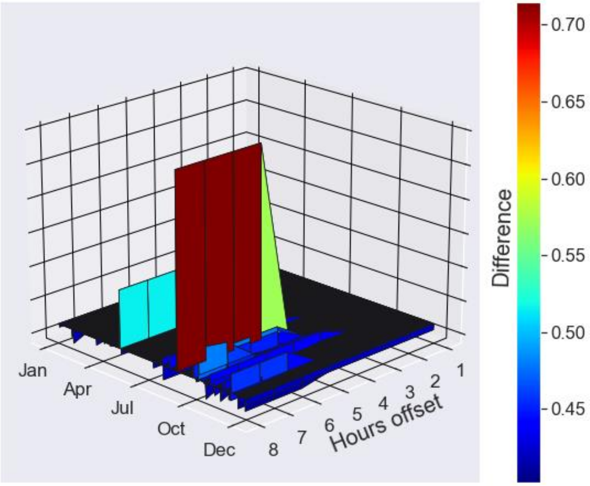
Energy Network Time Series

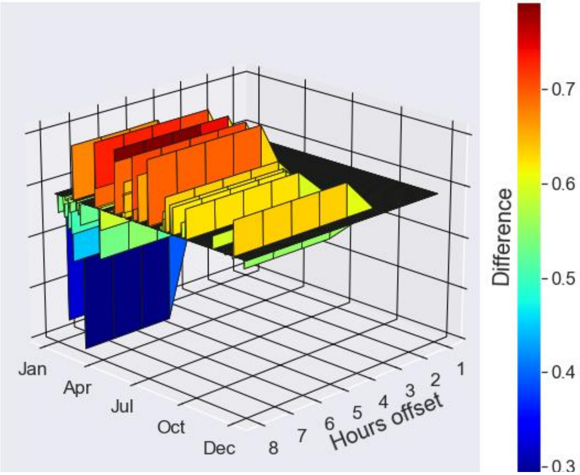


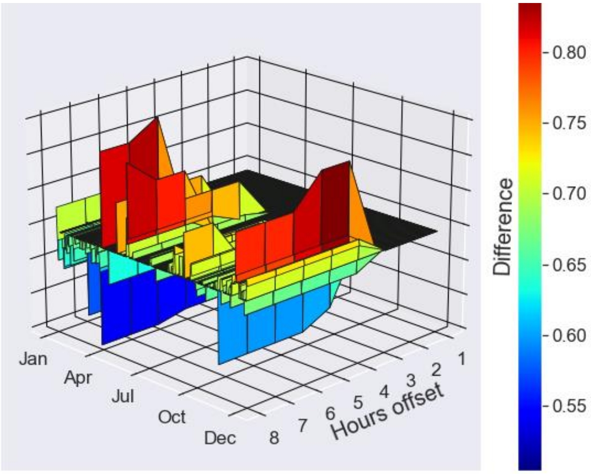


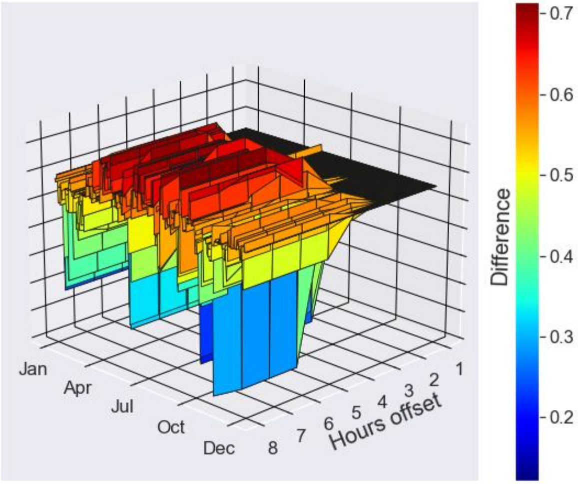
Petkovic, M., Chen, Y., Gamrath, I. *et al.* A hybrid approach for high precision prediction of gas flows. *Energy Syst* (2021). https://doi.org/10.1007/s12667-021-00466-4
Petkovic, M. Koch T., Zittel, J. Deep learning for spatio-temporal supply and demand forecasting in natural gas transmission networks, Energy Science and Engineering (2021). https://doi.org/10.1002/ese3.932
Zakiyeva, N., Petkovic, M. (2022). Modeling and Forecasting Gas Network Flows with Multivariate Time Series and Mathematical Programming Approach. In: Operations Research Proceedings 2021.

Energy Network Time Series: Nominations

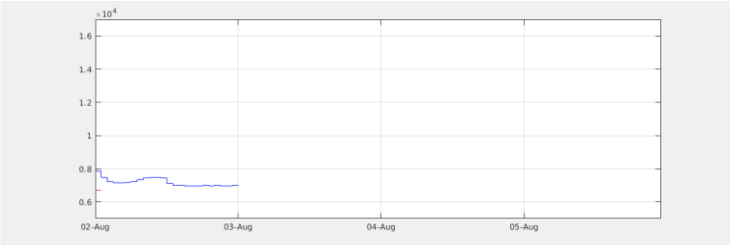


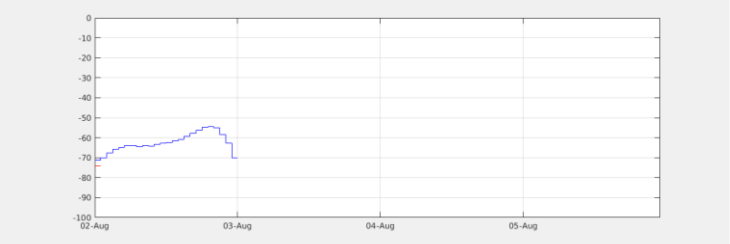


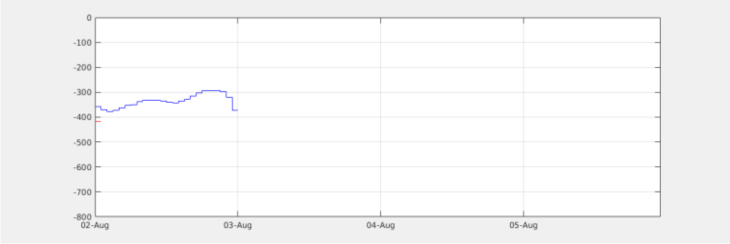




Forecasting energy time series – OPT method







Preprocessing

Solve the most relevant points with sophisticated forecast model,
use computationally less expensive model for less important points.

- training set $\{1, \ \dots, T\}$
- features $\rho_i(*) \in \mathbb{R}, \ i = 1, \ \dots, p$
- flow values $F_t, \ t \in \{1, \ \dots, T\}$
- historical data $*_t, \ t \in \{1, \ \dots, T\}$
- max. number of chosen features B
- weights w_i
- over- and under-estimator $e_t^+, \ e_t^-$
- decision variables x_i

Step 1 (offline computation):
Solve a MIP to compute sparse solutions leading to optimal feature sets for each node at each hour

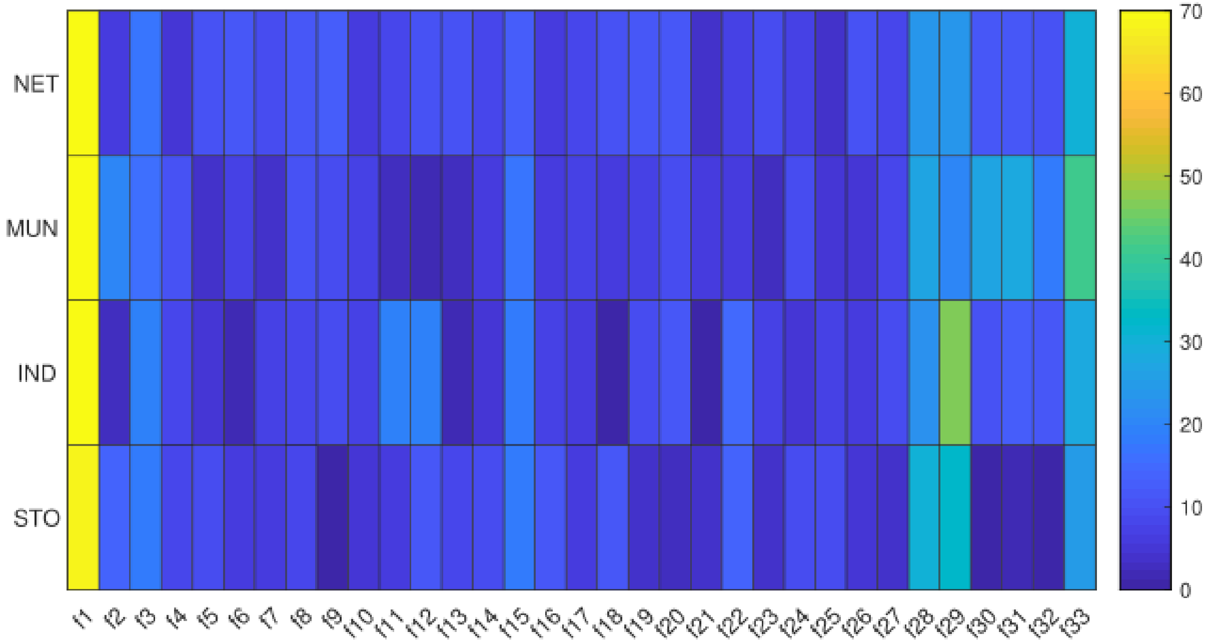
$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} e_t^+ + e_t^- & \\
\text{s.t.} \quad & \sum_{i=1}^{p} w_i \rho_i(*_t) - F_t = e_t^+ - e_t^- & \forall t \in \{1, \ \dots, T\} \\
& -2 * x_i <= w_i <= 2 * x_i & \forall i \in \{1, \ \dots, p\} \\
& \sum_{i=1}^{p} x_i <= B & \\
& w_i \in [-2, 2] & \forall i \in \{1, \ \dots, p\} \\
& x_i \in \{0, \ 1\} & \forall i \in \{1, \ \dots, p\} \\
& e_t^+, \ e_t^- \geq 0 & \forall t \in \{1, \ \dots, T\}
\end{aligned}
$$

