Data Analytics and Optimization in Steel Industry

Lixin Tang

Key Laboratory of Data Analytics and Optimization for Smart Industry (Northeastern University), Ministry of Education, China

November 21 2019
About Northeastern University - basic information

- Founded in 1923, a key university directly affiliated to MOE
- 985, 211, Double First Class Project
- Over 46,000 students: 29931 undergraduates
  12166 master degree students
  3986 doctoral students
- 2688 faculty members: 538 professors
- 4 Campuses: Nanhu Campus, Hunnan Campus,
  Qinhuangdao Campus, East Campus
**H plan**: a interdisciplinary development plan involving two main disciplinary groups of Northeastern University.
Outline

- Research Background
- System Modeling and Optimization Method
- Production Scheduling
- Logistics Scheduling
- Energy Optimization
- Data Analytics
1. Research Background —— Steel is a Key Driver of the World’s Economy

Data Analytics and Optimization
1. Research Background —— China is the Largest Steel Producer

- China has been the largest steel producer in the world for the last twenty consecutive years.
- In 2018, China's steel output has reached 928 million tons, accounting for 51.3 percent of the world's steel output.
- Steel industry has been one of the pillar industries in China’s national economy.
1. Research Background — Steel Production Process

**Features:** continuous and discrete production, huge devices, high-temperature operation, mass consumption of energy and resource.
1. Research Background —— Challenges Faced by Steel Industry

- High Resource Consumption
- High CO₂ Emission
- High Energy Consumption
- High Inventory

Steel Production

Steelmaking
Logistics
Hot rolling
Cold rolling

Data Analytics and Optimization
1. Research Background — Analytics and Optimization in Steel Industry

- Production
- Logistics
- Steel Manufacturing Engineering
- Energy
- Data

Data Analytics and Optimization
Outline

Research Background

System Modeling and Optimization Method

Production Scheduling

Logistics Scheduling

Energy Optimization

Data Analytics
2. System Modeling and Optimization Method

- New Characteristics
  - Complex physical and chemical process
  - Large variety and low volume products
  - Complicated logistics structure

Complicated Production Process
- Iron
- Steel
- Slab
- Coil
- Rolling Process

Large Variety and Low Volume

Huge Chemical Equipment

Complicated Logistics Structure

Data Analytics and Optimization
Jean-Jacques Rousseau:

The art of musicians does not lie in depicting images directly, but in placing the mind in the emotions that these objects can create in the mind.
2. System Modeling and Optimization Method — System Modeling

Various Demand
- Physical performance
- Chemical performance
- Mechanical performance

Customer Demand

Engineering Object
- Production
- Logistics
- Energy
- Equipment

Time Management
- Steelmaking
- Hot rolling
- Cold rolling
- Logistics

Space Management

Various Products
- raw material
- semi-finished product
- finished product

Data Analytics and Optimization
2. System Modeling and Optimization Method — System Modeling

Data Analytics and Optimization

运筹优化 (OR)

人工智能 (AI)

DAO
A system modeling method combining mathematical optimization and data analytics is proposed. The mathematical integer programming model for production scheduling is established, and complementary part is carried out through data analytics.

\[
\text{Max } Z = \sum_{j \in P} \sum_{i \in O_j} F_j X_{ij} - \sum_{j \in P} \sum_{i \in O_j} (C_{ij} X_{ij} + \sum_{k \in O_j \cap R_i} C_{jk} Y_{kj}) \\
\text{s.t.} \\
\sum_{j \in O_i} X_{ij} \leq 1 \quad \text{for } i \in N \\
\sum_{i \in O_j} h_i X_{ij} \leq H \quad \text{for } j \in P \\
\sum_{k \in O_j \cap R_i} Y_{kj} = 1 \quad \text{for } j \in P \\
Y_{kj} \leq Y_{kl} \quad \text{for } j \in P, k \in O_j, i \in O_j \cap R_k \\
X_{ij} = \sum_{k \in O_j \cap R_i} Y_{kj} \quad \text{for } j \in P, i \in O_j \\
X_{ij} \in \{0,1\} \quad \text{for } j \in P, i \in O_j \\
Y_{kj} \in \{0,1\} \quad \text{for } j \in P, k \in O_j, i \in O_j \cap R_k
\]
The problem is transformed into the optimization combination of multiple batch schemes of jobs, and the Set-Packing model is established;

A batch scheme of jobs is defined as an element, which includes the combination of jobs;

The sub-problems are to describes the generation rules of batch schemes of jobs;

Effectively reduce the number of variables and constraints and improve the solving efficiency of the model.

