Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 55-40 (2022) 295–300

From controlling single processes to the complex automation of process chains by artificially intelligent control systems: the ControlInSteel project

Marcus J. Neuer [*,**], **Moritz Loos** [*] **Francesca Marchiori** [****]
Valentina Colla [***] **Stefano Dettori** [***] **Joaquin Ordieres** [†]
Andreas Wolff [*]

[*] *VDEH Betriebsforschungsinstitut*
[**] *RWTH Aachen*
[***] *Scuola Superiori Sant'Anna*
[****] *RINA Consulting Centro Sviluppo Materiali*
[†] *Universidad Polytecnica de Madrid*

Abstract: The ControlInSteel project, a cooperation of four research institutes, revisited research projects of the last 20 years focusing on automation and control solutions applied to the downstream steel production route. During this investigation we found hints to those solutions, which were beneficial for specific problems. For our analysis, 46 projects were systematically reviewed. Taxonomies for the problem space, the solution space, the barriers and issues and the impact were developed and each project categorized along these taxonometrical dimensions. As a result, the interdependencies between solutions and impact could be analysed in a quantifiable way, which led to a new way of evaluating project success. It also brought new insights about the most promising techniques already applied and those techniques, that have been apparently not yet been applied to steel production, although being highly successful in other domains. This leads to potential future research chances for the steel production and their complex process chains. The paper will also finally demonstrate how a similar taxonometrical approach can be used to conserve expert knowledge in automation to feed a truly artificially intelligent control solution – not only exploiting machine learning methods but essentially using machine reasoning on top of the digitized expert knowledge to achieve improved process automation.

Keywords: Knowledge-based control, Data-fusion, Cyber-physical systems, Adaptive system and control

1. INTRODUCTION

Facing ecological, economical and even political challenges, the mission of modern automation becomes more diverse, facetted and difficult. Throughout the past years, the European Union financially funded research in steel processing via the Research Fund for Coal and Steel (RFCS) programme. Many projects conducted within this context focused on innovative automation and control solutions in the downstream process chain where steel is in its solid state. In retrospective, those projects also took place during the fundamental disruptions brought by increasing digitalisation, rise of Big Data and finally the industrial scope application of artificial intelligence. All these revelations provided the foundations and technologies for new innovations in process automation as well. Here, the dissemination project ControlInSteel comes into play. This project ran from July 2020 till December 2022 and reviewed two decades of the the most influential automation and control research in steel industry. Thus, it covered the relevant period in time, where these new technologies emerged and de-facto provided real impact at industrial sites for the first time.

The ControlInSteel project used a semantic analysis of the conducted projects. It considered the problem space T_1, the solution space T_2 (containing the according automation solution) and lastly the achieved impact T_3. Each T_λ is a mathematical set (where λ abbreviates 1,2 or 3). This allows us to map the projects that applied control methods onto the different T_λ. On the basis of these semantic sets, we can identify which method led to greatest impact in steel production. Additionally, we can find methods with high potential, never being tried out in steel research. In the course of the present paper, we demonstrate how this analysis was performed and to which interpretations it leads. In short, we have the following objectives:

Objective 1: to analyze the interdependencies between aggregates, solutions and impact.

Objective 2: to quantify how much impact was generated by a specific technique, which is equivalent to establish a

10.1016/j.ifacol.2023.01.088

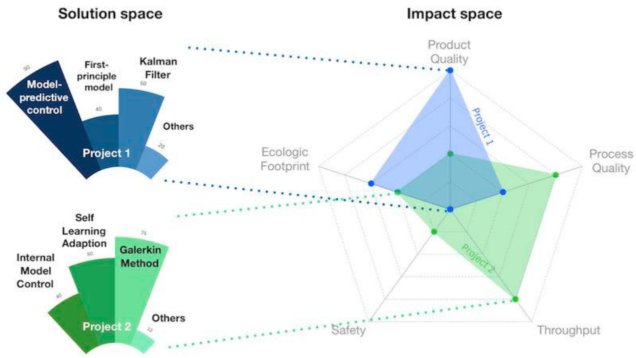

Fig. 1. Schematic of the envisioned analysis results: using the project analysis to connect control methodology with impact metrics.

conditional probability distribution

$$P = P(\text{"Emission Reduction"}|\text{"Hot Rolling"}, \text{"MPC"})$$

on top of our taxonomy sets. This allows to quantify the de-facto generated impact as a function of the used control method. Fig. 1 shows a schematic view on this mapping approach. In Sec. 2, we explain further details of the method. After this introduction, the paper will present theoretic considerations in Sec. 2, followed by details on the problem and solution space in Sec.3. Sec.4 focuses on the impact analysis and Sec. 5 concludes with the most important results optained by our approach.