Step 2 (online 24/7 at OGE):
Solve a LP to forecast hourly supply and the demand based on these feature sets

$$
\begin{aligned}
\min \quad & \sum_{t=1}^{T} e_t^+ + e_t^- & \\
\text{s.t.} \quad & \sum_{i=1}^{p} w_i \rho_i(*_t) - F_t = e_t^+ - e_t^- & \forall t \in \{1, \ \dots, T\} \\
& w_i \in [-2, 2] & \forall i \in \{1, \ \dots, p\} \\
& e_t^+, \ e_t^- \geq 0 & \forall t \in \{1, \ \dots, T\}
\end{aligned}
$$

Features

Feature	Description
$f_1(d,h) = \begin{cases} m(d,h-1), & \text{if } h > 0 \\ m(d-1,23) & \text{otherwise} \end{cases}$	Prior hour
$f_2(d,h) = m(d-1,0)$	First hour yesterday
$f_3(d,h) = m(d-1,23)$	Last hour yesterday
$f_{4:10}(d,h) = m(d-(1,2,..,7),h)$	The same hour 1,2,..,7 days ago
$f_{11:12}(d,h) = m(d-1,0(h))/m(d-2,0(h))$	Ratio first(same) hour yesterday first(same) hour 2 days ago
$f_{13:14}(d,h) = m(d-1,0(h)) - m(d-2,0(h))$	Difference first(same) hour yesterday, first(same) hour 2 days ago
$f_{15:21}(d,h) = 1/24(\sum_{h\in H} m(d-(1,2,...,7),h))$	Mean flow 1,2,..,7 days ago
$f_{22}(d,h) = f_{15}(d,h)/f_{16}(d,h)$	Ratio mean flow yesterday, 2 days ago
$f_{23}(d,h) = f_{15}(d,h)/f_{21}(d,h)$	Ratio mean flow yesterday, 7 days ago
$f_{24}(d,h) = f_{15}(d,h)/(1/24(\sum_{h\in H} m(d-8,h)))$	Ratio mean flow yesterday, 8 days ago
$f_{25}(d,h) = f_{15}(d,h) - f_{16}(d,h)$	Difference mean flow yesterday, 2 days ago
$f_{26}(d,h) = f_{15}(d,h) - f_{21}(d,h)$	Difference mean flow yesterday, 7 days ago
$f_{27}(d,h) = f_{15}(d,h) - (1/24(\sum_{h\in H} m(d-8,h)))$	Difference mean flow yesterday, 8 days ago
$f_{28}(d,h) = \begin{cases} 0, & \text{if } h = 0 \\ m(d,0) & \text{otherwise} \end{cases}$	First hour today
$f_{29}(d,h) = \begin{cases} 0, & \text{if } h = 0 \\ 1/h\sum_{i=0}^{h-1}(d,i) & \text{otherwise} \end{cases}$	Mean flow today
$f_{30}(d,h) = t_d - t_{d-1}$	Difference mean temperature today and yesterday
$f_{31}(d,h) = \begin{cases} 1, & \text{if } day \in \{Saturday, Sunday\} \\ 0 & \text{otherwise} \end{cases}$	Weekend
$f_{32}(d,h) = \begin{cases} 1, & \text{if } day \in \{Friday, Saturday\} \\ 0 & \text{otherwise} \end{cases}$	Evening
$f_{33}(d,h) = 1$	Offset



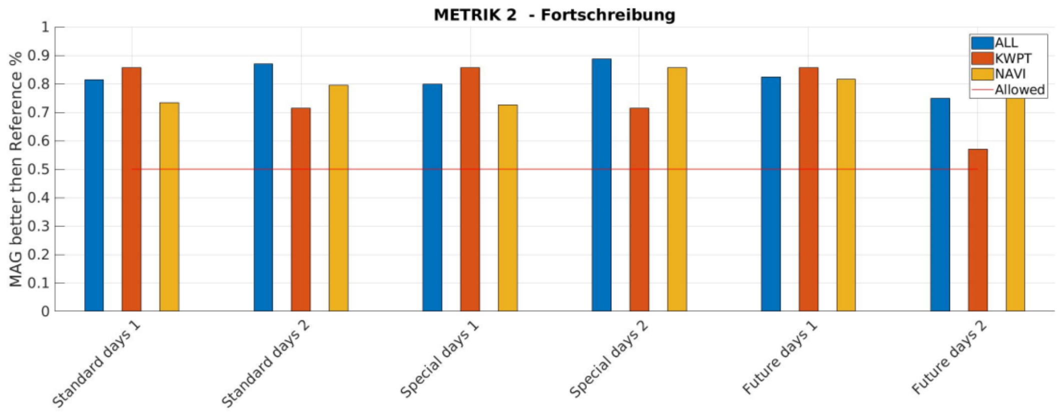
Forecast quality evaluation

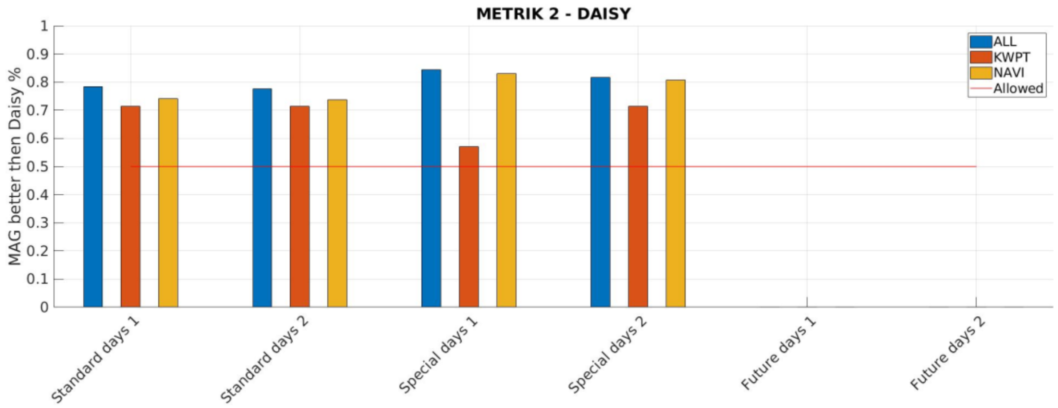
Standard days	Special days	Future days

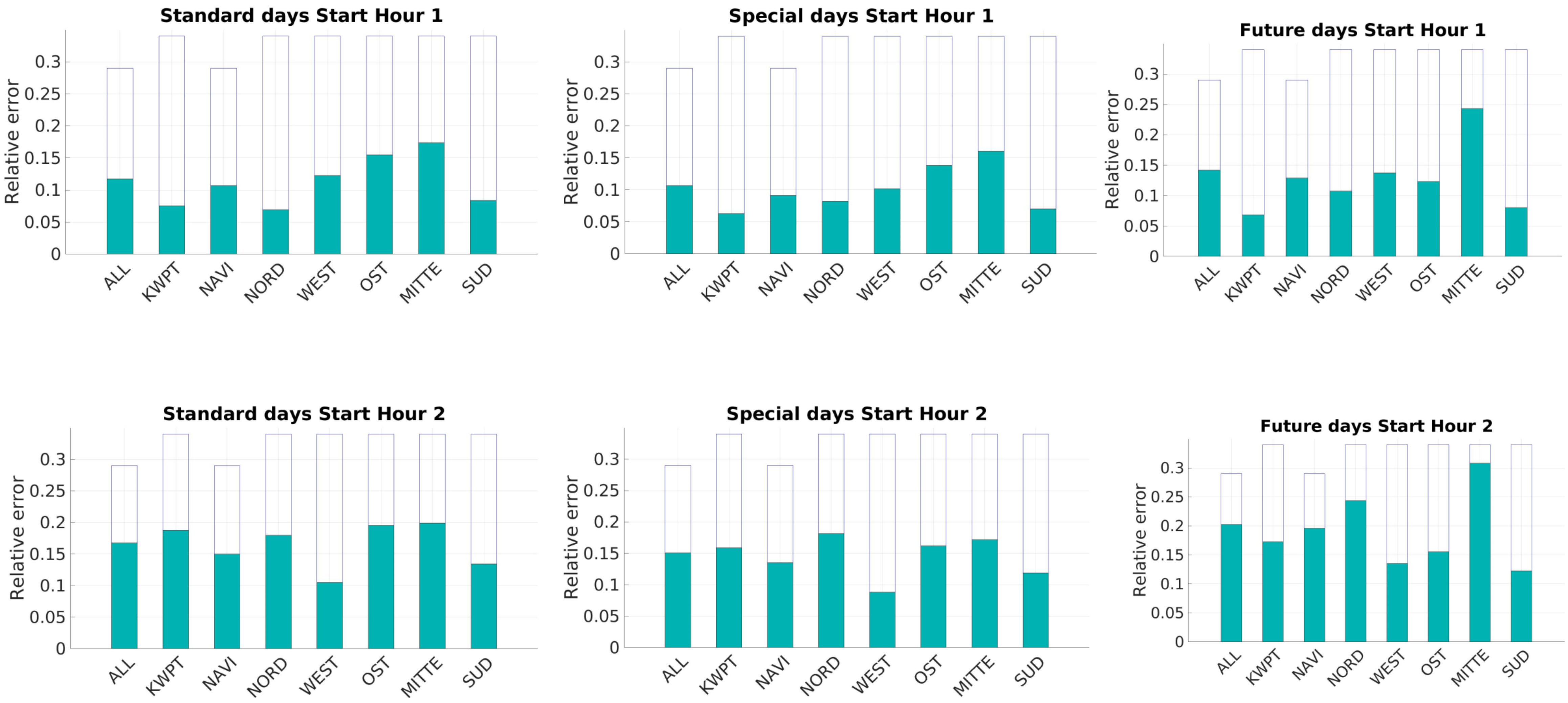
Start hour 1 – 02:30 (gas hour 21)	Start hour 2 – 09:30 (gas hour 4)