Set-Packing modeling

2. System Modeling and Optimization Method — System Modeling

Space-time network flow modeling

- The space-time is discretized into grid and depicted based on network graph. Each node represents a location, each edge indicates a crane's move between two locations in a stage;
- The spatial location includes all the locations in the storage area and the entry, exit and initial location of the crane;
- The scheduling of task sequence is transformed into the allocation of crane movement in stages, and an event-based space-time network model is established.

2. System Modeling and Optimization Method — Optimization Method

- Optimal
  - Exact Algorithms

- Integer Programming
  - Large scale
  - Multi-objective
  - Dynamic

- Near-Optimal
  - Meta-Heuristics

- Production and logistics system
  - Iron-making
  - Steel-making
  - Hot rolling
  - Cold rolling
  - Slab yard
  - Coil Warehouse
  - Warehouse

- Production Scheduling
  - Logistics Scheduling

Data Analytics and Optimization
2. System Modeling and Optimization Method — Exact Algorithms

Benders Decomposition Algorithm

Various Valid Inequalities
\[ \sum_{j \in \mathcal{N} \setminus \{i\}} u_{ij} \neq \sum_{j \in \mathcal{N} \setminus \{i\}} u_{rj} \quad \sum_{(r,s) \in G_i} y_{irs} \leq 0 \]

Combinatorial Benders Cuts
\[
\text{MILP}_{CB} := \left\{ \begin{array}{c}
A_{\text{MILP}} \quad v(A_{\text{MILP}}) \leq UB - \varepsilon
\end{array} \right\}
\]

Variable Reduction
\[ v[MP^k(IR)] > UB \quad v[MP^k_{LP}(IR)] > UB \]

Structure

Improving lower bound

Accelerating convergence

Reducing search space

2. System Modeling and Optimization Method — Exact Algorithms

Outer Approximation Algorithm (OA)

**Multi-generation Cuts**

\[ \alpha \geq f(x^k) + h^T \cdot y + (\lambda^k)^T \cdot (g(x^k) + H \cdot y) + (\mu^k)^T \cdot (A \cdot x^k + E \cdot y - b) \quad k \in KFS \]

\[ (\lambda^k)^T \cdot (g(x^k) + H \cdot y) + (\mu^k)^T \cdot (A \cdot x^k + E \cdot y - b) \leq 0 \quad k \in KIS \]

**Partial Surrogate Cuts**

\[ (\lambda^k)^T \cdot [Hy + Dw + g(v^k)] - (\mu^k)^T A_2 (v - v^k) \leq 0 \]

**Hybrid Strategy of OA and GBD**

**Structure**

- NLP
- NLP
- NLP
- MI LP + MC

- Accelerating convergence
- Tightening lower bound
- Improving efficiency

2. System Modeling and Optimization Method — Exact Algorithms

**OA Algorithm with Quadratic Cuts**

- Integrate the strategies of scaled quadratic cuts with multi-generation cuts for Outer Approximation

---

**Performance**

![Graph showing the performance of different algorithms](image)
2. System Modeling and Optimization Method — MetaHeuristics

Hybrid Multi-objective Evolutionary Algorithm

- Incorporating the Concepts of Personal Best and Global Best
- Propagating Mechanism
- Multiple Crossover Operators to Update the Population
- Avoiding local optimum
- Increasing robust
- Improving diversity

