



INFORMS Journal on Applied Analytics

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Optimization of Continuous Steel Annealing Operations Using Model Predictive Control at Tata Steel, India

Sujit A. Jagnade; , Sachin C. Patwardhan; , Kunal Kumar; , Arup K. Dey; , Sai K. Gudimetla; , Manish K. Singh, Ajay K. Jha; , Gyan Prakash; , Jose M. Korath;

To cite this article:

Sujit A. Jagnade; , Sachin C. Patwardhan; , Kunal Kumar; , Arup K. Dey; , Sai K. Gudimetla; , Manish K. Singh, Ajay K. Jha; , Gyan Prakash; , Jose M. Korath; (2025) Optimization of Continuous Steel Annealing Operations Using Model Predictive Control at Tata Steel, India. INFORMS Journal on Applied Analytics 55(1):48-65. <https://doi.org/10.1287/inte.2024.0183>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2025, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>





THE FRANZ EDELMAN AWARD
Achievement in Operations Research

Optimization of Continuous Steel Annealing Operations Using Model Predictive Control at Tata Steel, India

Sujit A. Jagnade,^a Sachin C. Patwardhan,^b Kunal Kumar,^c Arup K. Dey,^d Sai K. Gudimetla,^d Manish K. Singh,^e Ajay K. Jha,^d Gyan Prakash,^a Jose M. Korath^{a,*}

^a Automation Division, Tata Steel Limited, Jamshedpur 831001, India; ^b Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India; ^c Systems and Control Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India; ^d Cold Rolling Mill Operation, Tata Steel Limited, Jamshedpur 831001, India; ^e Design and Engineering Division, Tata Steel Limited, Jamshedpur 831001, India

*Corresponding author

Contact: sujit.jagnade@tatasteel.com,  <https://orcid.org/0000-0002-0857-1129> (SAJ); sachin@iitb.ac.in (SCP); kunalchoudhary9962@gmail.com (KK); akdey@tatasteel.com (AKD); s.gudimetla@tatasteel.com (SKG); manishk.singh@tatasteel.com (MKS); ajayjha@tatasteel.com (AKJ); gyan.prakash@tatasteel.com (GP); jose.korath@tatasteel.com,  <https://orcid.org/0009-0001-3538-8135> (JMK)

Accepted: October 30, 2024

<https://doi.org/10.1287/inte.2024.0183>

Copyright: © 2025 INFORMS

Abstract. In steel manufacturing, continuous annealing is a crucial heat treatment applied to cold-rolled steel strips to achieve a prescribed temperature, which will ensure quality in terms of mechanical properties. However, controlling this process is challenging because of slow furnace temperature dynamics, significant time delays, frequent changes in the steel mass flow rate, target annealing temperature changes induced by steel grade transitions, and multivariable interactions within the furnace zones. To address these challenges, Tata Steel, India, with consultancy support from the Indian Institute of Technology Bombay, developed a novel model predictive control (MPC) technology-based solution for a continuous annealing furnace, which produces automotive grade steels. The dynamic model was developed using data from both perturbation trials and scraping historical processes. We then converted the model to a discrete-time state-space form and used it to formulate an optimal control problem over a moving time window. The solution generates optimal furnace setpoints by solving this finite-horizon optimal control problem each minute, ensuring smooth temperature transitions during steel grade changes while avoiding operational constraint violations. Tata Steel successfully implemented an MPC-based real-time supervisory optimal control solution, which became fully operational in January 2023. The implementation of the solution has led to a significant improvement in the proportion of annealed products meeting the premium quality band ($\pm 5^\circ\text{C}$ of the target temperature), increasing from 30% (manually operated) to 50% (MPC operated), thereby ensuring better uniformity of properties. Furthermore, an 8% reduction in products outside the widest band ($\pm 15^\circ\text{C}$) has prevented the reprocessing of 13,000 tons of material from one line alone, annually. We have seen a consistent 8% reduction in specific fuel consumption per ton of steel. When considering Tata Steel's current installations and those under commissioning, these improvements translate to savings of US\$2.5 million and a reduction of 10,000 tons of CO₂ emissions annually.

Keywords: continuous annealing furnace • optimization • data-driven modelling • Kalman filter • model predictive control • Edelman Award • optimal control

Introduction

Steel has been an important material in the history of humanity. Over the centuries, it has played a crucial role in the growth of civilizations (e.g., by enabling infrastructure, agricultural implements, and arms). The industrial revolution further bolstered the position of steel as an indispensable material, and the steel manufacturing process underwent a transformation in the ensuing period. Steel touches every aspect of

human life from basic infrastructure, transportation, packaging, home appliances, and industrial machinery to advanced applications like nuclear and space technology. Evolving along with the application requirements, present-day steel is available in different varieties of strength, finish, and shape. Its versatility and robustness make it a fundamental material in the manufacturing sector worldwide. Steel production plays a critical role in the global economy, and per

capita consumption of steel is often used as an index in developmental economics. Many countries have developed dedicated policies to facilitate production and consumption within the steel sector. The increasing demand for steel, driven by urbanization, industrialization, and infrastructure expansion, underscores its significance in the global market. Although infrastructure forms a major segment by volume of steel production and consumption, automotive steel is the premium segment in terms of value.

Transportation, especially human mobility, is an essential aspect of daily life and automotives are invariably the primary enablers of the everyday movement of people worldwide. Given its scale, impact on safety, and fuel consumption, the automotive industry is always at the forefront of innovations in designs and materials. It has triggered demands for stronger and more durable steel with characteristics suitable for modern automobiles. In the case of automotive steel, value addition is concentrated toward the downstream processes such as cold rolling, annealing, and coating. Typically, more than 60% of the steel used in a passenger car is made of cold-rolled and annealed steel. Hence, the quality, safety, and durability of the car depends on the quality of the steel used. The steel product quality requirements and standards to be followed during the manufacturing of steel are well established as standards in the automotive ecosystem. The ability of steel manufacturers to produce steel with consistent quality is important in retaining existing customers and acquiring new customers.

The Indian steel industry is no exception to above trends and is currently embracing them, given the rapid economic growth path the country is on. Tata Steel is the largest steel company in India and ranks among the top global steel companies, boasting a crude steel capacity of 35 million tons per annum (MTPA). It is among the most geographically diversified steel producers, with operational activities and a commercial presence that extends across multiple regions worldwide. Tata Steel's overall operational capacity in India is 20.6 MTPA, encompassing flat and long steel products. Approximately 75% of its throughput consists of flat products, including hot-rolled and cold-rolled strips. Cold-rolled steel is utilized in various applications, such as automotives and white goods (i.e., large home appliances such as ovens, refrigerators, and dishwashers), making it a value-added product. Cold rolling of steel is a process wherein a steel strip or sheet at room temperature is passed through a number of rolling stands, which compress it under high pressure and reduce its thickness. This helps to produce a steel strip with a high-dimensional precision and surface finish. However, cold-rolled steel's ductility diminishes because of the work hardening phenomenon (Avner 1974). Work hardening is the phenomenon whereby material become hard and brittle because of the generation of large

numbers of internal imperfections during the application of high rolling stress. Prior to utilizing the steel sheet for applications involving a forming process, recovering its lost ductility is essential (Colvin and Juthe 2007). The enhancement of ductility involves subjecting cold-rolled steel to an annealing process before applying it in products, which represents a pivotal stage in the cold-rolling value chain. Annealing is heating the steel strip to a specific temperature wherein microstructural changes take place rendering it again ductile and formable. In the modern continuous annealing line (CAL) process, the steel strip passes in a continuous manner through an annealing furnace chamber with progressively increasing temperatures and is heated gradually to achieve a specified target temperature. Since the early 1980s, continuous annealing technology, which is superior to batch annealing in producing consistent and advanced high-strength steels, has been used increasingly to produce the steel needed in automotive and other high-end applications (Mould 1982). Additionally, continuous annealing presents various advantages that make it an appealing process for steel producers, including quality enhancement, reduced production costs, and improved product versatility.

Over the past three decades since 1990, India's economy has undergone a significant transformation because of changes in consumption trends. Notable was the high growth in automotives, passenger vehicles in particular, and white goods sectors. Responding to this change early on, in 2001, Tata Steel established a cold-rolling mill complex and CAL combined with galvanization at its plant in Jamshedpur, India, and it stands as a pioneer in the Indian automotive steel industry. It was the first to manufacture cold rolled, continuously annealed, and coated sheets for automotive applications in India. This technology contributes to the production of advanced high-strength steels for automotive and other premium applications, making it the most value-added product in Tata Steel's portfolio in the Indian steel market.

Tata Steel holds a leadership position in the Indian automotive steel market with a market share of 50%. This corresponds to \$1.68 billion of revenue, which is about 15% of Tata Steel's total revenue. With the completion of additional annealing lines, which are under commissioning, Tata Steel's market analysis/intelligence division estimates that this revenue stream will grow to \$2.96 billion, which is equivalent to about 23% of the company's total revenue. In addition, the Indian automotive steel market is witnessing dramatic growth and is becoming increasingly competitive. This has led Indian steel manufacturers to increase their focus on the quality, cost, and throughput of the manufacturing process as they also ensure efficiency in other aspects of the supply chain. For Tata Steel, the imperative is to retain its existing leadership position in the automotive steel sector and also to grow in volume and take advantage of the emerging

trends in value-added segments. Because of its impact on final product quality and position in the process chain, the annealing process is critical to product quality and often becomes a bottleneck in terms of plant throughput. As a result, optimal operation of the annealing process is of utmost importance, and there is significant technological and business interest in continuous annealing technology, with several lines in operation globally.