All boundary objects	KWPT boundary objects	NAVI boundary objects	Subnets

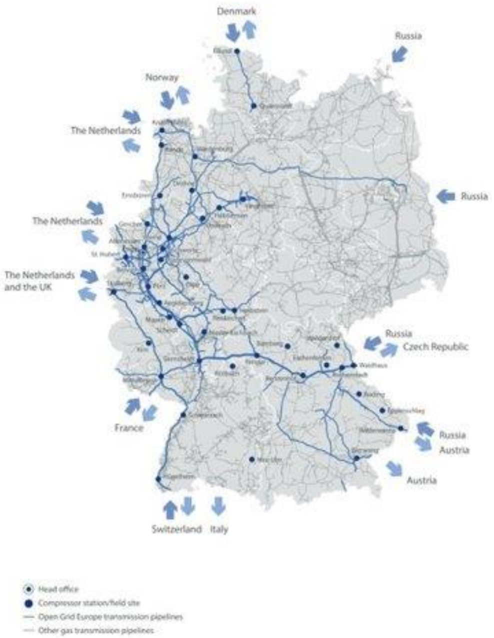
Daisy forecast	Reference forecast (Fortschreibung)

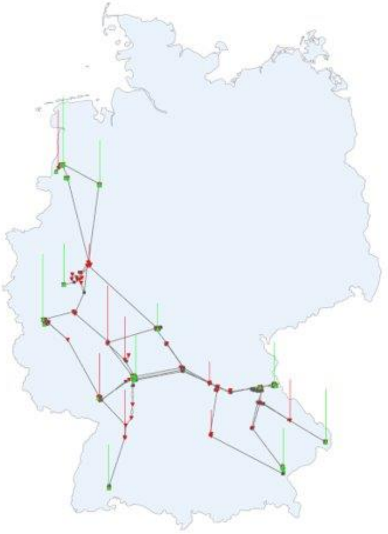


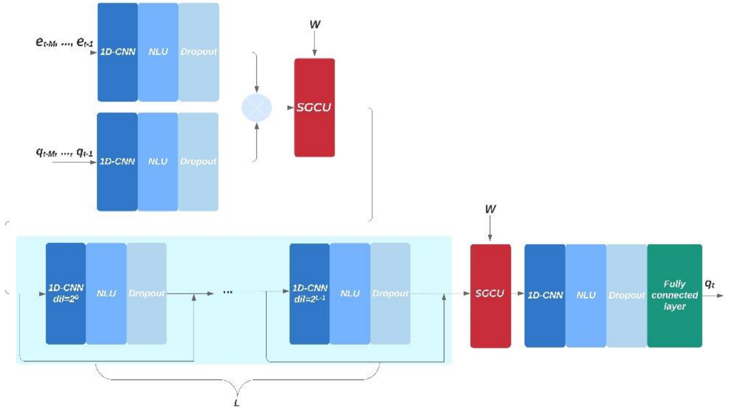


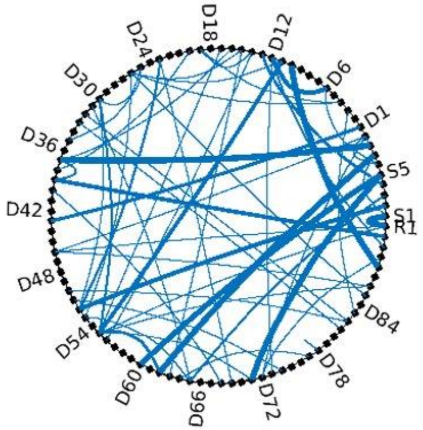

Standard days Start Hour 1
Special days Start Hour 1
Future days Start Hour 1
Standard days Start Hour 2
Special days Start Hour 2
Future days Start Hour 2
Relative error
0
0.05
0.1
0.15
0.2
0.25
0.3
ALL
KWPT
NAVI
NORD
WEST
OST
MITTE
SUD

Energy Network Time Series









$$\min_{w} \sum_{t=1}^{T} \sum_{i=1}^{N} (e^+_{t,i} + e^-_{t,i})$$

$$\text{s.t. } q_{i,t} - \sum_{j=1}^{N} \sum_{k=1}^{F_k} f_{t,j,k} \cdot w_{j,i,k} = e^+_{t,i} - e^-_{t,i}, \quad i = 1, ..., N, t = 1, ..., T$$

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{F_k} f_{t,j,k} \cdot w_{j,i,k} \leq B(t), \quad t = 1, ..., T$$

$$lb \cdot z_{i,j,k} \leq w_{i,j,k} \leq ub \cdot z_{i,j,k}, \quad i, j = 1, ..., N, k = 1, ..., Fk$$

$$\sum_{j=1}^{N} \sum_{k=1}^{F_k} z_{i,j,k} \leq L, \quad i = 1, ..., N$$

$$z_{i,j,k} \in \{0, 1\}, \quad i, j = 1, ..., N, k = 1, ..., F_k$$

$$e^+_{t,i}, e^-_{t,i} \geq 0, \quad i = 1, ..., N, t = 1, ..., T,$$

- **Additional constraints** arising from the application area
- **Resilience**
- **Unknown and/or changing network topology**

HNR-RB: Hierarchical Network Regression model with Relaxed Balance constraint

Objective

- A comprehensive understanding of the network dynamic
- Detects influential nodes in the network that demonstrate a strong effect on the future flows of other nodes

Challenges

- High dimensionality
- Unknown spatial-temporal dependence structure
- Constraints: demand and supply should be balanced

- Natural gas transmission network with N nodes
- Flow values $q_{t,i}$, $i = 1, ..., N, t = 1, ..., T$
- Features $f_{t,j,k}$
- Weights $w_{j,i,k}$

$$\hat{q}_{t,i} = \sum_{j=1}^{N} \sum_{k=1}^{F_k} f_{t,j,k} \cdot w_{j,i,k}, \quad i = 1, ..., N, t = 1, ..., T$$

$$e_{t,i} = q_{t,i} - \sum_{j=1}^{N} \sum_{k=1}^{F_k} f_{t,j,k} \cdot w_{j,i,k}, \quad i = 1, ..., N, t = 1, ..., T$$

HNR-RB: Idea

- Max number of selected features L
- Allowed flow disbalance $B(t)$
- Over- and under- estimation errors $e^+_{t,i}, e^-_{t,i}$
- Decision variables $z_{i,j,k}$

$$\min_w \sum_{t=1}^{T}\sum_{i=1}^{N}(e^+_{t,i} + e^-_{t,i})$$

$$\text{s.t. } q_{i,t} - \sum_{j=1}^{N}\sum_{k=1}^{F_k} f_{t,j,k}\cdot w_{j,i,k} = e^+_{t,i} - e^-_{t,i}, \quad i = 1,...,N, t = 1,...,T$$

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{F_k} f_{t,j,k}\cdot w_{j,i,k} \leq B(t), \quad t = 1,...,T$$

$$lb\cdot z_{i,j,k} \leq w_{i,j,k} \leq ub\cdot z_{i,j,k}, \quad i,j = 1,...,N, k = 1,...,Fk$$

$$\sum_{j=1}^{N}\sum_{k=1}^{F_k} z_{i,j,k} \leq L, \quad i = 1,...,N$$

$$z_{i,j,k} \in \{0,1\}, \quad i,j = 1,...,N, k = 1,...,F_k$$

$$e^+_{t,i}, e^-_{t,i} \geq 0, \quad i = 1,...,N, t = 1,...,T,$$

- $F_k \cdot N^2$ binary variables
- $F_k \cdot N^2 + 2\cdot T\cdot N$ continuous variables
- $2F_k \cdot N^2 + N\cdot T + T + N$ constraints

HNR-RB: Multistep model

$$\min_{w} \sum_{t=1}^{T} \sum_{i=1}^{N} (e_{t,i}^{+} + e_{t,i}^{-})$$

$$\text{s.t. } q_{i,t} - \sum_{j=1}^{N} \sum_{k=1}^{S} f_{t,j,k} \cdot w_{j,i,k} = e_{t,i}^{+} - e_{t,i}^{-}, \quad i = 1, ..., N, t = 1, ..., T$$