Performance

<table>
<thead>
<tr>
<th>Problems</th>
<th>HMDOEA $\bar{x}_{HMDOE}$</th>
<th>ADRSS $\bar{x}_{ADRSS}$</th>
<th>SMPSO $\bar{x}_{SMPSO}$</th>
<th>NSGA-II $\bar{x}_{NSGA-II}$</th>
<th>SPEA2 $\bar{x}_{SPEA2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZDT1</td>
<td>1.572e-04 ±1.6e-05</td>
<td>1.852e-04 ±4.6e-05</td>
<td>1.170e-04 ±5.5e-05</td>
<td>2.203e-04 ±6.6e-05</td>
<td>2.198e-04 ±1.6e-05</td>
</tr>
<tr>
<td>ZDT2</td>
<td>7.568e-05 ±3.1e-05</td>
<td>1.022e-04 ±3.6e-05</td>
<td>5.012e-05 ±4.2e-05</td>
<td>1.598e-04 ±1.6e-05</td>
<td>1.718e-04 ±1.6e-05</td>
</tr>
<tr>
<td>ZDT3</td>
<td>7.012e-05 ±2.1e-05</td>
<td>1.670e-04 ±2.3e-05</td>
<td>7.555e-05 ±2.3e-05</td>
<td>1.852e-04 ±2.1e-05</td>
<td>1.114e-04 ±2.3e-05</td>
</tr>
<tr>
<td>ZDT4</td>
<td>1.502e-04 ±4.5e-05</td>
<td>4.599e-04 ±1.5e-05</td>
<td>1.507e-04 ±4.5e-05</td>
<td>4.079e-04 ±5.1e-05</td>
<td>5.770e-04 ±3.1e-05</td>
</tr>
<tr>
<td>ZDT6</td>
<td>1.374e-02 ±1.0e-02</td>
<td>4.236e-04 ±1.2e-05</td>
<td>4.630e-05 ±1.0e-05</td>
<td>6.773e-04 ±1.0e-05</td>
<td>1.325e-02 ±1.0e-02</td>
</tr>
<tr>
<td>Kursawe</td>
<td>1.407e-03 ±3.5e-04</td>
<td>1.411e-03 ±3.5e-04</td>
<td>1.754e-03 ±2.5e-04</td>
<td>1.667e-03 ±3.5e-04</td>
<td>1.416e-03 ±3.5e-04</td>
</tr>
<tr>
<td>Kita</td>
<td>1.973e-03 ±1.0e-03</td>
<td>4.479e-04 ±2.4e-02</td>
<td>1.768e+01 ±3.6e-02</td>
<td>7.517e-03 ±2.2e-02</td>
<td>4.746e-03 ±2.4e-02</td>
</tr>
<tr>
<td>Constr</td>
<td>4.292e-04 ±1.5e-04</td>
<td>4.349e-03 ±3.5e-03</td>
<td>2.703e-02 ±3.5e-02</td>
<td>4.630e-04 ±1.5e-04</td>
<td>4.931e-04 ±1.5e-04</td>
</tr>
<tr>
<td>DTLZ2</td>
<td>3.461e-01 ±1.0e-01</td>
<td>7.151e-04 ±1.3e-03</td>
<td>3.748e-03 ±3.0e-03</td>
<td>1.354e-03 ±3.0e-03</td>
<td>1.277e-03 ±3.0e-03</td>
</tr>
<tr>
<td>DTLZ3</td>
<td>5.843e-01 ±1.1e-01</td>
<td>7.216e-04 ±1.3e-03</td>
<td>4.599e-03 ±3.6e-03</td>
<td>8.271e-01 ±1.1e-01</td>
<td>1.594e-00 ±1.1e-01</td>
</tr>
<tr>
<td>DTLZ4</td>
<td>4.705e-03 ±2.4e-03</td>
<td>4.853e-03 ±4.1e-03</td>
<td>5.707e-03 ±4.1e-03</td>
<td>5.016e-03 ±2.4e-03</td>
<td>4.820e-03 ±2.4e-03</td>
</tr>
<tr>
<td>DTLZ5</td>
<td>3.231e-04 ±4.6e-05</td>
<td>2.692e-04 ±3.5e-05</td>
<td>2.304e-04 ±3.0e-04</td>
<td>4.510e-03 ±4.6e-05</td>
<td>3.066e-04 ±4.6e-05</td>
</tr>
<tr>
<td>DTLZ6</td>
<td>4.807e-04 ±3.5e-04</td>
<td>8.250e-02 ±2.3e-03</td>
<td>4.816e-03 ±3.4e-03</td>
<td>1.316e-01 ±2.3e-02</td>
<td>1.234e-01 ±2.3e-02</td>
</tr>
<tr>
<td>DTLZ7</td>
<td>3.778e-03 ±1.3e-03</td>
<td>1.591e-03 ±1.3e-03</td>
<td>5.226e-03 ±1.3e-03</td>
<td>3.417e-03 ±1.3e-03</td>
<td>3.738e-03 ±1.3e-03</td>
</tr>
<tr>
<td>Vicenet</td>
<td>9.930e-03 ±2.1e-03</td>
<td>1.128e-02 ±2.6e-03</td>
<td>1.104e+02 ±2.6e-03</td>
<td>1.128e-02 ±2.6e-03</td>
<td>1.316e-02 ±2.6e-03</td>
</tr>
<tr>
<td>Vicen44</td>
<td>1.795e-03 ±3.0e-04</td>
<td>1.568e-03 ±1.3e-03</td>
<td>2.116e-01 ±2.6e-03</td>
<td>2.274e-03 ±2.6e-03</td>
<td>2.752e-03 ±2.6e-03</td>
</tr>
<tr>
<td>LZ09_F6</td>
<td>9.915e-02 ±2.4e-02</td>
<td>1.256e+00 ±1.6e-02</td>
<td>3.802e+00 ±2.5e-02</td>
<td>1.165e+00 ±2.6e-02</td>
<td>2.715e-01 ±2.4e-02</td>
</tr>
<tr>
<td>Binh4</td>
<td>5.696e-03 ±5.8e-04</td>
<td>6.171e-02 ±3.8e-03</td>
<td>1.142e+01 ±5.1e-03</td>
<td>6.116e-02 ±3.8e-03</td>
<td>6.204e-02 ±5.8e-04</td>
</tr>
<tr>
<td>Tamaki</td>
<td>3.716e-03 ±1.0e-03</td>
<td>3.848e-03 ±1.0e-03</td>
<td>7.333e-01 ±3.0e-03</td>
<td>3.760e-03 ±1.0e-03</td>
<td>2.531e-03 ±1.0e-03</td>
</tr>
</tbody>
</table>