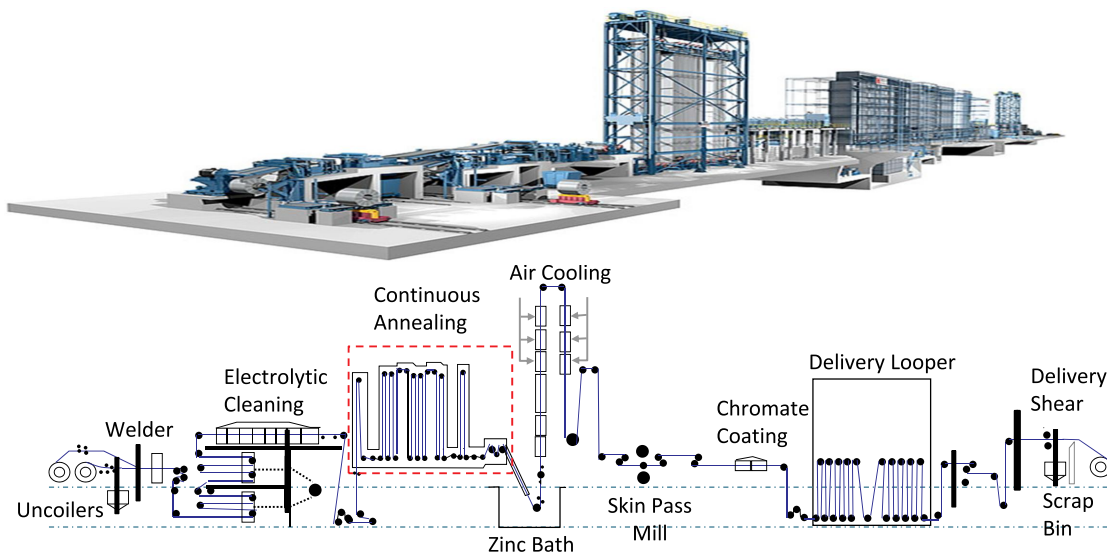
The task of optimizing the CAL operation needs to be viewed in the light of the complexity of the supply chain of an integrated steel plant. Given complex operations, for example, the movement of raw materials and finished products across geographies, handling of bulk material in solid and liquid form, and heavy logistics operations during the processing stage within the steel works, it takes about eight weeks on average to complete the conversion from raw material to finished dispatchable steel product. In addition, the variety of products made in terms of grade and section is substantial; as an example, the number of stock-keeping units (SKUs) in flat rolled products runs into a few tens of thousands. With such macro- and microlevel complexities in the supply chain, to meet on-time delivery commitments to customers, we use a well-established, robust, and flexible framework for production planning and scheduling. Although this provides optimal planning and scheduling at an overall plant level, when scheduling an individual line, the specificities and constraints of that particular line must be considered. Hence, the local optimization of each line demands a customized approach and must follow the overall production plan specified by the enterprise-level planning

optimization tool. Additional optimization at the local level must use this given plan as a constraint. The case we discuss in this paper addresses the CAL process, which is toward the end of the value chain and is one of the most value-added processes. Hence, it is a case of last-mile optimization wherein both constraints flowing from upstream and constraints imposed by the specificities of this line impact the operational performance of the line. A suboptimal operation during the last-mile process is costlier given the nullifying effect it can make on the value additions in previous stages. Hence, the optimal operation of CAL under all operating conditions is imperative.

Continuous Annealing Line—Operation, Control, and Challenges

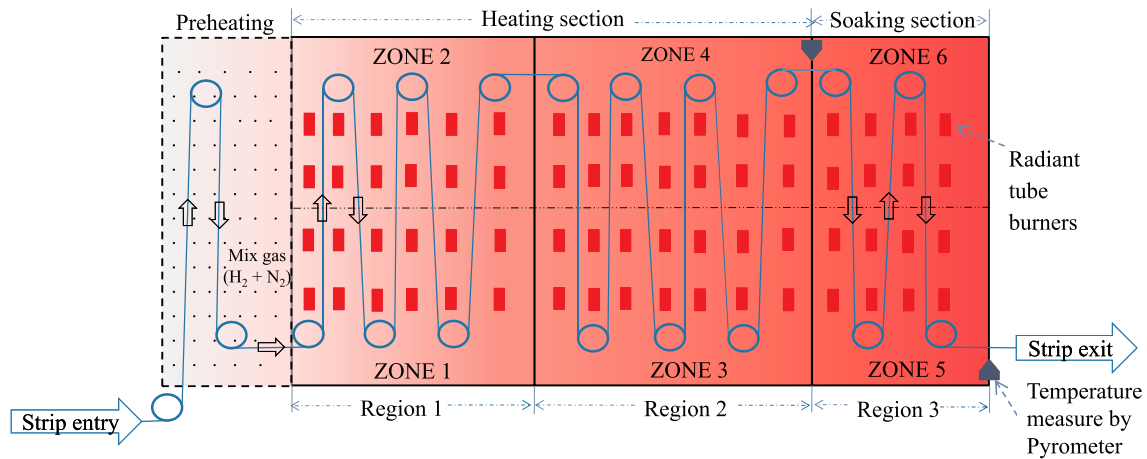
CAL typically comprises various interconnected unit operations, including coil unwinding, welding, electrolytic cleaning, the annealing furnace (i.e., CAL furnace), optional metal coating, temper rolling, chromate coating, and winding operated in tandem (see Figure 1). At the entry side of the CAL, the strip is uncoiled and entered into the line. To ensure continuous feeding, the head end of the next coil is welded to the tail end of the leading coil in the welding machine at the line entry. After entering the line, the strip first passes through the electrolytic cleaning section where the strip surface is cleansed of the lubricant emulsion applied during the upstream cold-rolling process. After cleaning, the strip enters the annealing furnace where it is heated progressively to the required temperature. After exiting from the furnace, it is cooled. The strip then passes through

Figure 1. (Color online) (Top) A Three-Dimensional Visualization of a CAL and (Bottom) a Schematic Representation of a Different Interconnected Process in a CAL



Notes. For additional information on annealing, see <https://www.sms-group.com/plants/annealing-and-galvanizing-lines-for-steel>. Zinc coating is used to protect the steel strip from corrosion. In the skin pass mill process, a very mild reduction in thickness is applied to achieve certain mechanical properties. Because the thickness reduction is extremely small (“skin deep”), it is called skin pass.

Figure 2. (Color online) Side View of the CAL Furnace



Notes. Small boxes within the zones represent radiant tube burners, which are parallel to the moving steel strip in the middle. We discuss zones in the subsection CAL Furnace Control.

an optional metal coating unit and subsequently undergoes temper rolling and chromate coating. It is then cut and wound to again form a coil in a winding unit at the delivery side of the CAL (Takeo 1985).

The annealing furnace is the central unit of the annealing line. The furnace consists of preheating, heating, and soaking sections (see Figure 2).

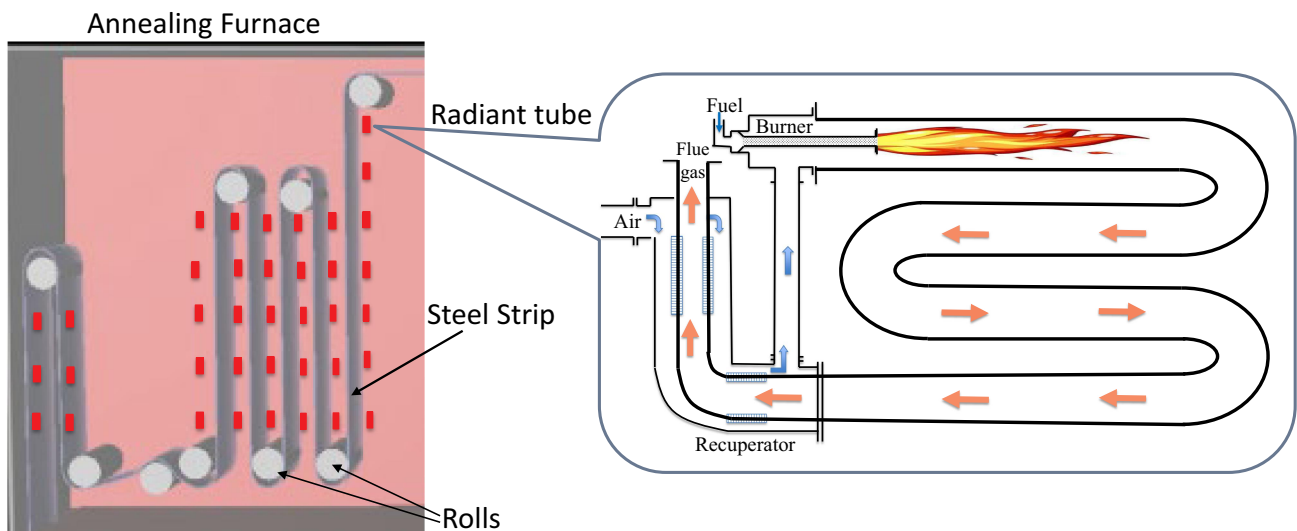
The preheating section does not have heating elements per se. This section is filled with a gaseous mixture of hydrogen and nitrogen. This recirculating gas mixture is heated by an external heat exchanger. As the strip pass through this gaseous medium, it is then heated by convection and radiation. The heating and soaking sections are equipped with heating elements called radiant tubes for the indirect heating of the passing strip (see Figures 2 and 3).

Fuel (i.e., coke oven gas) burns inside these radiant tubes; the heat is conducted through the tube wall, radiates out into the furnace chamber, and then heats up the passing strip. Most of the strip heating occurs in the heating section; hence, the maximum number of radiant tubes are in this section. This section is followed by a relatively shorter soaking section in which the heat penetrates from the surface into the strip cross section. Pyrometers are used to measure the strip temperature at the end of the heating and soaking sections to monitor and ensure that timely control actions are taken so that the strip achieves the specified target temperature (Imose 1984, Takeo 1985).

CAL Furnace Control

The primary objective of the annealing furnace is to heat each part of the incoming strip at a specified rate to a

Figure 3. (Color online) (Left) A Three-Dimensional Visualization of Steel Strip Running Over the Rolls, and (Right) a Schematic Representation of a Radiant Tube Burner



given target temperature, which is crucial for ensuring efficiency and product quality. This precise temperature control is crucial for achieving the desired metallurgical properties and product quality in the steel sheets produced. Depending on the steel grade of the coil and the desired mechanical properties of the final product, the quality control department prescribes an “annealing recipe” consisting of (1) ideal temperature setpoints (i.e., points or levels at which the temperature stabilizes) for the annealing furnace sections, (2) the target strip exit temperature for the steel grade, and (3) allowable maximum and minimum deviations in the strip exit temperature achieved. The furnace temperature profiles across the heating and soaking sections must be maintained under all circumstances to achieve the target strip temperature within a given tolerance. In addition, fuel efficiency and throughput must be maximized.

The temperature control, and hence the fuel flow control in the furnace, is aligned with the above functional structure and operational requirements. For purposes of temperature control, the heating section is divided into two regions and the soaking section is the third region. At any given point in time, fuel flow into radiant tubes of each region is regulated to achieve a desired furnace temperature in that region. To minimize the pressure drop in fuel flow across radiant tubes within a region, each region is divided into two zones, and fuel flow is through dedicated valves to each zone. However, from a furnace temperature control point of view, the two zones constituting one region must reach a single

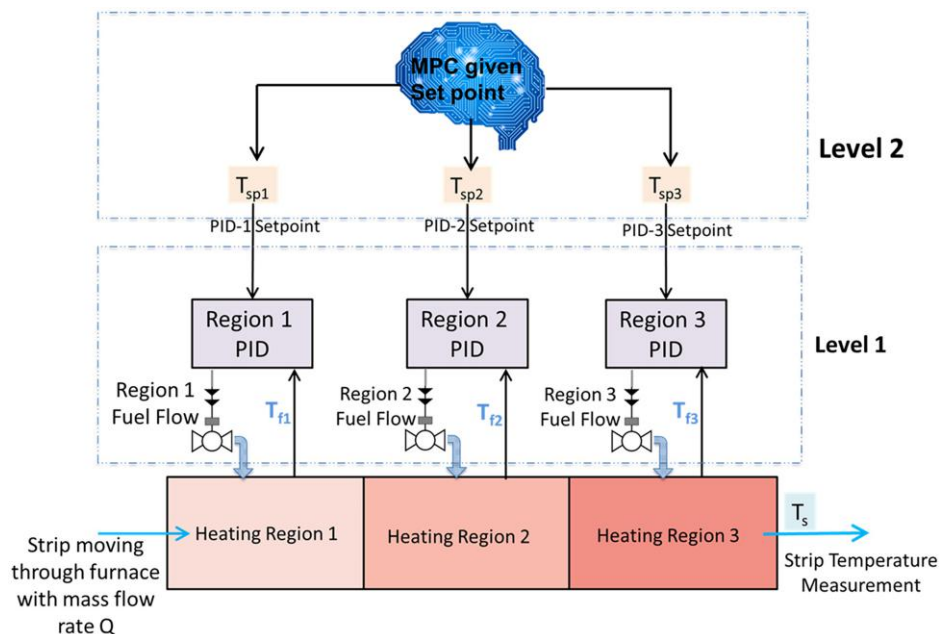
temperature, which is measured by thermocouples located at representative locations in each region.