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{S} f_{t,j,k} \cdot w_{j,i,k} \leq B(t), \quad t = 1, ..., T$$

$$lb \cdot z_{i,j,k} \leq w_{i,j,k} \leq ub \cdot z_{i,j,k}, \quad i, j = 1, ..., N, k = 1, ..., S$$

$$\sum_{j=1}^{N} \sum_{k=1}^{S} z_{i,j,k} \leq L(S), \quad i = 1, ..., N$$

$$z_{i,j,k} = \tilde{z}_{i,j,k} \quad i, j = 1, ..., N, k = 1, ..., S-1$$

$$z_{i,j,k} \in \{0, 1\}, \quad i, j = 1, ..., N, k = 1, ..., S$$

$$e_{t,i}^{+}, e_{t,i}^{-} \geq 0, \quad i = 1, ..., N, t = 1, ..., T$$

$$\hat{q}_{i,t+h} = \sum_{j=1}^{N} \sum_{k=1}^{S} f_{t+h,j,k} \cdot w_{i,j,k}$$

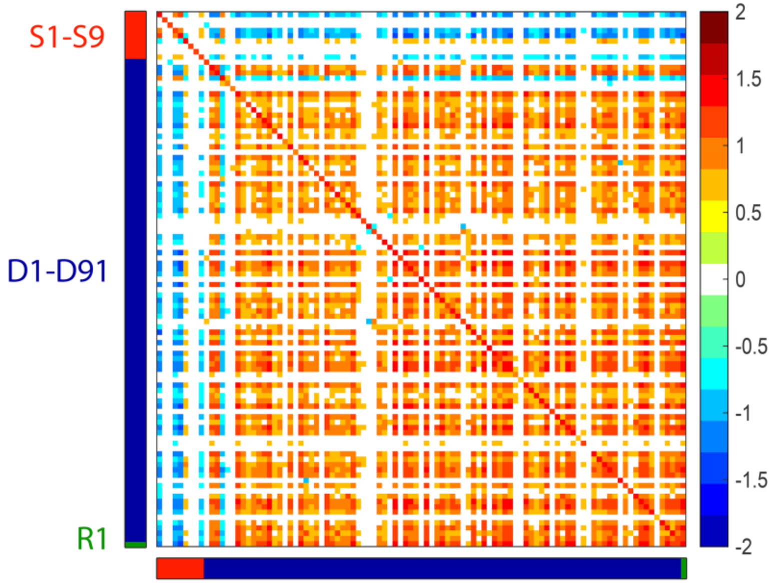
Data and experimental setting

- Out-of-sample forecast from 06:00 am for **1 to 24 hours ahead**
- Test period: **September to December 2022**

- **DATA:**
  - natural gas demand and supply flows (hourly time resolution)
  - network of **9 supply**, **91 demand** nodes + **artificial node**

- **MODEL setup:**
  - 3 steps with 3 features:
    - **previous hour**
    - **same hour yesterday.**
    - **mean flow yesterday**

  - Allowed disbalance 5% of hourly mean of total flow
  - Historical window 120 hours
  - Lower and upper bound are set to -2 and 2 respectively.



HNR-RB: Results

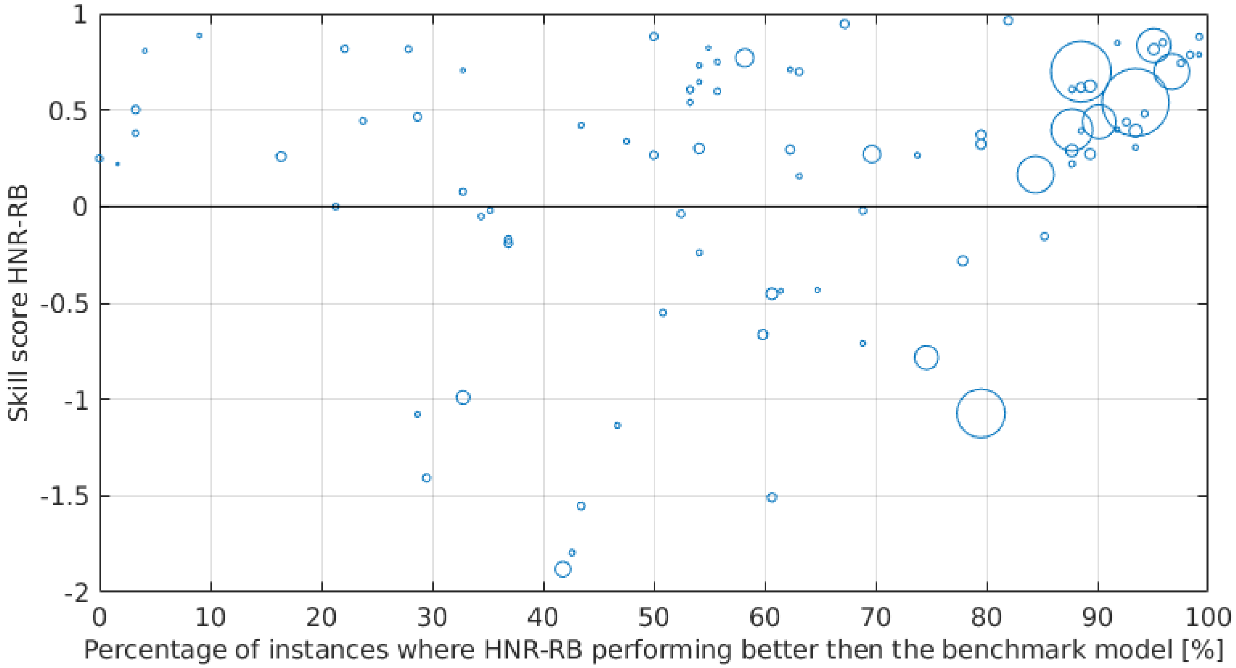
Benchmarks:

- ARIMA
- LSTM
- Network Autoregressive model with Balance constraint (NAC)[1]

Error metrics:

- **RMSE**
- **MAPE**
- **Skill Score:**

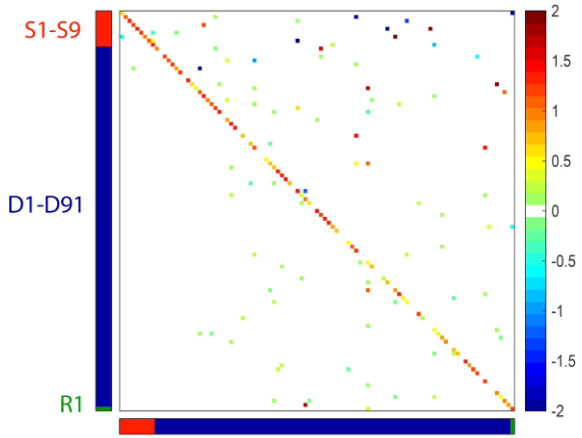
$$Skill_{HNR-RB} = 1 - \frac{MAPE_{HNR-RB}}{MAPE_{Benchmark}},$$



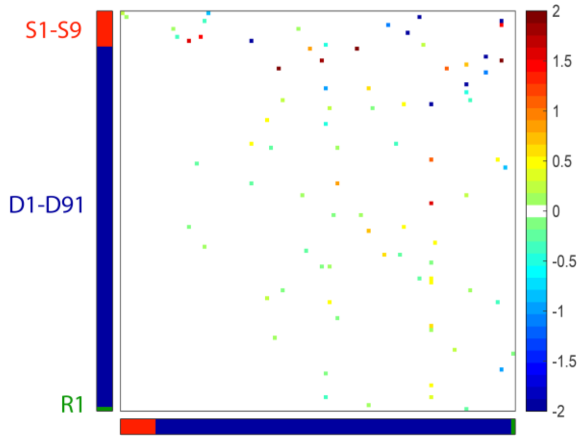
	RMSE				MAPE				Skill Score		
H	HNR-RB	ARIMA	LSTM	NAC	HNR-RB	ARIMA	LSTM	NAC	ARIMA	LSTM	NAC
1	117	701	3034	136	0.131	0.351	0.197	0.138	0.628	0.335	0.051
6	224	762	6662	988	0.160	0.322	0.210	0.219	0.503	0.238	0.269
12	384	954	8353	2346	0.188	0.361	0.287	0.349	0.479	0.345	0.461
24	581	1417	9810	5552	0.218	0.342	0.348	0.592	0.380	0.374	0.632

[1] Zakiyeva, N., Petkovic, M. (2022). Modeling and Forecasting Gas Network Flows with Multivariate Time Series and Mathematical Programming Approach. In: Operations Research Proceedings 2021.

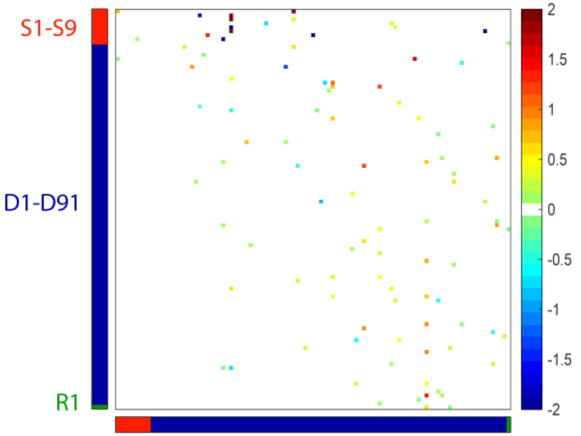
HNR-RB: Results



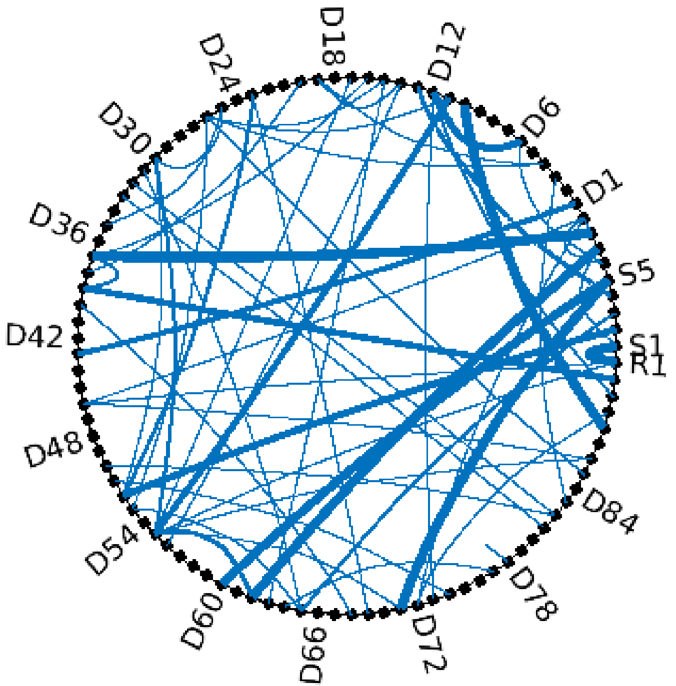
f1 previous hour



f2 same hour yesterday



f3 mean flow yesterday



Estimated network dependencies

The Leibniz Institute for Crystal Growth (IKZ) in Berlin-Adlershof is an international competence center for science & technology as well as service & transfer in the field of crystalline materials.