2. System Modeling and Optimization Method — MetaHeuristics

**Improved Differential Evolution Algorithm for Dynamic Scheduling**

**Incremental Mechanism for Initial Population Generation**

**Real-coded Matrix Representation**

\[
A = \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{15} \\
a_{21} & a_{22} & \cdots & a_{25} \\
\vdots & \vdots & \ddots & \vdots \\
a_{N1} & a_{N2} & \cdots & a_{NS}
\end{pmatrix}
\]

**Randomly Mutation Operator**

\[
v_{i,g} = x_{i,g} + F(x_{r1,g} - x_{r2,g}) + F(x_{r3,g} - \hat{x}_{r4,g}) + F(x_{r5,g} - x_{r6,g})
\]

**Performance**

- Improving efficiency
- Avoiding invalid solutions
- Expanding search space

Algorithm has a fast convergence speed

---

Differential Evolution with An Individual–Dependent Mechanism

Individual–Dependent Parameters Setting

\[ F_i = \text{randn}(\frac{i}{NP}, 0.1) \quad CR_i = \text{randn}(\frac{i}{NP}, 0.1) \]

Individual–Dependent Mutation Operator

\[ DI = \frac{1}{N} \sqrt{\sum_{i=1}^{N} \left( x_i - \frac{1}{N} \sum_{j=1}^{N} x_j \right)^2} \quad DF = \frac{1}{N} \sqrt{\sum_{i=1}^{N} \left( f(x_i) - \frac{1}{N} \sum_{j=1}^{N} f(x_j) \right)^2} \]

Perturbations with Small Probability

\[ d = L + \text{rand}(0, 1) \ast (U - L) \]

Self-adaptive allocation

Self-adaptive selection

Global search

Experimental demonstrate the algorithm’s outstanding performance

Outline

- Research Background
- Optimization Modeling and Method
- Production Scheduling
- Logistics Scheduling
- Energy Optimization
- Data Analytics
3. Production Scheduling — Steel Production

Production: Ironmaking/ Steelmaking/ Hot rolling/ Cold Rolling

Data Analytics and Optimization
3. Production Scheduling — Steelmaking Stage

Charge Batching

Order 1 → Charge
Order 2 → Charge

Cast Batching

Steel making → Continuous Casting

Data Analytics and Optimization
3. Production Scheduling —— Charge Batching of Steelmaking

3. Production Scheduling —— Cast Batching of Steelmaking

**Objective**

- Maximize tundish utilization
- Minimize total grade switch and width switch cost

**Decisions**

- Batch and sequence charges to form casts for the given tundishes
- Select a casting width for each charge in a cast

**Constraints**

- Grade switch constraint
- Width switch constraint
- Lifespan of tundish

---

3. Production Scheduling —— Steelmaking Scheduling

Steel making

Continuous Casting (CC)

Charge

Cast

ladle

tundish

CF-1
CF-2

RF-1
RF-2

CC-1
CC-2

Charge 1

Charge 2

Charge 3

waiting time

cast

time

Just-in-time idea

Solve machine conflicts in (SCC) production scheduling based on JIT idea

Four-level scheduling

- Level 1: cast sequences on the casters
- Level 2: sub-scheduling
- Level 3: rough scheduling
- Level 4: elimination of machine conflicts

Beneficial effects

- Improve productivity of large devices
- Shorten waiting-time between operations
- Cut down production costs

Traditional batching machines are mainly divided into three types: (1) burn-in (2) fixed batch (3) serial batching.