Relative to the temperature control philosophy, the desired furnace temperature to be maintained for each region is given as a setpoint to the basic control system (Level 1). A schematic representation of the legacy control architecture is shown at the bottom of Figure 4. This system involves three furnace regions, proportional, integral, and derivative (PID) controllers, PID-1, PID-2, and PID-3, which, in turn, provide fuel flow setpoints to arrays of low-level controllers. These low-level controllers, that is, programmable logic controllers (PLCs), in conjunction with safety interlocks, manipulate fuel flows into the radiant tube burners. These individual PID loops ensure closed-loop control of the temperatures of each furnace region. However, whenever a change in the strip exit temperature is required by design or the temperature moves from its given target, an operator needs to manually change the furnace regions’ temperature setpoints. Hence, the legacy control philosophy requires using an open loop for strip temperature control and a closed loop for furnace temperature control. We note that all setpoint execution is subject to limitations on the maximum radiant tube temperature and other related constraints.

Operational Challenges

The manual decision on a furnace temperature profile, implemented as setpoints for the PID loops, is the main source of suboptimality in different scenarios during the furnace operation, as we describe in the following paragraphs.

Figure 4. (Color online) The MPC Solution Was Developed as a Supervisory Optimal Control Layer (Level 2) to Provide an Automatic Setpoint to the Honeywell Legacy Control System (Level 1)



Notes. In manual operating mode, PID setpoints are specified by operators (T_{fj} represents the j th furnace region temperature). PID control loops are a conventional feedback control methodology.

Given the target temperature for a given grade of steel, an operator is guided by the annealing recipe, which suggests the temperature setpoints for different regions of the furnace. However, depending on the furnace operating conditions, the operator relies primarily on his or her knowledge and skill to determine the furnace temperature setpoints to achieve the target strip temperature. Observing the strip temperature at the furnace exit, the operator continuously readjusts the furnace setpoints in his or her attempt to achieve the target strip temperature. Such reactive interventions will have limitations on its efficacy in actual plant operating conditions. We next explain two scenarios that occur most frequently when the manual decisions on strip temperature setpoints are highly suboptimal.

The first scenario is related to the production plan, which is generated at the overall plant level to ensure that the delivery timelines and duration that each processing step demands are met. Once this plan has been developed, a schedule (i.e., the order in which SKUs are processed) is generated for each line in the value chain considering the particular line's specificities and constraints. Although each line has a core manufacturing process, many preprocesses and postprocesses, especially in the continuous lines, could also be present. The constraints of these pre- and postprocesses also need to be considered when we do the scheduling. In the context of the CAL, which consists of several tightly coupled units, the constraints of a preprocessing unit (e.g., welder) and a postprocessing unit (e.g., rolling mill) are activated because of the difference in thickness and width allowed between successive strips. Hence, although processing coils of similar target temperatures together would be better, the difference in width and thickness could impose constraints on welding and rolling between successive coils. The net result is that even the most optimized schedule will include transitions in target temperatures more frequently than we desire. This is an effect of the overall plant-wide production plan acting as a constraint on the scheduling of the individual line. Optimal handling of these transitions is a significant challenge for an operator because of the huge variety in transitions and the need for coordinated changing of setpoints of all furnace regions to ensure efficient heating of the strip.

The second scenario arises out of the interconnected nature of the line where various unit operations with different scales of dynamics operate in tandem. This makes it vulnerable to suboptimal operation because a disturbance from any unit could adversely affect another unit. Such disturbances lead to unforeseen speed (and/or throughput) changes, necessitating furnace temperature manipulation to address disturbances and maintain the target strip exit temperature.

Both of the above scenarios can also occur simultaneously, making the transition even more challenging. The

control problem is further complicated by the fact that the system has to be operated under various constraints arising from safety and the limits of the actuators. The grade transitions and/or disturbance attenuation of the furnace operation must be carried out under the following constraints:

- Radiant tube temperatures must be maintained below specified safety limits under all conditions.
- Maximum permissible fuel flow in each burner, and, in turn, each furnace region has an upper bound.
- Permissible rates of change of fuel flow rise and fall differ.

Moreover, the annealing furnace is a distributed parameter system (i.e., the strip temperature changes with time and space), which renders control tasks difficult. The conventional approach to modelling such systems is to approximate the effect of spatial variations by including input time delays (ITDs). Thus, for the control relevant dynamic models, we developed differential equations involving temperatures at the end of each furnace region in combination with large input time delays. We note that systems with large input time delays are inherently difficult to control. Thermal interactions within different zones of the furnace further complicate the control problem.

The fallout of suboptimal management of such transitions is the underheating or overheating of steel strips, resulting in compromised temperature compliance (hence, loss of uniformity in mechanical properties), loss of throughput, and excess fuel consumption, which increase costs and CO₂ emissions. Achieving a tighter control of the strip temperature around the target ensures uniformity in mechanical properties throughout the strip length.

CAL must be able to handle the scenarios we describe above in an optimal manner with minimum adverse effect on product quality and furnace operational efficiency (i.e., fuel consumption and throughput). However, with the legacy control architecture, which consists of three PID controllers that act independently with no coordination among them, and with demand for continuous operator input of the furnace temperature setpoints (T_{sp1} , T_{sp2} , T_{sp3}) to maintain the strip exit temperature, achieving a smooth and satisfactory control performance is difficult, especially in the case of transitions and disturbances. Hence, an efficient and robust solution, which could control the strip temperature in a closed loop and in an optimal manner and could also be integrated with the basic control system, became imperative. We did an exhaustive literature review and explored various options for the optimal control architecture, which we cover in the Review of Continuous Annealing Furnace Control Literature section below.

Review of Continuous Annealing Furnace Control Literature

The conventional approach to controlling an annealing furnace is the use of multiple-loop PID controllers

(Åström and Hagglund 1995, Martineau et al. 2004). However, because of a lack of coordination between their control actions, dealing with difficulties arising from multivariable interactions, systematically meeting operating constraints, and achieving optimal operations can be difficult. The best method available to handle such complex multivariable control problems with operating constraints is model predictive control (MPC) technology, which emerged in the process industry in the early 1990s (Qin and Badgwell 2003). MPC is a real-time optimal control formulation, which can handle systems with multiple inputs and operating constraints. Recognizing the potential, some academic and industrial research groups have utilized MPC technology for optimal control of the annealing furnace operation. Yoshitani and Hasegawa (1998) report the application of adaptive generalized predictive control (AGPC) to the annealing furnace. These authors control the strip exit temperature and average furnace temperature using the aggregate fuel flow. They develop a discrete-time grey-box autoregressive with exogenous input (ARX) model using input–output data and use it in formulating an AGPC scheme. Model parameters are estimated online to compensate for process nonlinearities. Although the results of actual plant implementations are encouraging, the use of a single aggregate flow rate restricts the ability to manipulate different zones of the furnace separately. Moreover, the authors do not discuss constraint handling using AGPC. Bitschnau et al. (2010) propose to use a linearized mechanistic model for developing an MPC process for an annealing furnace. They propose to manipulate fuel flows in each zone separately thereby retaining the ability to implement distributed manipulation. Moreover, input and output constraints are handled systematically. The efficacy of the proposed MPC scheme is demonstrated using simulation studies. Niederer et al. (2016) and Strommer et al. (2018) propose to use a first-principles model (i.e., physics-based mechanistic model) directly for developing a nonlinear MPC scheme for an annealing furnace. In addition to servo and regulatory control objectives, the controller maximizes throughput and minimizes energy consumption. The efficacy of their proposed approach is demonstrated using a high-fidelity simulator of an industrial annealing furnace. The computation time needed to simulate the dynamics in real time for a nonlinear MPC implementation can prove to be a stumbling block in the implementation process. Wu et al. (2015) also propose to use a detailed mechanistic model for optimal control of an annealing furnace. To simplify online computations, they carry out piecewise linearization of the nonlinear mechanistic model. Although the direct use of a detailed mechanistic model for formulating an MPC scheme is attractive from a control viewpoint, the development and validation of a reliable

mechanistic model can be a time-consuming and economically unattractive option.

MPC-Based Supervisory Control Solution

Our objective in this project was to build an advanced process control solution that would automatically generate optimal furnace temperature setpoints so that the strip reached the required target temperature at the furnace exit. Such advanced control solutions are offered by established vendors in the domain. However, these solutions are proprietary in nature, requiring dependency on the original equipment manufacturer for customization and fine-tuning as business needs evolve. Tata Steel has a dedicated in-house team engaged in the development of real-time optimization solutions for manufacturing processes across the value chain. Hence, developing the MPC-based solution for the annealing process and providing the team with advanced training for transporting the capability to other manufacturing processes was a strategic decision. Therefore, we developed a supervisory (cascade) control structure in which the MPC system acts as a primary controller (Level 2), and it supervises the secondary controllers (i.e., Level 1 of the legacy control system), as we show in Figure 4.

The literature includes reports of solutions wherein the MPC bypasses the PID controllers and directly manipulates the fuel flow setpoints (Yoshitani and Hasegawa 1998, Wu et al. 2015). In our current project, the MPC solution was developed to play a supervisory role and manipulated only the furnace temperature setpoints (T_{sp1} , T_{sp2} , T_{sp3}). This strategy facilitated a smooth change management process in the plant operations ecosystem. This choice implied that, when MPC was disconnected, the furnace operators would revert to manually changing the setpoints (i.e., T_{sp1} , T_{sp2} , T_{sp3}) to control the operation. More importantly, when the MPC controller was working in automatic mode, the operators could continuously verify whether the actions being planned by MPC were in line with their experience and intuition. Thus, MPC received measurements of the furnace temperatures (T_{f1} , T_{f2} , T_{f3}), mass flow rate (Q), and strip exit temperature measurement (T_s) from the legacy Honeywell system, and manipulated setpoints of PID-1, PID-2, and PID-3 were communicated to the legacy system; see Figure 4. Figure 4 also shows that the steel mass flow rate, Q , acts as a disturbance input to all three regions of the furnace. Note that fluctuations in the steel mass flow rate can occur because of changes in the strip width, thickness, and line speed. Because these are measured disturbances, MPC can proactively take feedforward control action when line speed, thickness, or width changes occur. This feature does not exist in the legacy control system.

An MPC scheme consists of following components
(1) control-relevant dynamic model used for online

predictions, (2) Kalman filter used for dealing with state estimation and model plant mismatch computation (see section State Estimation and Handling of Model Plant Mismatch Using Kalman Filtering below), and (3) constrained optimization formulation, which is solved in real time repeatedly over a moving time window. Figure 5 shows a schematic representation of an MPC scheme. In this section, we describe each of these components in detail.