The R&D spectrum ranges from basic and applied research topics to pre-industrial research tasks.

Crystal Materials

Electronics	Optics	Scientific Instruments	Energy Storage:
• **Semiconductors:** Silicon crystals are fundamental in making **microchips** and **transistors.** • **Piezoelectric Devices**: Quartz crystals are used in watches, sensors, and oscillators	• **Lenses and Lasers:** Crystals like sapphire and ruby are used in high-precision optical instruments and lasers.	• **X-ray Crystallography**: Used to determine the structure of complex molecules like proteins.	• **Battery Materials**: Crystalline structures are essential in the development of high-capacity batteries.

Challanges of crystal growth

Large furnaces and grown crystals	• multicrystalline (mc) silicon ingot weighting up to 1600kg, single crystalline silicon ingot of 300 mm diameter, few meters long
Several simultaneously acting driving forces	• buoyancy, rotational force, Lorentz force, surface tension, viscous force etc.
Long transient processes	• mc Si crystal growth lasts ca. 1 week
Numerous temporally changing process parameters	• heating power, gas flow etc.
High operating temperature	• melting temperatures Tm(Si) = 1683 K
Contamination restrictions	
Can be hardly studied by costly plant trials	

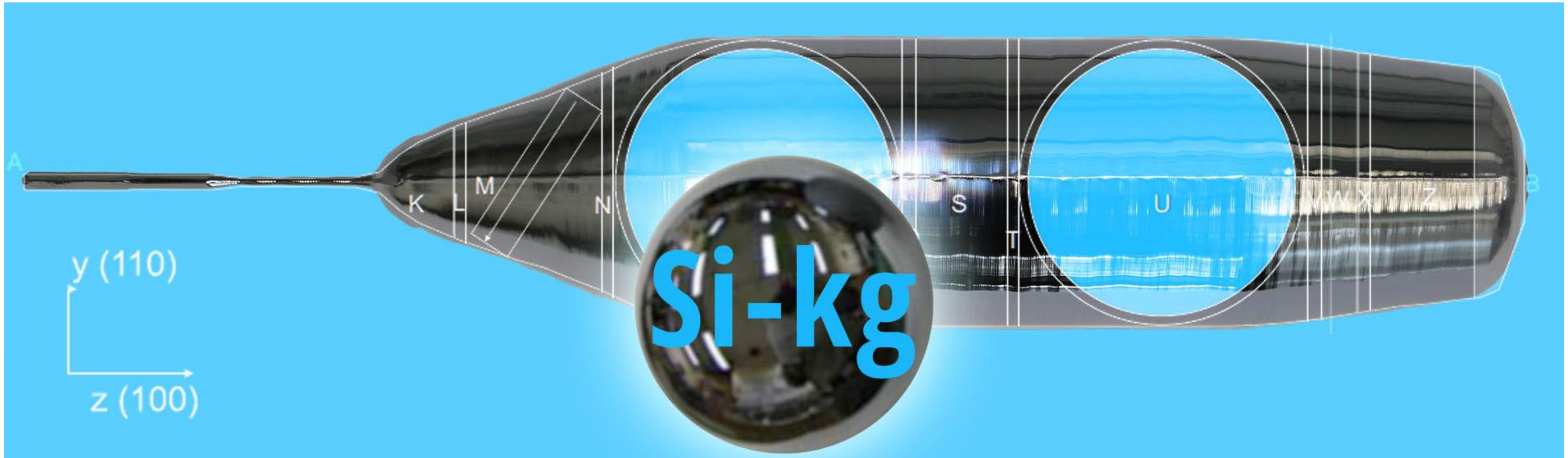
The real alternative for deeper insight into the complex transport phenomena taking place in bulk crystal growth and guidance/assistance in development of novel processes can be obtained only by CFD simulations and optimization

New Kilogram

The international system of units (SI system) was revolutionized over the last years by connecting all SI units to fundamental natural constants instead of artificial objects.
The last missing unit was the kilogram (kg) mass unit.
"Big K", the Platin-Iridium block in the French Museum Louvre in Paris defining the kg for around 130 years, had lost weight.

For this purpose, the PTB (Physikalisch-Technische Bundesanstalt) pursued an approach of "counting" atoms in a silicon sphere.
Within the scope of the international Avogadro Project and national Kg projects, the by the Float-Zone technique defect-free bulk crystals of highly enriched (up to 99.999 %) silicon-28.

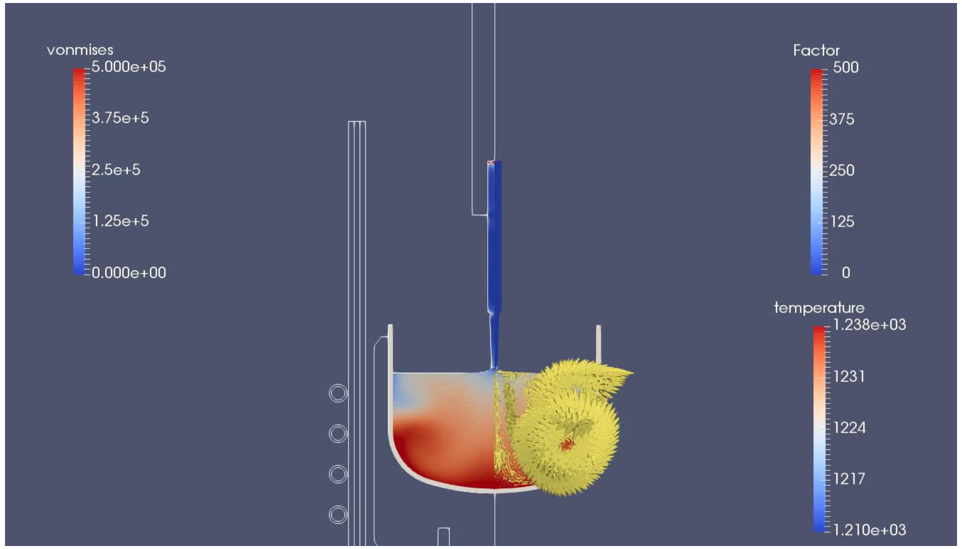
Only the IKZ with its expertise in crystal growth is able to achieve this precision in isotopically pure silicon crystal growth today.



Cz-Si crystal growth

The Czochralski method, is a crystal growth method used to obtain single crystals:

- semiconductors (e.g. silicon, germanium and gallium arsenide)
- salts
- synthetic gemstones



Miller, Wolfram, IKZ Berlin

Monocrystalline silicon (mono-Si) grown by the *Cz method* is the basic material in the production of integrated circuits and semiconductor devices.

Mono-Si is widely used in solar cells due to its nearly perfect crystal structure, offering the highest light-to-electricity conversion efficiency.

Advantages

High purity and uniformity of crystals.

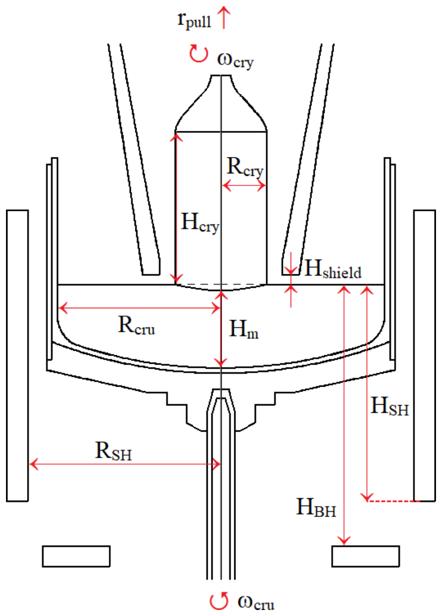
Can produce large-diameter crystals, reducing costs in mass production.

Challenges

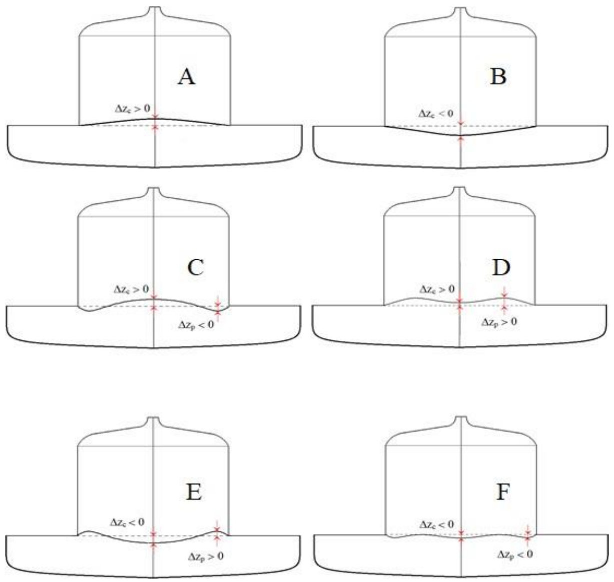
Equipment cost and energy-intensive process.