A new kind of batch scheduling

We analyze the semi-continuous batch scheduling problem, and present optimal algorithm.

Characteristics of Semi-Continuous Batching Scheduling

- Classical Batching Machine Scheduling
  - Handle several jobs simultaneously
  - The same batch begin processing at the same time
  - The same completion time

- The new Semi-Continuous Batching Machine Scheduling
  - Enter and leave the machine one by one
  - Respective start time
  - Respective completion time

The heating process of Tube-billets in heating furnace.

3. Production Scheduling —— Hot Rolling Scheduling

**Decisions**

Sequence of adjacent jobs to be processed

**Objective**

Minimize the total changeover costs

Minimize $\sum_{i=1}^{N+M} \sum_{j=1}^{N+M} C_{ij} X_{ij}$

Subject to $\sum_{i=1}^{N+M} X_{ij} = 1, \quad j \in \{1, 2, ..., N+M\}$

$\sum_{j=1}^{N+M} X_{ij} = 1, \quad i \in \{1, 2, ..., N+M\}$

$\sum_{i \in S} \sum_{j \in S \setminus \{i\}} X_{ij} \leq |S| - 1, \quad S \subset \{1, ..., N+M\}, \quad 2 \leq |S| \leq N+M - 2$

3. Production Scheduling — Cold Rolling Scheduling

- Maximize Reward
- Minimize Mismatching cost
- Equipment Constraints
- Matching Constraints

Form batches for each empty furnace

Select a median coil for each batch

Outline

- Research Background
- Optimization Modeling and Method
- Production Scheduling
- Logistics Scheduling
- Energy Optimization
- Data Analytics
4. Logistics Scheduling — Logistics in Steel Plant

 Logistics: Loading(Un)/ Transportation/ Shuffling/ Storage/ Stowage

Data Analytics and Optimization
4. Logistics Scheduling —— Shuffling

**Shuffling Problems in Steel Plants**

Assign a storage slot for each shuffled item during retrieving all target items in the given sequence.

**Decisions**

Suitable storage positions for shuffled items.

**Objective**

Minimize shuffling and crane traveling.

**For general case**

Greedy heuristic.

**For special cases**

Polynomial algorithms (optimal solutions).

---

4. Logistics Scheduling — Reshuffling and Stacking

- For statistic and dynamic reshuffling problem, an improved mathematical formulation and a simulation model are established, respectively;
- Five polynomial time heuristics and their extended versions are proposed and analyzed theoretically;
- The proposed heuristic outperforms existing methods.

4. Logistics Scheduling —— Crane Scheduling

Crane scheduling problem
Determines the transportation sequence for all demanded coils and shuffled position for each blocking coil.

Decisions
Retrieval sequence of the target coils and shuffled positions for blocking coils

Objective
Minimize the time by which the retrieval of all target coils is completed

For general case
Heuristic algorithm & worst-case analysis

For special cases
Polynomial algorithms (optimal solutions)

4. Logistics Scheduling —— Ship Stowage Planning

Minimize the moment imbalance

Minimize the shuffling

Minimize the dispersion of coils for the same destination

Structural constraints

Weight restriction constraints

Operational constraints

4. Logistics Scheduling —— Coordinated Scheduling

Integrated production, inventory, and delivery problem

- Case 1: order delivery is splittable
- Case 2: order delivery is non-splittable

Most General Problems
- No approximation algorithms exist with a constant worst-case ratio
- Combined column generation and tabu search heuristic algorithms

Integrated Production & Two-Stage Distribution Scheduling

- Obtain a joint schedule of job processing at the plant and two-stage shipping
- Optimize a performance measure that takes into account both delivery timeliness and total transportation costs

Offline problems involving a single production line
- Optimal dynamic programming algorithms