Control-Relevant Dynamic Model Identification

The MPC scheme uses a dynamic model for online forecasting of the future behaviour of a plant over a moving time window $[t, t + T_p]$, where t represents the current time, and T_p represents the prediction horizon. Thus, a reliable dynamic model that can forecast the future behaviour of CAL over a time window of 30 to 40 minutes is at the heart of MPC. Such a model can also be developed using the first principles-based approach (Wu et al. 2015, Niederer et al. 2016). However, for a complex system such as an annealing furnace, this can be a time-consuming and expensive option, making an MPC implementation infeasible from an economic viewpoint. An alternative approach popularly used in the industry is to develop black-box or grey-box dynamic models using operating data (Qin and Badgwell 2003). Using this approach, the system under consideration is stimulated by injecting deliberate input perturbations, and the observed system output response is used to construct a differential or difference equation model that fits the input–output perturbation data (Ljung 1999). The model identification problem is cast as finding a set of model parameters that minimize the sum of squared prediction errors between the measured output and model predictions. An examination of Figure 4 reveals the following:

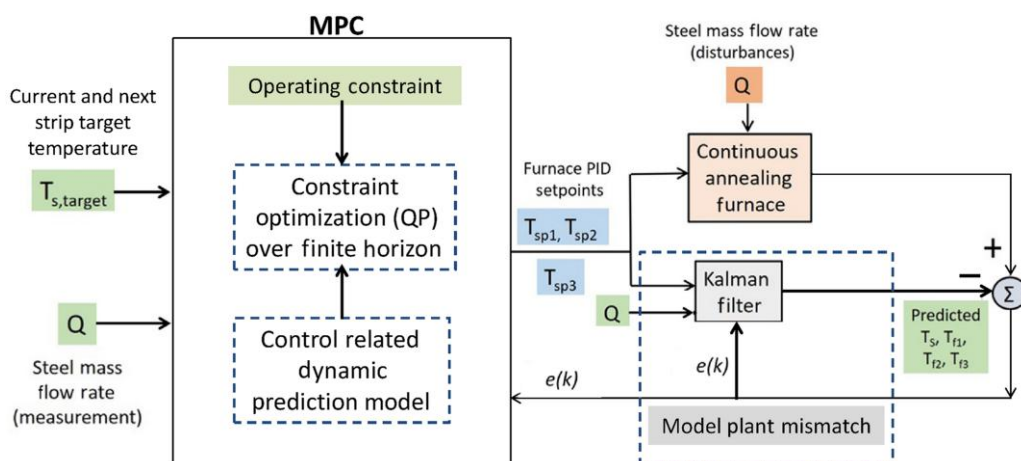
- Region 1 PID setpoint change affects the furnace zone temperatures in all three regions.
- Region 2 PID setpoint change affects the furnace temperatures of Region 2 and Region 3 only.
- Region 3 PID setpoint change affects the furnace zone temperature of Region 3 only.
- Steel mass flow-rate change (caused by either a change in line speed, thickness, or width) influences temperatures in all three regions.

Based on these observations, a sequential multiple-input, multiple-output (MIMO) model that consists of two components, (1) a furnace temperature model and (2) a strip temperature model, is proposed. Figure 6 shows a graphical representation of this model and Table 1 presents input and output variables associated with the two submodels.

In Figure 6 ($T_{sp1}, T_{sp2}, T_{sp3}$) and Q are external inputs driving the furnace temperature model. The operators G_{ij} in the boxes represent transformations that capture the effect of the i th input signal on the j th model output. Each G_{ij} operator is modelled as an ordinary differential equation (ODE) with an input delay. This model is referred to as the ODE with input time delay (ODE-ITD) model in the rest of the text; see Appendix A and Åström and Hagglund (1995) for details. The outputs of the furnace temperature model (i.e., predicted T_{f1}, T_{f2}, T_{f3}) and the steel mass flow rate, Q , are inputs to the strip temperature model. These “inputs” again undergo transformation through second-order-plus-time-delay (SOPTD) operators and together influence the exit strip temperature, T_s . Note that each input has a different influence on the strip temperature (with different time delays and gains), and these individual effects are captured separately.

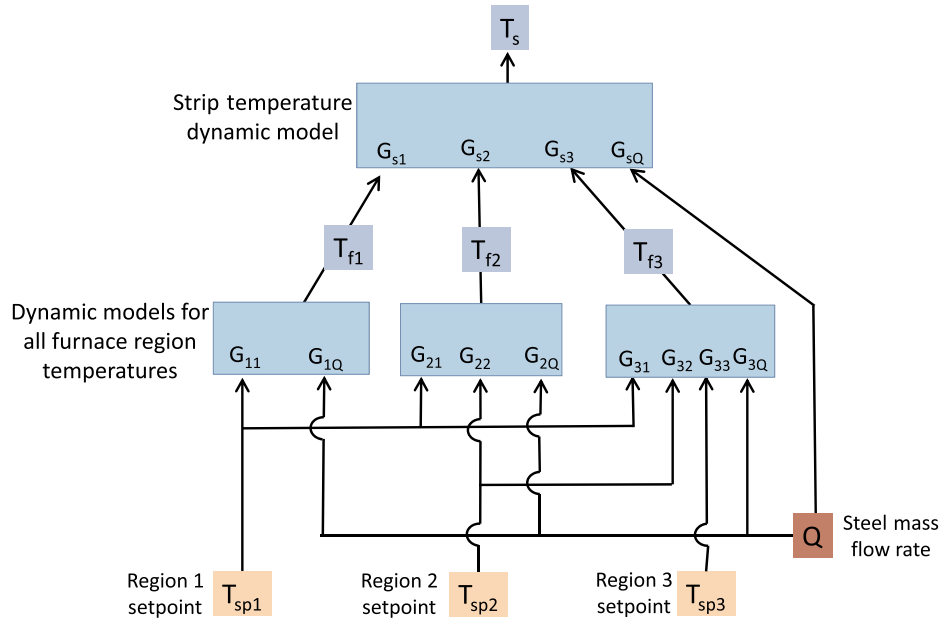
The first step in the model development exercise is determining the structure of an individual ODE-ITD

Figure 5. (Color online) The Information Flow Between the Dynamic Predictive Model, Constraint Optimization, Kalman Filter, and CAL Furnace in an MPC Scheme



Note. The term Σ represents the difference between measured and predicted values.

Figure 6. (Color online) The Sequential Dynamic Black-Box Model for the CAL Furnace Shows the Information Flow Between the Furnace Temperature and the Strip Temperature Submodels



Note. Terms G_{ij} represent the transfer function, which captures the effect of the i th input signal on the j th model output.

operator, G_{ij} , which relates to an input and output pair. In principle, we can identify MIMO models directly from input–output data in which all inputs are activated simultaneously. However, determining the structure of an individual SOPTD operator when all inputs are excited simultaneously is difficult. Thus, we decided to conduct perturbation trials by separately introducing a sequence of step changes in external inputs T_{sp1} , T_{sp2} , T_{sp3} , and Q . These tests were conducted with minimal disturbances to the production schedule. The perturbation data collected using the legacy Honeywell system were available at a sampling interval of three seconds. The data collected during an individual step test were used for model identification. The model development was performed in two steps. Initially, a preliminary high-order ARX model was identified from the input–output data using linear regression (Ljung 1999). We used the step response of the deterministic component of the identified high-order ARX model to generate

estimates of orders, time delays, and coefficients of the ODE-ITD model. Subsequently, utilizing this preliminary information, we identified the MIMO ODE-ITD model relating T_{sp1} , T_{sp2} , T_{sp3} , and Q with controlled furnace region temperatures T_{f1} , T_{f2} , and T_{f3} by minimizing the sum of squared prediction errors (Ljung 1999) using MATLAB’s System Identification Toolbox. Data used for this model development were a mix of step test data and selected segments of historical operating data when grade changes were made, and disturbance input Q was perturbed during the operation.

The modelling exercise was carried out sequentially; that is, parameter identification of the furnace temperature model was carried out first. Subsequently, parameters of the ODE-ITD model represented by G_{s1} , G_{s2} , G_{s3} , and G_{sQ} in the strip temperature model were identified using a combination of step test data, selected segments of historical operating data, and predicted data of (T_{f1}, T_{f2}, T_{f3}) obtained using the furnace temperature

Table 1. The Table Lists Input and Output Variables Associated with the Furnace Temperature and Strip Temperature Dynamic Black-box Submodels

Furnace temperature model		Strip temperature model	
External inputs	Model outputs	Inputs	Model outputs
PID-1 setpoint T_{sp1}	Predicted T_{f1}	Predicted T_{f1}	Predicted strip exit temperature T_s
PID-2 setpoint T_{sp2}	Predicted T_{f2}	Predicted T_{f2}	
PID-2 setpoint T_{sp3}	Predicted T_{f3}	Predicted T_{f3}	
Steel mass flow rate (Q)		Steel mass flow rate (Q)	

model. We again did this by minimizing the sum of squared prediction errors (Ljung 1999) using MATLAB's System Identification Toolbox.

Finally, the combined model obtained through the proposed sequential modelling strategy consists of a set of coupled ODE-ITD-related external inputs, T_{sp1} , T_{sp2} , T_{sp3} , and Q , with the controlled output for MPC (i.e., the strip exit temperature T_s .) We note that the annealing furnace is a slow process and the response time to a step change in an input is in the order of 30 minutes. As a consequence, MPC requires online forecasting over a window of 30 to 40 minutes to make decisions that can achieve smooth setpoint tracking and to avoid constraint violations over the future horizon. Thus, the ODE-ITD model was converted into a digital control-oriented discrete-time state-space model of the form using a sampling interval of $T = 1$ minute (see Appendix A for details). This model is also used for the development of the Kalman filter and MPC scheme.

State Estimation and Handling of Model Plant Mismatch Using Kalman Filtering

To formulate an MPC scheme using ODE-ITD models, it is necessary to estimate the current state of the system. Given the model presented in Equations (A.1) and (A.2) in Appendix A, using the input signals and measurements, there are many ways to reconstruct the state vector by combining the model predictions with the measurement. In this work, we have chosen to use the Kalman filter, which minimizes the covariance of the state estimation error (Åström and Wittenmark 1997, Patwardhan et al. 2006). The main idea is to use the state-space model in Equations (A.1) and (A.2) for predicting the output using the model presented in Equation (B.1) in Appendix B, and information available up to instant $(k - 1)$ to compute the model residuals. This model residual signal carries information about unmeasured disturbances and is used to correct the predicted state estimates (for details, see Equation (B.2) in Appendix B; Åström and Wittenmark 1997). The corrected state estimate is then used for MPC optimization problem formulation at sampling instant k . Moreover, the residual signal also carries information about the model plant mismatch, which is nothing but the difference between the actual plant condition and model prediction. These errors arise from approximations made while developing the control relevant dynamic model. These typically appear as low-frequency drifting mean signals in the model residuals, $e(k)$. Thus, as Figure 5 shows, the model residual signal, $e(k)$, is also used in the MPC formulation for determining compensation for the model approximation errors after removing the high-frequency component using a robustness filter (see Appendix B). This filter plays a crucial role in increasing

the robustness of the MPC controller in the face of model plant mismatch.