Risk of introducing defects during crystal growth.

Definition of variables



interface shape types



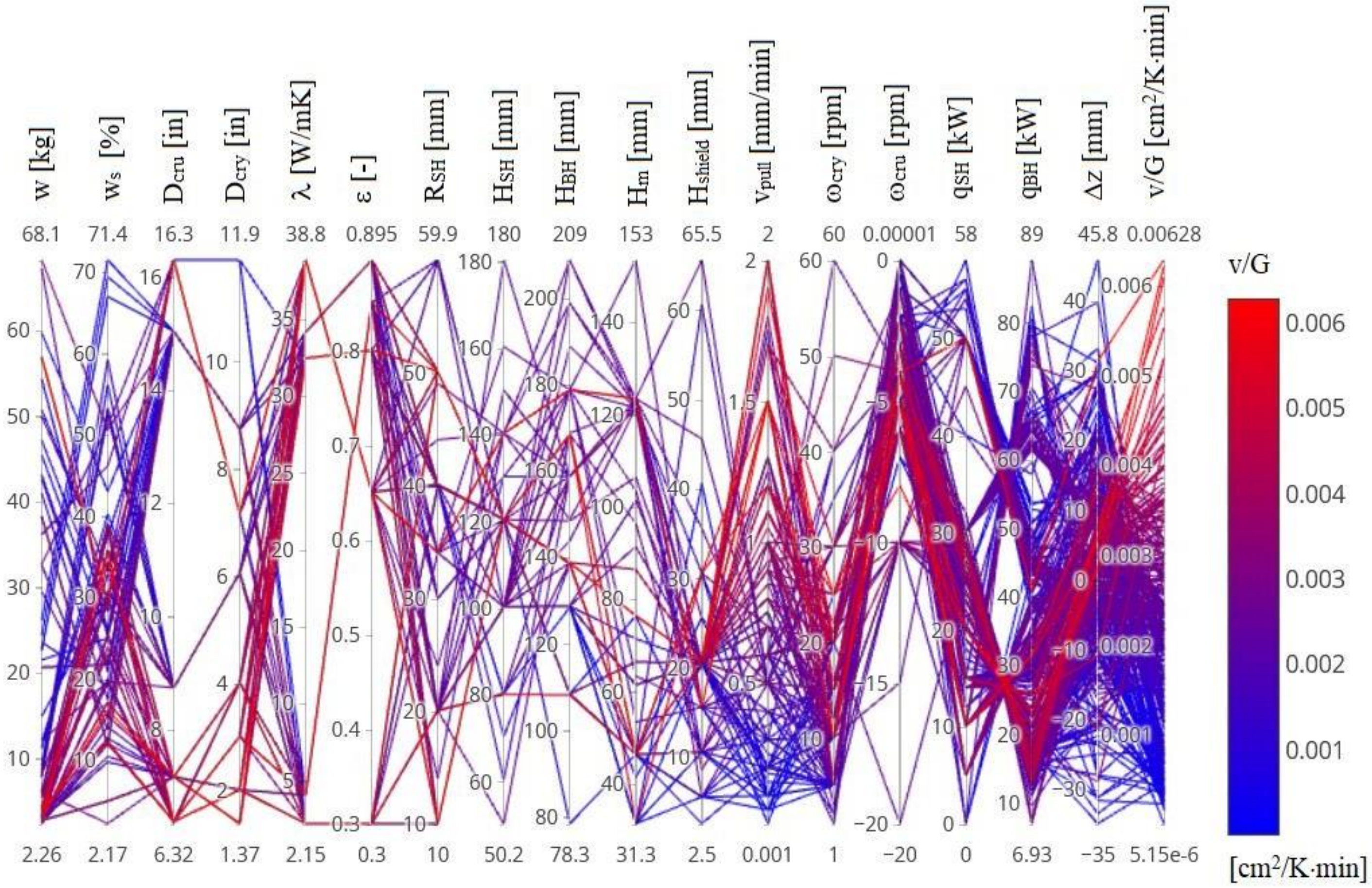
o1 Δz / Rcry [-]
o2 v/G /1.34·10cm /K·min [-]
o3 T_{min}-1685 K [K]

Variable	Furnace parameters	Unit	Variable	Process parameters	Unit
x1	Si weight	kg	x11	Pulling Rate	mm/min
x2	fraction of Si crystalized	%	x12	Crystal Rotation	rpm
x3	crucible diameter	mm	x13	Crucible Rotation	rpm
x4	crystal diameter	mm	x14	Side Heater power	KW
x5	Lcruc_SH	mm	x15	Bottom Heater power	KW
x6	Hcruc_SH	mm			
x7	Lcruc_BH	mm			
x8	distance between shield and melt surface	mm			
x9	radiation shield material	-			
x10	lambda shield	W/mK			

i_1 Re for crucible [-]
i_2 Re for crystal [-]
i_3 Gr [-]
i_4 St [-]
i_5 Si weight [kg]
i_6 % of Si solidified [%]
i_7 λ of the radiation shield [W/m·K]
i_8 ε of the radiation shield [-]
i_9 R_{SH} / R_{cry} [-]
i_{10} H_{SH} / H_m [-]
i_{11} H_{BH} / H_m [-]
i_{12} H_{shield} / H_{cry} [-]

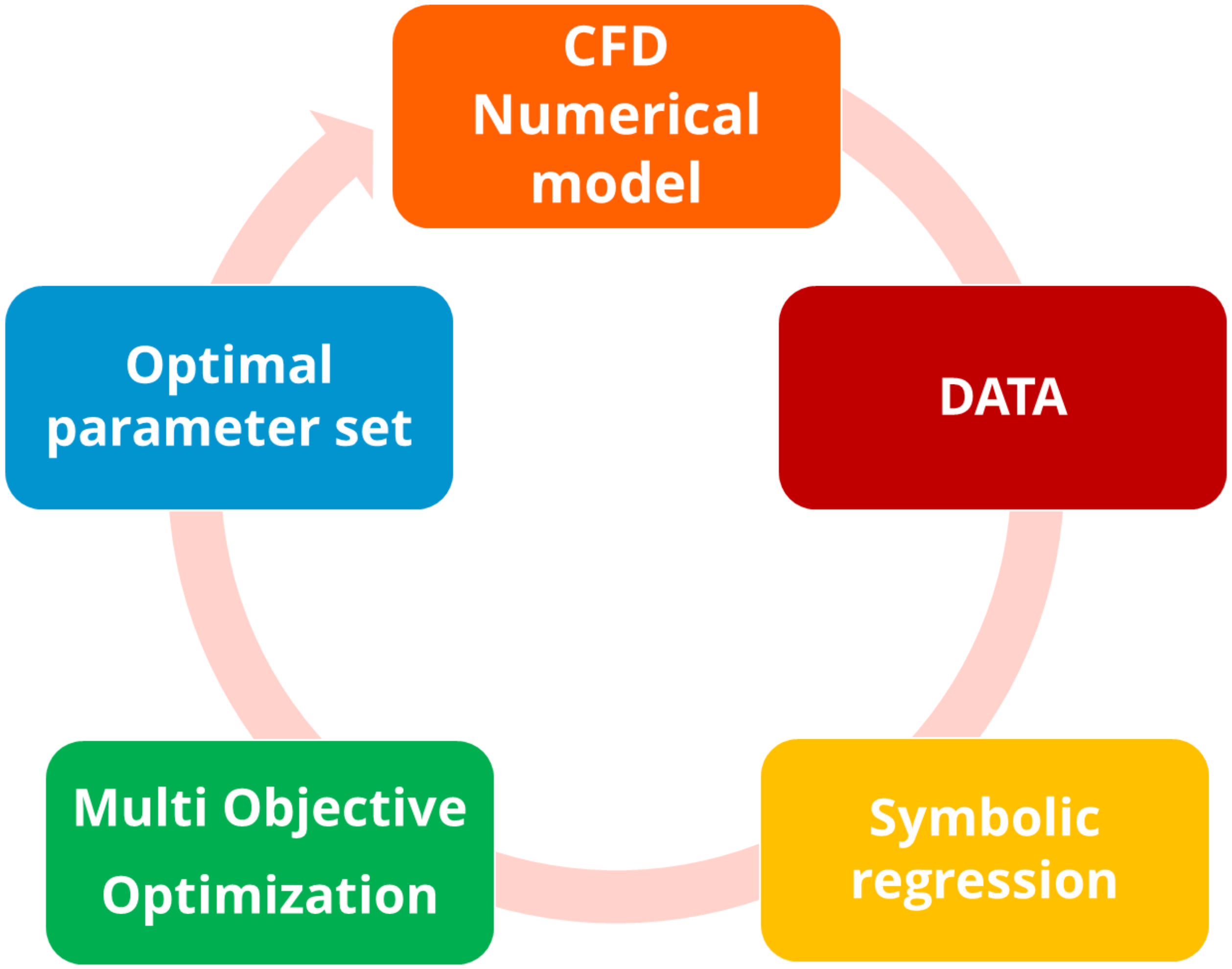
different growth recipes in different furnace geometries were simulated