Offline problems involving multiple production lines
- Fast heuristics
- Worst-case & asymptotic performance

Online problems
- Online algorithms
- Competitive ratios analytics
5. Energy Optimization

Goal
- Determine energy purchased and secondary energy generated
- Minimize total energy cost

Challenges
- Frequent change of production environment
- Unstable ratio of energy consumption and regeneration

Energy supply
- Electricity
- Heavy oil
- Oxygen
- Water
- BFG
- COG
- LDG
- Nitrogen
- Pulverize coal
- Natural gas

Energy regeneration
- BFG
- COG
- LDG
- Steam

Iron making
Steelmaking
Continuous casting
Hot rolling

Data Analytics and Optimization
5. Energy Optimization —— Dynamic Energy Allocation

**Objectives**
- Minimizing emission
- Minimizing cost
- Maximizing income

**Constraints**
- Production
- Balance
- Demand
- Price

**ADP Algorithm**
Accomplish dynamic energy allocation

- The proposed energy allocation method shows obvious superiority in effectiveness and stability than static method.
5. Energy Optimization —— Comprehensive Scheduling of Gas System

**Comprehensive scheduling of gas system**

- **Determine**: mixing plan of BFG, COG, LDG
- **Goal**: maximize value flow by controlling heat value

---

**Minimizing emission**

**Objective**: 
\[
\text{Min } Z = \sum_{j} \sum_{t} c_j F_{jt} + c^Q \sum_i \sum_{t} Q_{iti}
\]

\[
\alpha_{il2} x_{tij1} = \alpha_{il1} x_{tij2} \sum_{l} x_{ijlt} \leq \sum_{l} e_{ilt} A_{ilj}
\]

\[
\alpha_{il3} x_{tij1} = \alpha_{il1} x_{tij3}
\]

\[
\alpha_{il3} x_{tij2} = \alpha_{il2} x_{tij3} \sum_{l} e_{ilt} = 1
\]

**Gas mixing ratio constraints**

---

**Minimizing supply shortage**

\[
\sigma_{it} = \Omega_{it} - \sum_{j} \sum_{l} (b_{j} \cdot x_{ijlt})
\]

\[
Q_{it} = \text{Max} \left\{ 0, \sigma_{it} \right\}
\]

\[
O_{ijt} = f_{ijt} \cdot S_{it}
\]

**Capacity & demand constraints**

---

**Pipeline balance, flow, holding constraints**

\[
\delta_{ij} = \sum_{i} O_{ijt} + (H_{j,t-1} - H_{j}^0) - \sum_{i} \sum_{l} x_{ijlt}
\]

\[
H_{jt} = \text{Min} \left\{ \max \left\{ \delta_{jt} + H_{j}^0, H_{j}^0 \right\}, H_{j}^1 \right\}
\]

\[
H_{j}^0 \leq H_{jt} \leq H_{j}^1
\]

\[
\phi_{j}^0 \leq \sum_{t} O_{ijt} \leq \phi_{j}^1
\]

\[
\phi_{j}^0 \leq \sum_{t} x_{ijt} \leq \phi_{j}^1
\]

**Data Analytics and Optimization**
5. Energy Optimization —— Steam Scheduling

Objectives

- Maximize electricity generation upon demand

\[ z = \max \sum_i \sum_j \left( u_i + v_i x_{t_i,j} + w_i R_{t_i} \right) \]

Supply capacity constraints

\[ a_i^0 < \sum_{j=1}^i x_{ij} < a_i^1, \ b_i^0 \leq x_{ij} < b_i^1, \ r_i^0 \leq R_i \leq \min \left( x_{a_i}, r_i^1 \right), \ q_i^0 \leq Q_i \leq \min \left( x_{a_i}, q_i^1 \right) \]

\[ x_{ij} = \min \left\{ a_i^0, \max \left( a_i^1, S_j^D - \sum_{i \in I} \sum_{j \in J} \left( x_{i,j} + R_i + Q_i \right) \right) \right\} \]

Fluctuation, safe flow constraints

\[ F_i^D = \max \left( 0, \sum_i \sum_{j \in J} \left( x_{ij} + R_i + Q_i - e_i \right) \right) \quad F_i^Z = \max \left( 0, \sum_i x_{ij} - e_i \right) \]

\[ \sum_i \sum_{j} \left( x_{ij} + R_i + Q_i \right) - \sum_i \sum_{j \in J} \left( x_{t-1,i} + R_{t-1,i} + Q_{t-1,i} \right) \leq \delta^D \]