Constraint Optimization over a Finite Horizon

At the heart of the MPC formulation are predictions carried out using the model presented in Equation (B.1) over a moving time window $[k, k + P]$, where $P = T_p/T$ represents the prediction horizon, and k represents the current time instant. Given a sequence of future input moves over the horizon $[k, k + P]$, dynamic response of the system can be predicted over the future horizon repeatedly using the difference equation (B.1) in conjunction with the filtered state estimate generated by the Kalman filter and the filtered model residual signal; see Appendix B and Patwardhan et al. (2006). Figure 7 shows a schematic representation of the forecasted outputs in response to future inputs.

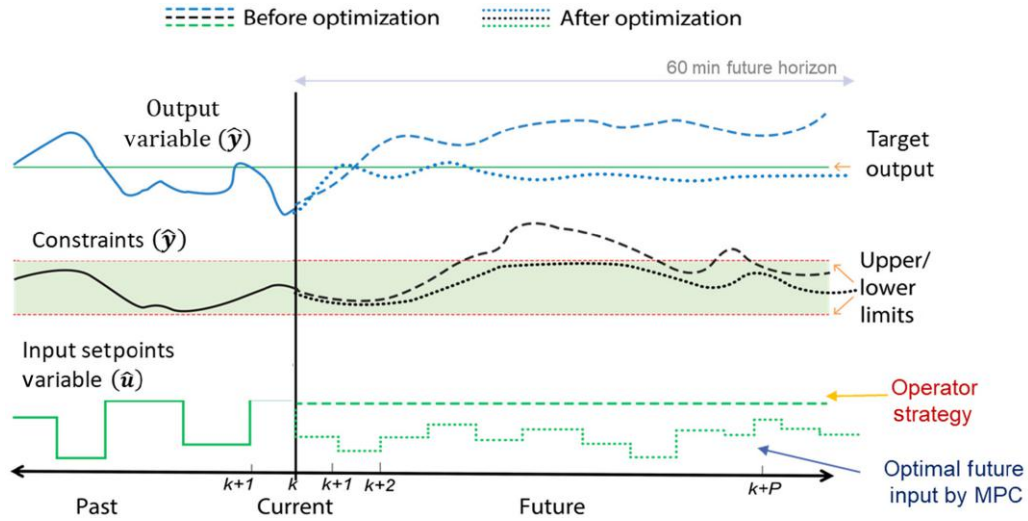
Errors between the predicted and actual outputs are used to formulate the MPC optimization problem. Thus, at sampling instant k , MPC is formulated as a constrained optimization problem that minimizes a cost function consisting of the following terms:

- sum of squared prediction errors between the predicted outputs and the future target profile of strip temperature over the window $[k, k + P]$;
- a penalty term for the rate of change of future manipulated input moves.

MPC is subject to constraints arising from the prediction model, and a set of inequality constraints arising from (1) bounds on future inputs and (2) bounds on the rate of change of future inputs. The prediction horizon used in the MPC formulation is 60; that is, forecasted behaviour over a horizon of 60 minutes is considered for each decision made by the MPC formulation. Details of the controller formulation can be found in Appendix B. As we indicate in Figure 5, using some algebraic manipulations, the constrained optimization problem is recast as a quadratic programming (QP) problem and solved online once each minute. The controller is implemented in the moving horizon framework (Qin and Badgwell 2003). This implies that although optimal future input moves over the entire horizon $[k, k + P]$ are computed by the QP solver, only one move (i.e., $U(k)$) is implemented on the plant. At instant $(k + 1)$, the MPC problem is reformulated over window $[k + 1, k + P + 1]$ and is resolved at the next sampling instant and so on. Thus, the controller implementation turns out to be an unending sequence of optimization problems; that is, 1,440 QP problems (one per minute) are solved every 24 hours.

As we indicate above, each grade of steel requires a target annealing temperature to be maintained at the furnace exit, and about 40 different grades of steel are processed in the annealing furnace. Thus, the furnace operation has to be shifted to an appropriate operating regime every few hours when a grade change occurs. The difference between these strip target temperature

Figure 7. (Color online) Visualization of the Dynamic Model-Based Forecasting and Setpoint Optimization over a Moving Time Window in Our MPC Formulation



changes can be as high as 100 degrees. The supervisory MPC dynamically alters the optimal temperature profile along the length of the strip by manipulating setpoints T_{sp1} , T_{sp2} , and T_{sp3} so that the strip exit temperature target is met despite the measured and unmeasured disturbances that frequently upset the plant operation. In addition, when grade transitions occur, the controller ensures that underannealing or overannealing of trailing and leading strips are minimized because information about the target strip temperatures for the leading and trailing strip is provided to the MPC in advance before the change is about to occur. The dynamic model is capable of predicting strip temperature transients in the immediate future by using real-time operational data, current and future coil target setpoints given by an annealing recipe, production sequence, and the steel mass flow rate; thus, the model manages a smooth transition from one grade to another.

Representative Results of Grade Change and Regulatory Control

To demonstrate the effectiveness of the proposed control scheme, we present a snapshot of the furnace operation for approximately seven and half hours between 11:09 a.m. to 6:30 p.m. on February 4, 2021; see Figures 8, 9, and 10. During this period, 19 coils of two grades of steel (CQ-14 and EIF-5) were processed. The snapshot involved two grade transitions, that is, from CQ (Commercial Quality) to EIF (Extra deep drawing Interstitial Free) and EIF to CQ (see Figure 8), and multiple line speed, thickness, and width changes (see Figure 9). Thus, the MPC was expected to manage grade transitions while simultaneously dealing with the disturbances.

As we mention in the section Implementation Challenges, the production schedule necessitates unavoidable transitions in the furnace temperature profile between successive strips. In Figure 8, we can see that

Figure 8. (Color online) Representative Results of Regulatory Control on Strip Exit Temperatures Using MPC During Simultaneous Grade Transitions and Mass Flow Changes

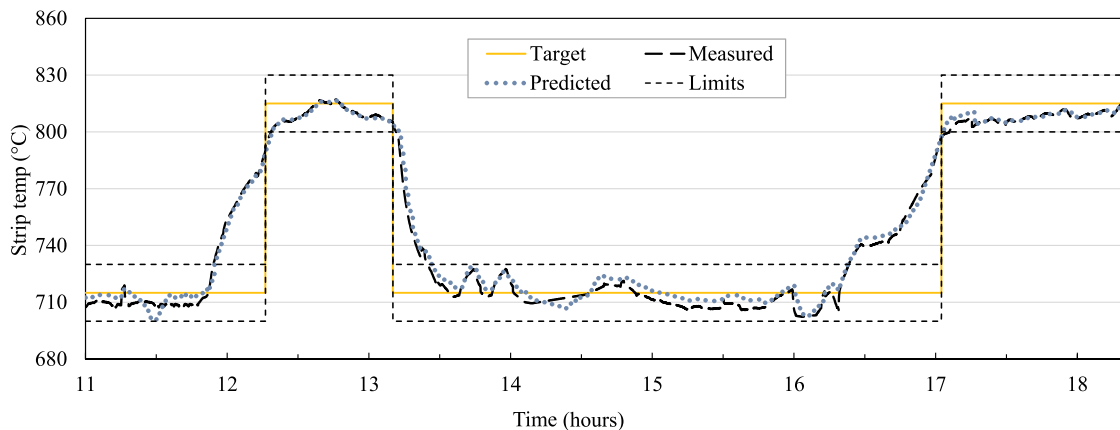
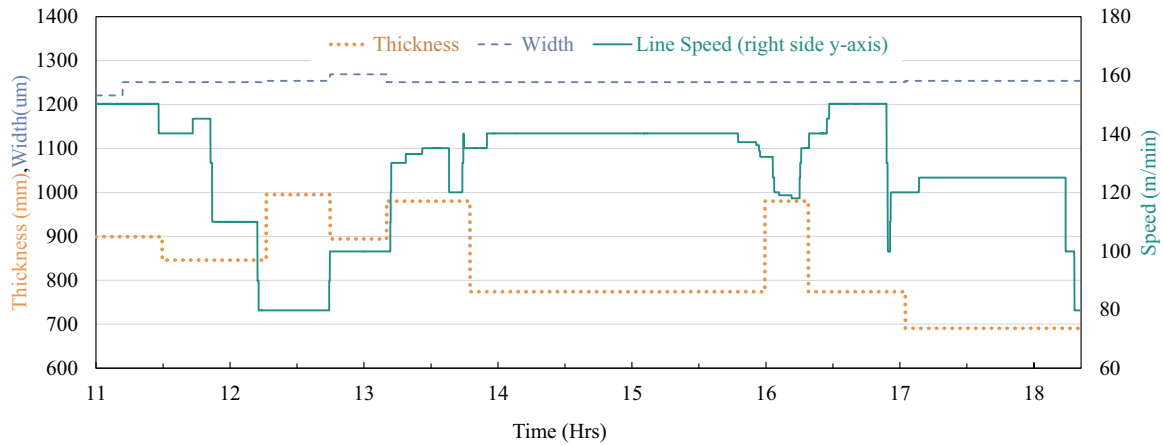


Figure 9. (Color online) Representative Results of Regulatory Control Using the MPC During Simultaneous Grade Transitions and Mass Flow Changes: Disturbance Input (Steel Strip Line Speed, Width, and Thickness)

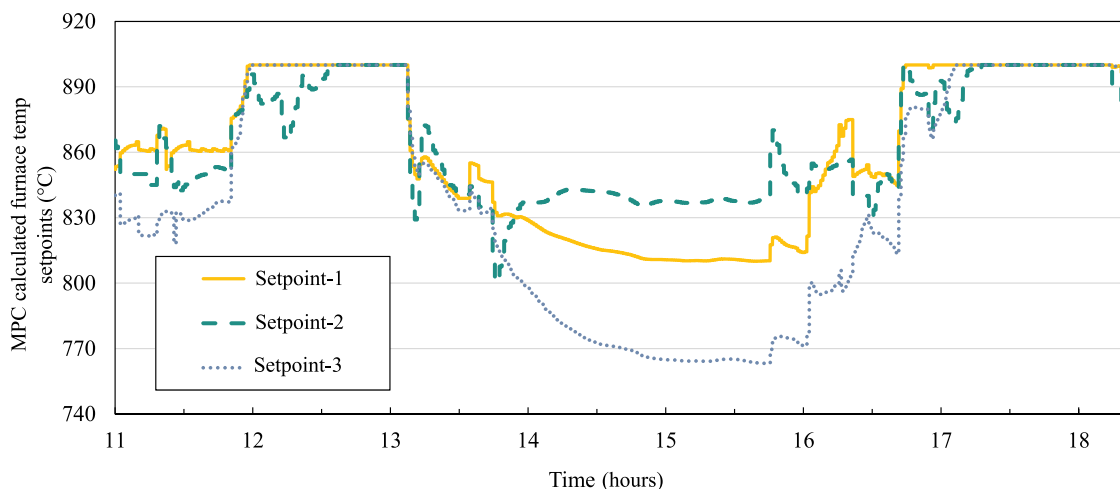


the target temperature increases from 710 to 815, remains at that level for some time, and then falls to 710. After a few hours, the target temperature requirement rises to 815. As we explain above, the unit operations (e.g., welding and rolling, which are part of the line) impose limits on differences in thicknesses and widths of successive strips. Hence, given the variety of SKUs scheduled for production, such increases in target temperatures are unavoidable, as are the furnace temperature transitions. This particular scenario demonstrates the challenges associated with plant operations and the need for an optimal control solution.