Training data from CFD simulations for machine learning



N. Dropka, K. Böttcher, G.K. Chappa , M. Holena, Crystal Research and Technology(2024) 23003422300342.

SyMO: A hybrid approach for multi-objective optimisation of crystal growth processes



SyMO: Symbolic regression

$$\mathrm{Term}_1 = \sqrt{\mathrm{ReLU}(-25.228 + x_{15}) \cdot \left| \frac{(x_2 + x_{13})^3}{-30.825 + \frac{x_{14}^6}{\left(0.907 \cdot x_1^2 + x_5\right)^3} + x_{12}} \cdot x_7 \right| \cdot e^{x_{13}}} \quad (1.1)$$

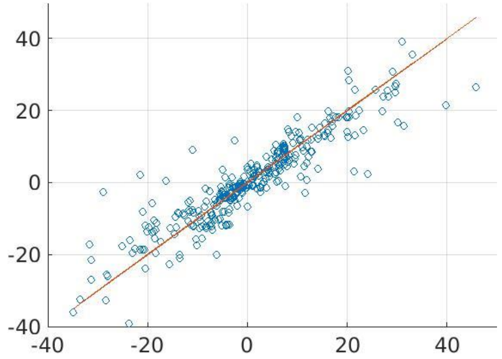
$$\mathrm{Term}_2 = \left(x_{11} + \frac{x_{12}}{\sqrt{\left(\frac{e^{x_{11}/x_9}}{x_1} + x_7\right)^3}} \right)^2 \quad (1.2)$$

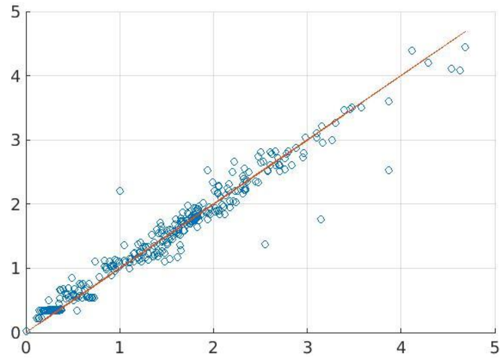
$$\Delta z = (x_{11} + x_4) \cdot \log\left((\mathrm{Term}_1 + \mathrm{Term}_2)^2\right) \cdot \sqrt{x_9} \quad (1.3)$$

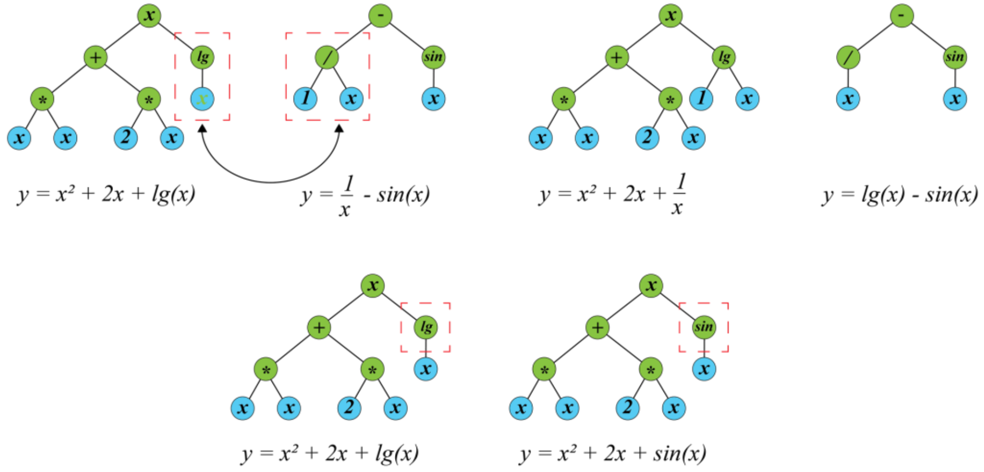
$$T_1 = \left[\mathrm{sigm}\left(0.612 \cdot |(-2.765 + x_2) \cdot x_1| \cdot \left(\frac{1}{x_4}\right)^3 \right) \right]^{x_{11}} \quad (2.1)$$

$$T_2 = \left(\mathrm{sigm}\left(\frac{x_{14}}{x_6 + \left(\frac{1}{x_{11}^2 \cdot 1.329}\right) / \left(\frac{x_{13}^2}{x_{12}} + \frac{1}{x_1 - 1.899}\right)^{2 \cdot x_9}} \right) \right)^2 \quad (2.2)$$

$$\Delta\Gamma = \left| \frac{\sqrt[3]{x_{11}^2 \cdot T_1 \cdot T_2}}{0.284} \right| \quad (2.3)$$

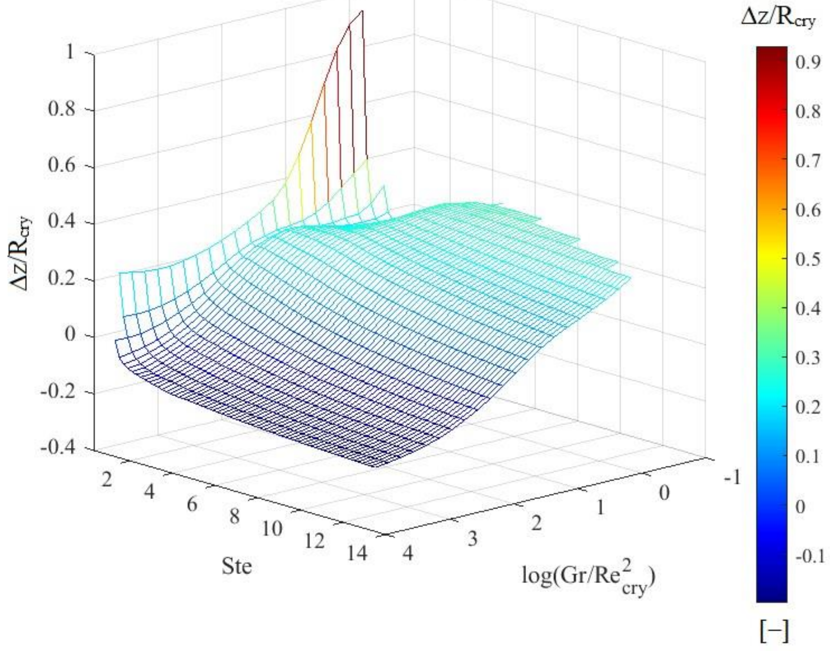




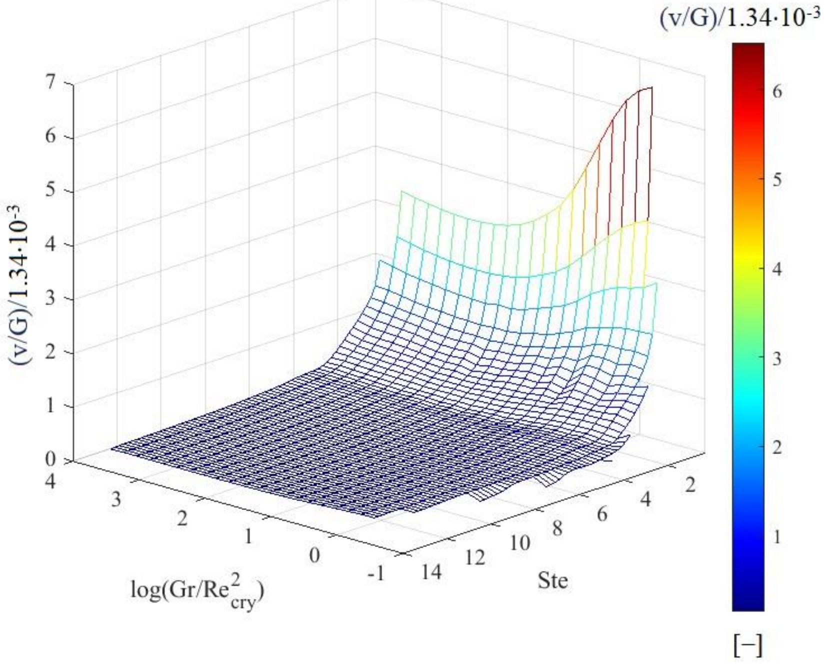


Impact of process parameters via Gr/Re_{cry}^2 and Ste numbers

interface deflection



v/G



Re_{cru}=-8.15·10³, w=32.6 kg, w_s=50.96%,
TiC radiation shield with λ=38.8W/mK and ε=0.65,
R_{SH}/R_{cru}=1.19, H_{BH}/H_m=3.35, H_{SH}/H_m=2.63 and H_{shield}/H_{cry}=0.06