Steam demand constraints

\[ \eta^Z \sum_i x_{ij} > S_i^Z \quad \eta^D \sum_i \sum_{j} \left( x_{ij} + R_i + Q_i \right) > S_i^D \]
5. Energy Optimization —— Oxygen Scheduling

**Task**

Dynamically balance and optimize the oxygen system

**Supply Modes**

- Supplied by oxygen generator
- Supplied by liquid oxygen system

**Minimize operating cost of oxygen system**

\[
Z = \sum_{i} \sum_{t} \left( c_i \cdot F_{ti} + c_i^A \cdot A_{ti} + c_i^Y \cdot Y_{ti} + \frac{1}{2} \gamma_{ti} \cdot c_i \cdot 0.7B_i \right)
\]

**Oxygen generators capacity, operating requirements**

\[
|O_{ti} - O_{t-1,i}| \leq \beta_{ti} \varepsilon \quad G_{ti} = G_{t-1,i} + Y_{ti} - D_{ti}, \quad G^0 \leq G_{ti} \leq G^1,
\]

\[
\gamma_{ti} = \max \left\{ 0, (\beta_{ti} - \beta_{t-1,i}) \right\} \quad d_i = \sum_{t} D_{ti}, \quad d_i < \sum_{t} G_{t-1,i}
\]

**Pipeline pressure, fluctuation limitations**

\[
(H_t - H_{t-1}) + \sum_{j=1}^{S_{ij}} < \sum_{i} A_{ti} \quad H^0 \leq H_t \leq H^1
\]

\[
\frac{|H_t - H_{t-1}|}{H_{t-1}} \leq \delta \quad A_{ti} \leq \beta_{ti} a_i \quad A_{ti} < O_{ti}
\]

**Oxygen demand constraints**

\[
\sum_j S_{ij} + \sum_{i} Y_{ti} + (H_t - H_{t-1}) + F_{i} = \sum_{i} O_{ti}
\]
Outline

- Research Background
- Optimization Modeling and Method
- Production Scheduling
- Logistics Scheduling
- Energy Optimization
- Data Analytics
Case 1. Prediction of Burden Distribution

Burden distribution

Temperature of the burden surface

The relationship between temperature and burden distribution

Data Analytics and Optimization
6. Data Analytics

Case 2. Temperature Prediction of Molten Iron in Transportation

Task
Predict the temperature of molten iron in transportation process from blast furnace stage to converter stage.

Prediction based on multi-model
- Multiple regression model
- Multiple LSSVM modeling based on estimation of distribution algorithm
- Hybrid modeling based on Kalman filter
6. Data Analytics

Case 3. Dynamic Prediction of BOF Steelmaking Process

Challenges

- Continuous prediction requirement
- Unstable performance of single model
- Dynamic adjustment requirement

Dynamic analytics method

- Multi-stage modeling strategy
- Dynamic model with feedback
- Hybrid kernel function
- Differential evolution algorithm

### Case 3. Dynamic Prediction of BOF Steelmaking Process

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Indexes</th>
<th>SVR</th>
<th>RVM</th>
<th>KELM</th>
<th>DSAE</th>
<th>M-DHLSSVM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>RMSE</td>
<td>MAPE (%)</td>
<td>MAXE (%)</td>
<td>RMSE</td>
<td>MAPE (%)</td>
</tr>
<tr>
<td>T</td>
<td></td>
<td>2.27E+01</td>
<td>1.35E+00</td>
<td>5.08E+01</td>
<td>9.99E-02</td>
<td>1.53E+01</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td>1.26E+01</td>
<td>1.47E-01</td>
<td>1.60E+02</td>
<td>4.03E-01</td>
<td>2.37E+01</td>
</tr>
<tr>
<td>Mn</td>
<td></td>
<td>1.55E+01</td>
<td>8.29E-01</td>
<td>6.34E+01</td>
<td>1.45E-01</td>
<td>2.29E+01</td>
</tr>
<tr>
<td>Si</td>
<td></td>
<td>4.03E+01</td>
<td>8.51E-02</td>
<td>4.62E+01</td>
<td>8.15E-02</td>
<td>1.26E+01</td>
</tr>
<tr>
<td>S</td>
<td></td>
<td>4.00E+00</td>
<td>5.96E-01</td>
<td>1.56E+01</td>
<td>3.94E-02</td>
<td>3.45E+00</td>
</tr>
<tr>
<td>P</td>
<td></td>
<td>1.13E+01</td>
<td>2.12E-01</td>
<td>1.56E+01</td>
<td>3.94E-02</td>
<td>3.45E+00</td>
</tr>
<tr>
<td>No. of Best</td>
<td></td>
<td>1/18</td>
<td>1/18</td>
<td>0/18</td>
<td>2/18</td>
<td>14/18</td>
</tr>
</tbody>
</table>