Figure 8 reveals that MPC is able to manage the grade transitions very well by maintaining the operation close to the target and within the acceptable temperature tolerance limits for a majority of the operation. Moves outside the specified temperature limit occur only during

transitions from one grade to another, which are inevitable. Moreover, Figure 8 reveals that the prediction model in conjunction with the Kalman filter does a reasonable job of capturing the strip temperature dynamics and the model plant mismatch is reasonably small. Note that MPC starts preparation for the grade change in advance because it is aware of the future setpoint trajectory when the grade changes occur. Moreover, when the specified strip temperature is achieved for a specific grade, it is maintained close to the target strip exit temperature setpoint despite multiple disturbances in the line speed, strip thickness, and strip width (i.e., steel mass flow rate). The corresponding manipulated PID setpoint trajectories generated by MPC are shown in Figure 10. It is evident from this figure that MPC pushes all three furnace setpoints close to the constraint boundaries for moving and maintaining the strip exit

Figure 10. (Color online) Representative Results of Regulatory Control on PID Setpoints Using the MPC During Simultaneous Grade Transitions and Mass Flow Changes



temperature in the high-temperature zone. When an operator tried to do a similar manipulation manually, this action would often result in tripping the furnace operation because of hitting the maximum temperature limits of the radiant tubes. We did not observe such tripping behaviour when MPC managed the furnace operation in the high-temperature region.

Impact and Benefits

The MPC-based supervisory optimal control solution was implemented in the second half of 2021; however, the online trials were severely affected by the COVID-19 pandemic-related restrictions. With the subsiding of the pandemic by mid-2022, more frequent trials were resumed, and the solution was made fully operational beginning in January 2023. Figure 11 depicts month-to-month usage of the MPC-based annealing furnace operation from January to December 2023.

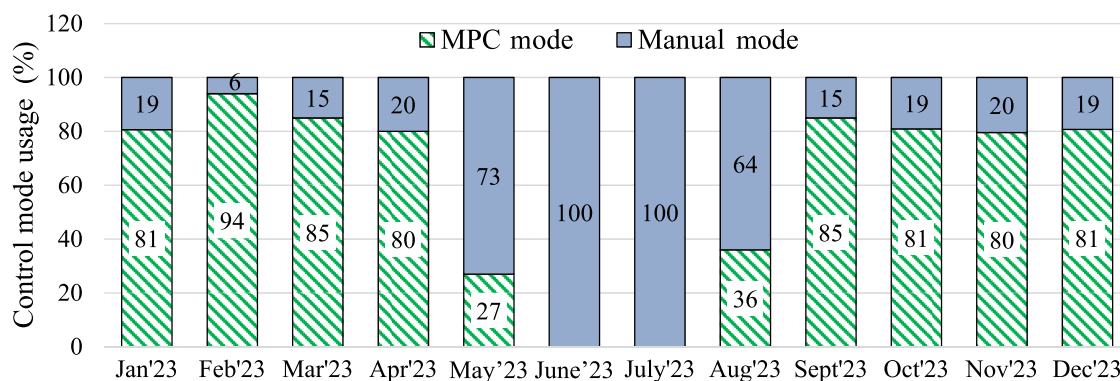
This figure shows that MPC controlled the furnace approximately 85% of the time in a given month. Note that manual mode operation was necessary to address mechanical or electrical issues of furnace hardware, and there was a maintenance period necessitating a major shutdown from May to July 2023.

The greatest impact of the MPC solution is on product quality. In annealing, compliance to the target temperature with minimum variation is the primary objective, while also ensuring throughput and fuel efficiency. Achieving the target temperature ensures making the required microstructural changes in the steel, and hence achieving the desired mechanical properties. Like any manufacturing process, there is a specification limit, which allows some tolerance around the target temperature; in this case it is $\pm 15^{\circ}\text{C}$. If the strip temperature deviates beyond this tolerance, the corresponding strip length is treated as rejected with implications on the process yield. Given the nature of transitions and disturbances that the annealing line experiences, the generation of some amount of material that is outside of

specifications is inevitable especially during grade transitions. However, compared with manual setpoint management during such transitions, optimal setpoint management by the MPC solution considerably reduces the deviation of strip temperature. As Figure 12 shows, for a one-year period both before and after implementation, the MPC-controlled operation resulted in an improvement of 8% (from 77% to 85%) in material that is within specification limits. This equates to salvaging about 13,000 tons of material annually. In addition, whenever such out-of-specification material is produced within a coil, it is reprocessed in another line to remove the noncompliant portion before being dispatched to the customer. After deploying the solution, we experienced a reduction in reprocessing of about 32%, which resulted in eliminating the logistic and operational costs.

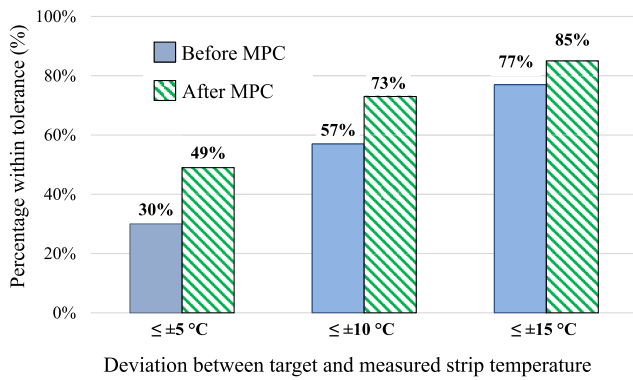
The most significant improvement in quality we observed was the ability of MPC to control the temperature within a narrow band around the target. The strip length with temperature within $\pm 5^{\circ}\text{C}$ increased by almost by 20% from 30% to 49%; see Figure 12. Given the variety of SKUs that the plant processes, the frequency of transitions could be quite high depending on the daily production schedule, thus leading to more product lengths that are outside of specifications. Hence, controlling 50% of the strip length within a very tight band of $\pm 5^{\circ}\text{C}$ and about 75% of strip length within $\pm 10^{\circ}\text{C}$ demonstrates an exceptionally high level of process capability, almost a benchmark. A tighter compliance to the target temperature along the coil length means more uniformity of mechanical properties. Moreover, MPC's ability to control the temperature better than the specification limits implies an improved process capability. In the automotive ecosystem, the ability of steel producers to ensure consistent product quality through better control of the manufacturing process is critical for retaining existing customers and gaining new ones.

Figure 11. (Color online) Monthly Usage of MPC-Based Optimal Control During January–December 2023



Note. A maintenance period necessitating a major shutdown occurred from May to July 2023.

Figure 12. (Color online) Comparison of Strip Temperature Controllability in Different Temperature Bands Around the Target Strip Temperature Before and After the MPC Implementation



Note. We did the comparison for one-year periods both before and after the MPC implementation.

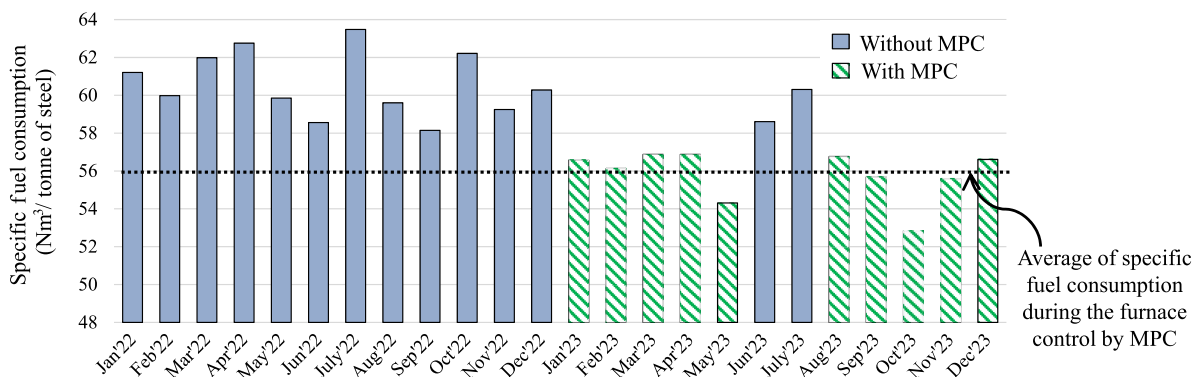
Our legacy system incorporates open-loop control of strip temperatures; as a result, the operator must manually adjust the setpoints during transitions and disturbances. Because of the suboptimality associated with manual setpoint management, which is often reactive, the fuel consumption is high. In contrast, MPC incorporates smooth as well as optimal control action in all operating transients. This prevents coils from becoming underannealed or overannealed and also reduces the specific fuel consumption, which refers to the total fuel consumed per ton of steel produced (see Figure 13). Since the implementation of model-based control in January 2023, we have observed a consistent and significant decrease of 8% per ton of steel in specific fuel consumption. This improvement in fuel efficiency underscores the tangible benefits of adopting model-based control, which has led to more sustainable and cost-effective operations. The reduction in specific fuel consumption not only contributes to environmental sustainability but

also demonstrates the potential for significant cost savings in the production process. When considering Tata Steel’s current installations and those under commissioning, these numbers translate to savings of US\$2.5 million and a reduction of 10,000 tons of CO₂ emissions annually.

Although the reduction in CO₂ emissions may look small in the overall context of CO₂ emissions from the steel industry, it needs to be viewed in the technological and operational constraints of the steel industry as a whole. For almost two centuries, the manufacturing of iron and steel has been dominated by the blast furnace (BF) and basic oxygen furnace (BOF). Although many alternative technologies have been discovered, almost 70% of the steel manufactured globally today is still manufactured using these processes because of their thermal and chemical efficiency, although they are major contributors to CO₂ emissions in steel production. Until we find a viable alternative in terms of scale and economic efficiency, dependence on the BF and BOF will continue to dominate the steel production process; hence, we continue to focus our efforts on reducing CO₂ emissions at every stage in each of the existing processes. In upstream processes like iron making, we focus on reducing usage of raw materials like metallurgical coke, which entails a higher carbon footprint. For downstream processes (e.g., rolling and annealing, which do not have an alternative), we put our efforts into optimizing the process as best we can by using advanced technology for production and control. Digitalization is a key enabler in this respect because advanced analytics and optimization solutions help to closely monitor and operate the lines at the optimal level at all times.

It is in this overall context that the multifaceted benefits of the MPC solution, encompassing both environmental sustainability and enhanced market competitiveness, become significant. The reduction in CO₂ emissions aligns with the company’s commitment to environmental

Figure 13. (Color online) Comparison of Monthly Specific Fuel Consumption Before and After the Implementation of the MPC-Based Furnace Process



Notes. A maintenance period necessitating a major shutdown occurred from May to July 2023. Specific fuel consumption refers to the total fuel consumed per ton of steel produced. Thus, it is a normalized measure of fuel consumption independent of the production volume of any month.

responsibility, whereas the improvement in product quality reflects a dedication to meeting and exceeding customer expectations. This combination of environmental and commercial benefits demonstrates the positive impact of the initiatives undertaken, positioning Tata Steel as a leader in sustainable manufacturing practices and customer satisfaction within the industry.