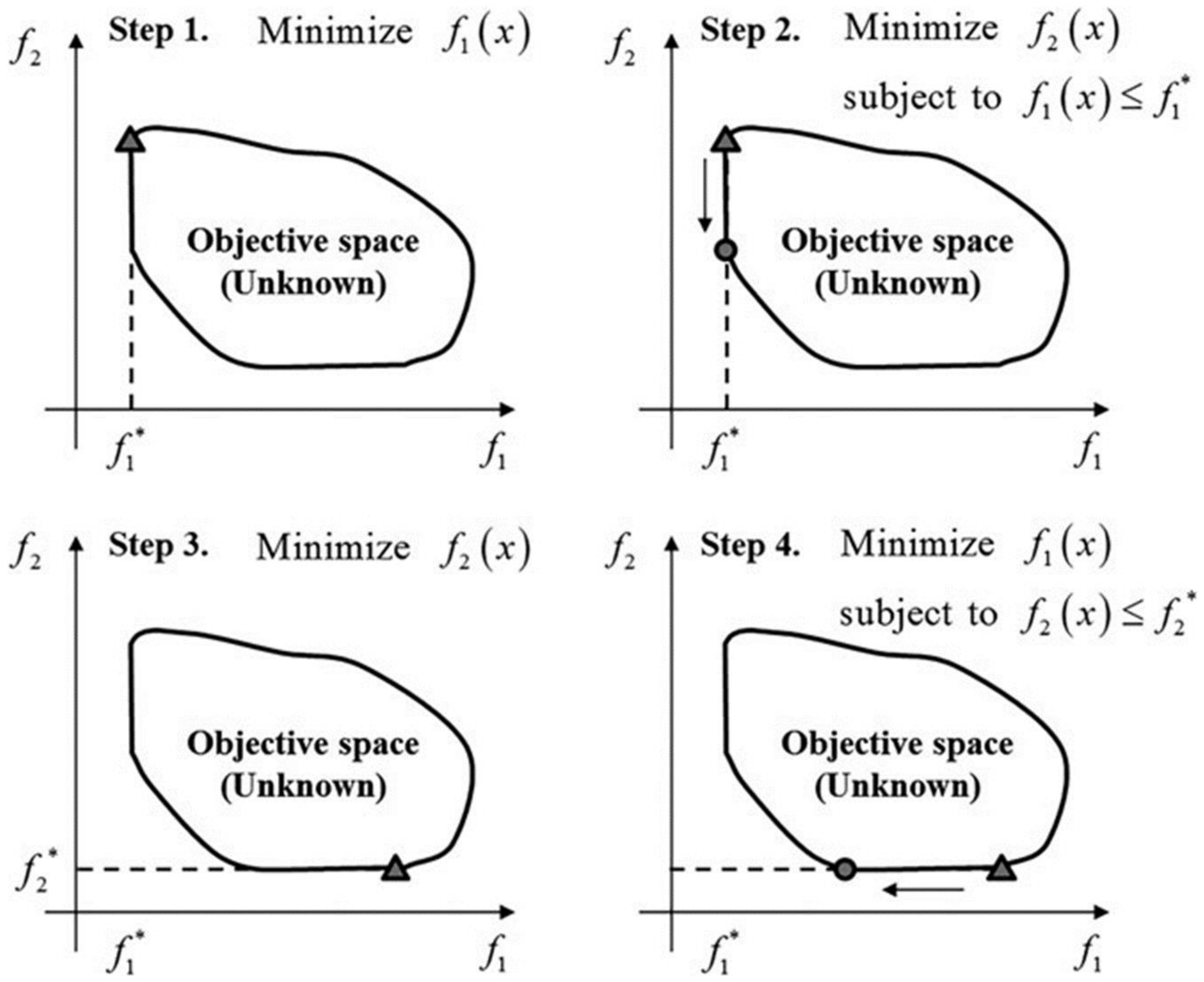
SyMO: Multiobjective Optimization

Let us assume an optimization problem with n decision variables $x_1, x_2, \cdots x_n$.
The standard form of a multi-objective optimization problem is:

$$\begin{aligned} \min \mathbf{F}(\mathbf{x}) &= (f_1(\mathbf{x}), f_2(\mathbf{x}) \cdots f_k(\mathbf{x})) \\ \text{s.t. } h_i(\mathbf{x}) &= 0 \quad i = 1, \cdots, l \\ g_j(\mathbf{x}) &\leq 0 \quad j = 1, \cdots, s \\ lb_m \leq x_m &\leq ub_m \quad m = 1, \cdots, n \end{aligned}$$

where:

- k: number of objectives
- l: number of equality constraints $h_i(\mathbf{x})$
- s: number of inequality constraints $g_i(\mathbf{x})$
- $\mathbf{F}(\mathbf{x})$: vector of objective functions $f_i(\mathbf{x})$.
- $\forall$ decision variable x_m is bounded $[lb_m, ub_m]$



Leksicographic MOO
https://doi.org/10.1002/nme.6013

SyMO: Results

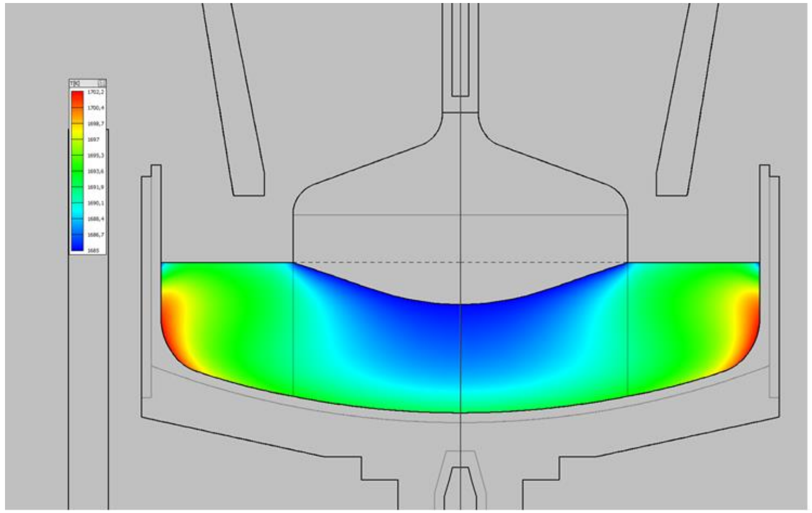
We employ hierarchical MOO model with three objectives:

- **Minimizing** Δz
- **Minimizing** $\Delta\Gamma$
- **Maximizing x11**, e.g. Maximizing crystal growth speed where Δz and $\Delta\Gamma$ are given as constraints from SR equations
- X1 - X10 fixed
- X11 – X15 optimized

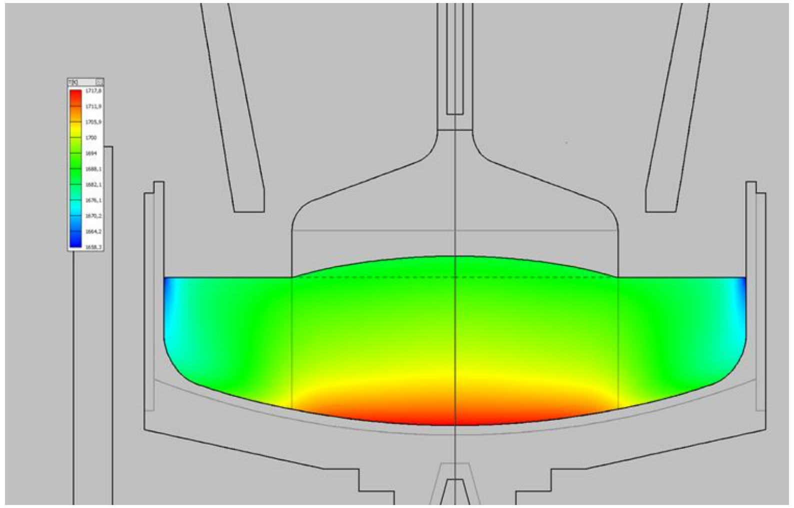
For fixed geometry parameters $x_1, \ldots, x_{10}$ we optimize process parameters to minimize Δz and $\Delta\Gamma$ while maximizing pulling rate x_{11}:

$$\min_{x_{11}, x_{12}, x_{13}, x_{14}, x_{15}} [\Delta z, \Delta\Gamma, -x_{11}]$$
$$\text{s.t. } \Delta z = \text{Equation 1.3}$$
$$\Delta\Gamma = \text{Equation 2.3}$$

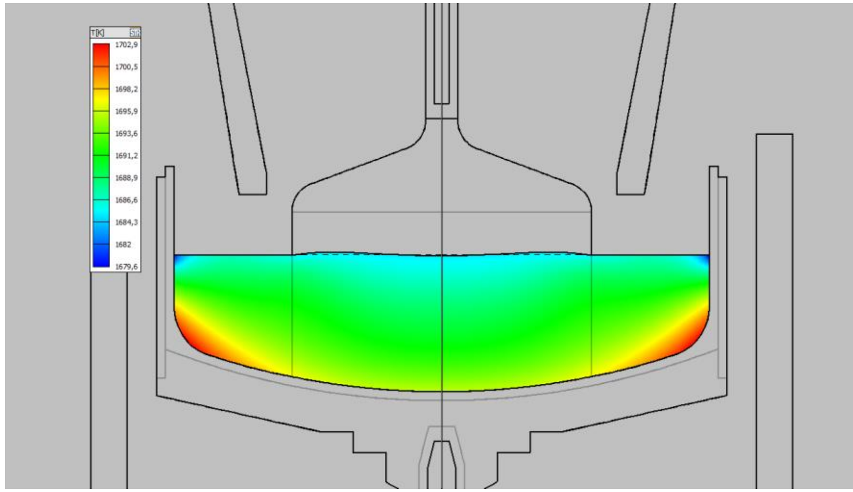
Additional constraint: $x_{14} + x_{14} \geq \mathrm{H}$



Original parameters



Model1: Optimal parameters



Model2: Optimal parameters

N. Dropka, M. Petkovic, K. Böttcher, M. Holena, Unraveling conditions for W-shaped interface and undercooled melts in Cz-Si growth: a smart approach, Journal of Crystal Growth, 2024, https://doi.org/10.1016/j.jcrysgro.2024.127897 ikz

Leibniz Association

Leibniz Forschungsverbund Berlin e.V.

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

MODAL FORSCHUNGSCAMPUS

From Energy Systems to Material Science: Optimization as a Common Denominator for a Sustainable Future

Questions?

CO@WORK, 25.09.2024, Berlin, Germany