Dynamic prediction results by M-LSSVM, M-HLSSVM and M-DHLSSVM.

6. Data Analytics

Case 4. Temperature Prediction of Reheat Furnace

Features of Heating Process
- Dynamic
- Non-linear

Mechanism Model
- Difficult to obtain
- Obvious prediction error

Mechanism Model
- LS-SVM is used to compensate for the prediction deviation of the slab temperature
- Significantly improve the model prediction accuracy
6. Data Analytics

Case 5. Strip Quality Analytics of Continuous Annealing

Difficulties in current production process modeling

- Production process is very complex
- Exact mechanism model is not available
- Dimensions of data are more than 30
- Dimensions of data are strongly correlated

Fluctuations of strip quality have caused great economic loss to the cold rolling mill
6. Data Analytics

Case 5. Strip Quality Analytics of Continuous Annealing

Multi-objective Ensemble Learning

- Least Square Support Vector Machine (LSSVM)
- Multi-objective Evolutionary Algorithm
  → Sub-learner in the ensemble learning
  → Evolving the ensemble learning model

Data Analytics and Optimization
6. Data Analytics

Case 6. Process Monitoring and Diagnosis of Continuous Annealing

Motivation

- Tension fluctuations
- Running deviation
- Breakdown of production line

Data Analytics

- Monitoring Model
- Diagnosis Model
- Monitoring the strip tension
- Locating the roller where the deviation occurs

Data Analytics and Optimization
6. Data Analytics —— Energy analytics

**Descriptive**
- Energy consumption and regeneration data
- Filter out exceptional data
- Complement the missing data
- Obtain actual process data

**Diagnostics**
- Solution
- Benchmark Analysis
- Bottleneck Identification
- Energy Analysis
- Analyze cause
- Identify bottleneck
- Analyze consumption
- Energy saving

**Predictive**
- Energy regeneration
- Predictive Diagnostics
- Energy saving
- Analyze cause
- Identify bottleneck
- Analyze consumption

- **Iron smelting**
  - Liquid iron
  - High temp
- **Steel making**
- **Heating furnace**
  - Slab
  - High temp
- **Hot rolling**
  - Slab
  - High temp
- **Hot rolling**
  - Hot coil
  - High temp
- **Acid rolling**
  - Cold coil
  - Normal temp
- **Continuous annealing**

- **Energy regeneration**
- **Coal Gas**
- **Pulverized Coal**
- **Oxygen**
- **Natural Gas**
- **Electric**
Conclusion and On-going Research

Intelligent Perception and Data Analytics for Cyber-physics System

Data Analytics and Optimization

Internet-of-Things-based IntelliSense

Cyber-physics System

- Steel/ Nonferrous
- Energy
- Petroleum-Chemical
- Mining/ Logistics

First principle

Meta-heuristic

Optimization

Plant-wide Production and Logistics Planning

Production/ Logistics Batching and Scheduling

Process Optimization and Optimal Control

Data Analytics

Descriptive
Diagnostics
Predictive

Production data
Logistics data
Energy data
Equipment data
Conclusion and On-going Research

Data Analytics and Optimization

- Planning: Production and inventory
- Scheduling: Production with batching
- Process optimization: Operations
- Descriptive: Production process
- Diagnostics: Equipment
- Predictive: Product quality
- Image: Recognizing
- Speech: Understanding
- VR/AR: Visualizing
Conclusion and On-going Research

With No Physical Background

Optimization

Data Analytics

Data Analytics

Refine

Logistics, Resources

Petrochemistry, Energy

Steel Industry

rooted in industries
dedicated to academy

Data Analytics and Optimization
Thank You!

Reporter: Lixin Tang

Key Laboratory of Data Analytics and Optimization for Smart Industry (Northeastern University), Ministry of Education, China

November 21, 2019