Implementation Challenges

The implementation challenges were twofold. First, for both software and hardware implementation, it was necessary to seamlessly communicate with the already existing distributed control system (DCS). Second, we had to convince the plant operators and management to accept the change from a decades-old manual operation to an automated optimal control solution. We briefly describe both challenges in this section.

Interfacing with the Legacy Distributed Control System

Implementing the proposed supervisory optimal control solution required a meticulous engineering of the interfaces by taking into consideration the diverse interlocks (i.e., precautionary measures instituted to ensure that a situation never gets out of control, even if that means stoppage of the plant) associated with safety and the inherent complexity of the controller. This process involved significant attention to detail to ensure the seamless integration and functionality of the new control layer with the legacy furnace control system. Our focus was on maintaining the safety protocols and addressing the complexities of the underlying continuous and logic controllers to ensure a smooth and effective transition to the supervisory optimal control operation. It required a comprehensive understanding of the safety interlocks related to combustion in the

radiant tube burners and intricate controller dynamics to facilitate a harmonious integration.

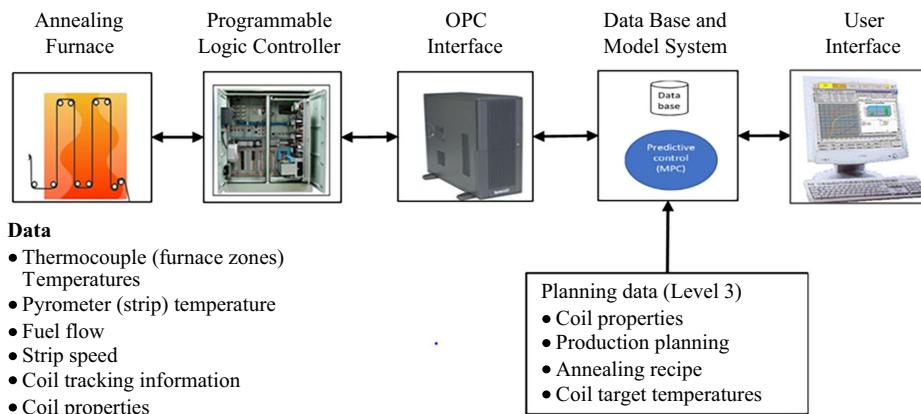
We deployed a real-time data acquisition module by establishing the interface between the PLC and an open platform communications (OPC) server software. OPC is the standard method used for data exchange between a higher-level computing layer and lower-level controllers (e.g., PLCs, DCSs). Figure 14 shows an overview of the system architecture. The PLC located in the basic automation system (Level 1) receives the measured values from various sensors and also includes the subordinate controllers for the actuator. The PLC provides the material tracking and the measured process values; examples include strip temperature, furnace temperatures, mass flow of fuel, strip velocity, and tracking information of leading and trailing coils. The data acquisition module reads the data from the PLC through an OPC client application every three seconds and shares these data with the MPC application.

The MPC calculation runs once each minute in the supervisory compute layer (Level 2) using time-averaged values of process values as inputs.

Change Management

After integrating the solution, the transition from manual operations to automatic control required a gradual and structured change management process. This process involved conducting multiple training and knowledge-sharing sessions with frontline operators and line managers, considering that manual operations had been the standard practice for two decades. The trials initially spanned a few hours, and we shared the performance analysis results with the operating team. Based on the team's feedback, we fine-tuned the solution, gradually increasing the duration of uninterrupted automatic control to cover multiple shifts over multiple days. By January 2023, model-based automatic control

Figure 14. (Color online) The Architecture of the Integrated MPC-Based Cascade Control System Highlights the Information Exchange Between the MPC and the Legacy Honeywell System Through the OPC Server



with average monthly usage of 85% had become the norm; manual operation was used only during exceptional conditions. This transition marked a significant shift in operational practices, necessitating careful planning and communication to ensure a smooth adoption of automatic control. The gradual increase in the duration of uninterrupted automatic control reflects the successful adaptation to the new operational paradigm, ultimately leading to the establishment of model-based automatic control as the standard practice.

This systematic transition not only demonstrates the adaptability of the operational workforce but also highlights Tata Steel's commitment to embracing technological advancements for improved efficiency and performance in industrial operations. It is important to note that this transition also led to improvements in operational efficiency and reliability. Freed from having to frequently manually adjust the setpoints, the operations team members were able to do more analysis and initiate improvement processes. Furthermore, the comprehensive training and knowledge-sharing sessions contributed to a more informed and skilled workforce, capable of effectively managing and maintaining the new automated system.

Future Work and Broader Implications

Currently, at the Tata Steel facility in Kalinganagar, Odisha, we are in the process of commissioning three similar production lines. As a result, our goal is to seamlessly integrate this innovative solution into the control system for these lines. Leveraging our newly acquired capability to develop predictive control systems using data-driven techniques, we have the potential to apply this methodology to a wide range of manufacturing processes within the steel industry. Notably, this approach is particularly well suited for addressing the complexities of industrial processes such as the BF (for iron making) and coke ovens (for making metallurgical coke), where traditional first-principles modelling methods may prove economically unattractive because of the costs associated with model development. By leveraging data-driven predictive control, we are positioned to significantly enhance the operational efficiency and performance of these vital industrial processes in the steel industry. This shift toward intelligent automation also underscores a broader industry trend toward embracing digitalization and advanced control systems to drive productivity and competitiveness. The successful adoption of model-based automatic control serves as a testament to Tata Steel's commitment to pioneering improvements and innovation in its manufacturing processes.

Conclusions

The continuous annealing furnace is a crucial unit operation in steel production, which recovers the formability

and ductility of cold-rolled steel strips. Controlling this process is challenging because of complex furnace temperature dynamics, frequent steel grade changes, and multivariable interactions within the furnace zones. To achieve optimal control of the strip exit temperature, we developed a novel MPC technology-based solution and successfully implemented it on the continuous annealing furnace that produces automotive grade steel at Tata Steel, India. The systematic implementation of the solution has led to a significant improvement in the proportion of annealed products meeting the premium quality standard, and substantial reductions in fuel consumption and CO₂ emissions. This achievement demonstrates the positive impact of integrating advanced control systems in industrial operations, highlighting the potential for improved resource efficiency and environmental stewardship within the manufacturing sector.

Acknowledgments

The authors express their gratitude to Carrie Beam, Manoj Chari, and Frederick Zahn for their coaching in the preparation and final presentation of this paper. Their time and guidance are sincerely appreciated. Special thanks to Anil Pujari (chief of manufacturing, Tata Steel) and K. P. Madhavan (professor emeritus, IIT Bombay) for providing constructive feedback and support, which were of great help during the development and implementation phases.

Appendix A. Development of a Control-Relevant Dynamic Model

Let $y(t)$ and $u(t)$ define perturbation output and input variables, respectively, in the neighbourhood of some steady state operating point. Then, a second-order Laplace transfer function model with input delay relating $y(s)$ and $u(s)$ is represented as $y(s) = G(s) u(s)$ (Åström and Hagglund 1995), where

$$G(s) = \frac{\beta(\tau_z s + 1)\exp(-ds)}{(T^2 s^2 + 2\tau\xi s + 1)}. \quad (A.1)$$

Here, β represents steady state gain, where the inverse of τ_z represents a zero of the system, τ represents a time constant, ξ represents a damping coefficient, and d represents the time delay. In time domain, this model can be represented as a second-order differential equation with an input delay (i.e., second order ODE-ITD) as follows (Åström and Hagglund 1995):

$$\tau^2 \frac{d^2 y}{dt^2} + 2\tau\xi \frac{dy}{dt} + y(t) = \beta \left(\tau_z \frac{du(t-d)}{dt} + u(t-d) \right). \quad (A.2)$$

Here, $(\tau, \xi, \beta, \tau_z, d)$ are parameters of the model, which are estimated using perturbation data. In a similar manner, first-order or third-order transfer ODE-ITD single-input, single-output (SISO) models can be derived (Åström and Hagglund 1995).

In the context of the furnace temperature model, consider perturbation output and input vectors defined as follows:

$$Y_f(t) = \begin{bmatrix} Y_{f,1}(t) \\ Y_{f,2}(t) \\ Y_{f,3}(t) \end{bmatrix} = \begin{bmatrix} \hat{T}_{f1}(t) - T_{f1,s} \\ \hat{T}_{f2}(t) - T_{f2,s} \\ \hat{T}_{f3}(t) - T_{f3,s} \end{bmatrix},$$

$$U(t) = \begin{bmatrix} U_1(t) \\ U_2(t) \\ U_3(t) \end{bmatrix} = \begin{bmatrix} T_{sp1}(t) - T_{f1,s} \\ T_{sp2}(t) - T_{f2,s} \\ T_{sp3}(t) - T_{f3,s} \end{bmatrix},$$

$D(t) = Q(t) - Q_s$. Here, $(T_{f1,s}, T_{f2,s}, T_{f3,s})$ represent a steady state operating point of the furnace temperatures, and Q_s represents a steady state steel mass flow rate. A MIMO model relating the furnace temperature outputs with furnace PID setpoints (inputs) is represented as the following matrix equation in the Laplace domain:

$$Y_f(s) = \begin{bmatrix} G_{11}(s) & 0 & 0 \\ G_{21}(s) & G_{22}(s) & 0 \\ G_{31}(s) & G_{32}(s) & G_{33}(s) \end{bmatrix} U(s) + \begin{bmatrix} G_{1Q}(s) \\ G_{2Q}(s) \\ G_{3Q}(s) \end{bmatrix} D(s). \quad (\text{A.3})$$

Here, each $G_{ij}(s)$ represents a first/second/third-order SISO transfer function and equivalently a first/second/third-order ODE-ITD relating j th input with i th output. Similarly, defining perturbation output $y(t) = T_s(t) - T_{s,s}$, the strip temperature model is proposed as follows:

$$y(s) = [G_{s1}(s) \ G_{s2}(s) \ G_{s3}(s)] Y_f(s) + G_{sQ}(s) D(s), \quad (\text{A.4})$$

where the output vector $Y_f(s)$ of the furnace temperature model is an input to the strip temperature model along with $D(t)$; also see Figure 6. Again, each $G_{ij}(s)$ represents a SISO first/second/third-order transfer function or equivalently a first/second/third-order ODE-ITD. Parameters $(\tau_{ij}, \xi_{ij}, \beta_{ij}, \tau_{z,ij}, d_{ij})$ of each SISO transfer function in models (A.3) and (A.4) are estimated using perturbation data generated by conducting perturbation experiments on the plant and excitation data obtained from historical records and MATLAB's System Identification Toolbox. The furnace and strip temperature models, (A.3) and (A.4), are subsequently converted into digital control-oriented discrete-time state-space models and further combined into a single-state-space model of the form

$$X(k+1) = \mathbf{A} X(k) + \mathbf{B} U(k) + \mathbf{H} D(k) + w(k), \quad (\text{A.5})$$

$$Y(k) = \mathbf{C} X(k) + v(k), \quad (\text{A.6})$$

using a sampling interval of $T = 1$ minute. Here, X represents the state vector, U represents the vector of manipulated inputs (three furnace zone temperature setpoints), D is disturbance input (i.e., perturbations in steel mass flow rate), $Y = \begin{bmatrix} Y_f \\ y \end{bmatrix}$ represents the controlled and measured outputs (i.e., three furnace region temperatures and the strip exit temperature), w represents uncertainties in the state dynamics, v represents the measurement noise, and index k represents a discrete sampling time instant (k represents the present, $k+1$ is one sample in the future, and $k-1$ represents one sample in the past).

The dimension of the state vector is determined by the orders of the denominator polynomials and time delays in the transfer function model (A.3). Because of the large time delays appearing in the identified transfer function model for the annealing furnace, the dimension of the X vector is 84×1 , and \mathbf{A} , \mathbf{B} , \mathbf{C} , and \mathbf{H} are matrices of dimension (84×84) , (84×3) , (84×1) , and (4×84) , respectively. Details of the continuous time model to discrete-time model conversion and state-space realization of ODE-ITD models can be found in Åström and Wittenmark (1997).

Appendix B. Kalman Filter and Model Predictive Control

To formulate an MPC scheme using a state-space model, it is necessary to estimate the current state, $X(k)$, of the system. Let us denote it by $X_e(k|k)$, that is, an estimate of $X(k)$ using information up to instant k . In this work, we have chosen to use the Kalman filter, which minimizes the covariance of the state estimation error (Åström and Wittenmark 1997, Patwardhan et al. 2006). The main idea is to use the state-space model for predicting the output using model (B.1) and information available up to instant $(k-1)$, that is, $Y(k|k-1) = \mathbf{C} X_e(k|k-1)$, and use the model residuals, $e(k) = Y(k) - Y(k|k-1)$, to correct the predicted state estimates using a gain matrix, \mathbf{L} , together with the residual signal as follows:

$$X_e(k|k-1) = \mathbf{A} X_e(k-1|k-1) + \mathbf{B} U(k-1) + \mathbf{H} D(k-1), \quad (\text{B.1})$$

$$X_e(k|k) = X_e(k|k-1) + \mathbf{L} (Y(k) - \mathbf{C} X_e(k|k-1)). \quad (\text{B.2})$$

The Kalman gain matrix, \mathbf{L} , is computed by solving algebraic Riccati equations; see Åström and Wittenmark (1997) for details. The filtered state estimate, $X_e(k|k)$, is then used for the MPC formulation. Moreover, the residual signal, $e(k)$, also carries information about the model plant mismatch and unmeasured disturbances. This signal is filtered to remove the high-frequency noise as follows: $e_f(k) = \lambda e_f(k-1) + (1-\lambda)e(k)$, where $0 < \lambda < 1$ is a tuning parameter used for achieving robustness with respect to the model plant mismatch and additionally used in MPC predictions. At instant k , given a future sequence of manipulated input guess vectors $\{U(k|k), U(k+1|k), \dots, U(k+P-1|k)\}$, we can predict the future behaviour of the system as follows:

$$Z(k+j+1|k) = \mathbf{A} Z(k+j|k) + \mathbf{B} U(k+j|k) + \mathbf{H} D(k+j) + \mathbf{L} e_f(k), \quad (\text{B.3})$$

$$Y(k+j+1|k) = \mathbf{C} Z(k+j+1|k) + e_f(k), \quad (\text{B.4})$$

for $j = 0, 1, 2, \dots, P-1$ starting with the initial condition $Z(k) = X_e(k|k)$. Now, given a strip target strip temperature setpoint trajectory, $r(k+j|k)$, we define future prediction error as

$$E(k+j) = r(k+j|k) - y(k+j|k) \text{ for } j = 1, 2, \dots, P$$

and deviation inputs as $\Delta U(k+j) = U(k+j|k) - U(k+j-1|k)$. The MPC problem at sampling instant k is formulated as a

constrained optimization problem as follows (Patwardhan et al. 2006):

$$\arg \min_{\{U(k|k), \dots, U(k+P-1|k)\}} \sum_{j=1}^P (w_E E(k+j)^2 + \Delta U(k+j-1)^T W_U \Delta U(k+j-1))$$

subject to prediction models (B.3) and (B.4) together with $Z(k) = X_e(k|k)$:

$$\begin{aligned} U_L &\leq U(k+j|k) \leq U_L \text{ for } j = 0, 1, 2, \dots, P-1, \\ \Delta U_L &\leq \Delta U(k+j|k) \leq \Delta U_L \text{ for } j = 0, 1, 2, \dots, P-1, \\ Y_L &\leq Y(k+j|k) \leq Y_L \text{ for } j = 1, 2, \dots, P. \end{aligned}$$

Here $w_E > 0$ and W_U is a positive definite tuning matrix, and P represents the prediction horizon. The inequality constraints specify limits on inputs, rate of change in inputs, and predicted outputs over the future horizon. Using multiple algebraic manipulations, this optimization problem is converted into a QP problem and solved using the *quadprog* function in MATLAB's Optimization Toolbox. Only the optimal input $U(k) = U^*(k|k)$ is implemented on the plant, and the problem is reformulated and resolved over the time window $[k+1, k+P+1]$ (i.e., moving horizon control).

References

- Åström KJ, Hagglund T (1995) *PID Controllers: Theory, Design and Tuning*, 2nd ed. (Instrumentation Society of America, Durham, NC).
- Åström KJ, Wittenmark B (1997) *Computer Controlled Systems: Theory and Design*, 3rd ed. (Prentice-Hall, Upper Saddle River, NJ).
- Avner SH (1974) *Introduction to Physical Metallurgy*, 2nd ed. (McGraw Hill, London).
- Bitschnau L, Jakubek S, Kozek M (2010) Constrained model predictive control of a continuous annealing furnace. *Proc. ASME 2010 Dynamic Systems Control Conf.* (American Society of Mechanical Engineers, New York), 285–292.
- Colvin FH, Juthe A (2007) *The Working of Steel Annealing, Heat Treating and Hardening of Carbon and Alloy Steel*, 2nd ed. (McGraw Hill, New York).
- Imose M (1984) Heating and cooling technology in continuous annealing. *Trans. Iron Steel Inst. Japan* 25(9):911–932.
- Ljung L (1999) *System Identification: Theory for the User*, 2nd ed. (Prentice Hall PTR, Upper Saddle River, NJ).
- Martineau S, Burnham K, Haas O, Andrews G, Heeley A (2004) Four-term bilinear PID controller applied to an industrial furnace. *Control Engng. Practice* 12(4):457–464.
- Mould PR (1982) An overview of continuous-annealing technology for steel sheet products. *J. Metals* 34(5):18–28.
- Niederer M, Strommer S, Steinboeck A, Kugi A (2016) Nonlinear model predictive control of the strip temperature in an annealing furnace. *J. Process Control* 48(December):1–13.
- Patwardhan SC, Manuja S, Narasimhan S, Shah SL (2006) From data to diagnosis and control using generalized orthonormal basis filters. Part II: Model predictive and fault tolerant control. *J. Process Control* 16(2):157–175.
- Qin SJ, Badgwell TA (2003) A survey of industrial model predictive control technology. *Control Engng. Practice* 11(7):733–764.
- Strommer S, Niederer M, Steinboeck A, Kugi A (2018) Hierarchical nonlinear optimization-based controller of a continuous strip annealing furnace. *Control Engng. Practice* 73(April):40–55.
- Takeo F (1985) Recent technological progress in high-speed continuous annealing. *Trans. Iron Steel Inst. Japan* 25(4):278–293.
- Wu H, Speets R, Heeremans F, Ben Driss O, van Buren R (2015) Non-linear model predictive control of throughput and strip

temperature for continuous annealing line. *Ironmaking Steelmaking* 42(8):570–578.

- Yoshitani N, Hasegawa A (1998) Model-based control of strip temperature for the heating furnace in continuous annealing. *IEEE Trans. Control Systems Tech.* 6(2):146–156.

Sujit A. Jagnade is a senior technologist of intelligent systems and mathematical modelling group in the Automation Division, Tata Steel Limited, India. He obtained his masters in chemical engineering from the Indian Institute of Technology Kanpur, India, in 2017. His work mainly focuses on optimization, advanced process control, and process modelling, employing both physics-based and data-driven (artificial intelligence/machine learning) techniques to enhance real-time manufacturing processes.

Sachin C. Patwardhan is a professor of chemical engineering and an associate faculty of systems and control engineering at the Indian Institute of Technology (IIT) Bombay. He received his PhD in systems and control engineering from IIT Bombay in 1994. His research interests are in the application of artificial intelligence/machine learning techniques for developing data-driven dynamic models, adaptive and model predictive control, online fault diagnosis, and real-time economic optimization.

Kunal Kumar is a research engineer in the Advanced Research Center division of GE Vernova, India. He earned his PhD in system and control engineering from the Indian Institute of Technology Bombay in 2021. His research focuses on applying artificial intelligence and machine learning techniques for soft sensor development in process control, graphical modelling, and dual model predictive control.

Arup K. Dey is a senior manager (scheduling and delivery) in Integrated Planning and Services group of Supply Chain Division, Tata Steel, India. He holds a diploma in electrical engineering from the State Board of Technical Education. His key experience is in the operations of continuous annealing, metal coating, and cold rolling.

Sai K. Gudimetla is head (operations) of the continuous galvanizing line of the Cold Rolling Mill Complex at Tata Steel, Jamshedpur, India. He obtained his MTech in industrial engineering from Nagpur University, India, in 2008, and his MBA in marketing from XLRI Jamshedpur, India, in 2012. His professional experience is primarily on galvanizing and continuous annealing.

Manish K. Singh is the general manager of design and engineering at Tata Steel, India. He obtained his BSc in electrical engineering from the National Institute of Technology, Jamshedpur, India, in 1989. He has extensive experience in integrated electrical maintenance, process automation, digitalization, design, and engineering in the steel industry.

Ajay K. Jha is chief of the Cold Rolling Mill Complex at Tata Steel, Jamshedpur, India. He completed his BTech in mechanical engineering at Bangalore University in 1994. He has close to three decades of experience in operations of cold rolling and downstream processing in the steel industry.

Gyan Prakash is chief of automation, One IT, at Tata Steel, India. He received his BTech in electrical engineering from the National Institute of Technology, Jaipur, India, in 1995. Currently he leads the digitalization in manufacturing at Tata Steel India, and his past experience includes integrated electrical maintenance and process automation in the steel industry.

Jose M. Korath is chief of intelligent systems and mathematical modelling group in the Automation Division, Tata Steel, India. He received his masters in metallurgy from the Indian Institute of Science, Bangalore, India, in 1994 and PhD in process system engineering from the University of Sydney, Australia, in 2009. He works on the development of analytic solutions for manufacturing processes using classical modelling techniques and advanced artificial intelligence/machine learning algorithms.