2. THEORETIC BACKGROUND

2.1 Semantic extraction

Semantic approaches have some history in steel industry, whereas it is safe to say, that semantics have though never been fully embraced by this industry yet. For analysing research, Zhu et al. (2022) present a recent work on the where they used a semantic approach that integrated expert reliability. In contrast, our present work covers a rather retrospective analysis of already finished (or nearly finished) projects - thus, taking advantage of the reported information at the end of a project.

In Fig. 1 we show an abstract schematic about how projects methodology can be mapped on impact. Similarly (not shown here) one can map the problem space or the barrier spaces. What is necessary to extract this information?

- Preparation: First, you have to define some closed set of works, in our case a defined set of 46 research projects (which can be found at our webpage www.controlinsteel.com), that clearly reported on methodology, problems, impact and transfer.
- Vocabulary synchronisation: One would assume that control theory usually uses identical terminologies, but in fact, multiple synonyms exist and there are sevaral occasions, where the vocabularies of different works have to be carefully synchronised. One example is the use of the principal component analysis (PCA) which is often also termed as Karhunen-Loeve transform. This is just one of many examples here.

Fig. 2. Mindmap of the problem space taxonomy T_1, including channels of interaction.

Fig. 3. Impact dimensions and the overall distribution $P(j, T_3)$ of the reviewed automation and control projects.

- Taxonomy development: Starting with such a consolidated and synchronised vocabulary, we generated taxonomies for the different spaces: for the problem space, the solution space and the impact space. The previously mentioned PCA is an example from the solution space. Some taxonomies required also the derivation of criteria.
- Taxonomical mapping of the projects: In this step, we went through all projects and allocated them to the taxonomies.

Based on the taxonomies apply rules of ontologic reasoning, to inversely map whether the elements of the taxonomies are linked with each other. This requires first, to get a formal description of our semantic sets. Let $\Pi = \{0, 1, 2, ..., 45\}$ be the set of numbers marking our $N = 46$ projects. T_λ is any of our taxonomies, where λ is defined as in Sec.1. The taxonomy operator $\mathcal{T}_\lambda$, retrieves the projects feature vector when mapped onto the taxonomy dimensions,

$$\mathbf{f}_\lambda(i) = [(f_{\lambda 0}(i), f_{\lambda 1}(i), ...] = \mathcal{T}_\lambda(i), \tag{1}$$

where $i \in \Pi$ is a project, and $f_{\lambda j}(i)$ is the associated feature weight of project i within the taxonomy T_λ regarding the jth entry in T_λ. For extracting a singular element j of T_λ for project i we write for convenience,

$$f_{\lambda j}(i) = \mathcal{T}_\lambda(i, j). \tag{2}$$

For the problem space and the solution space these values f are either 0 or 1: formally $f_{\lambda,i} \in [0, 1]$ for $i \in T_\lambda$, if $\lambda = 1$ or $\lambda = 2$. For the impact space, where we use a floating point distribution, $f_{3j} \in \mathbb{R}$ for each $j \in T_3$. The straightforward distributions can then be obtained by summing through the necessary sets,

$$P(j, T_\lambda) = \sum_{i \in \Pi} \sum_{k \in T_\lambda} f_{\lambda j}(i) \delta_{kj} \tag{3}$$

using the Kronecker notation to formalise counting the contribution of the jth entry to the overall group of all projects. Similarly, the conditional distribution can be obtained for $\rho \neq \lambda$ by

$$P(j, T_\lambda | m = 1 \text{ with } m \in T_\rho) = \sum_{i \in \Pi} \sum_{k \in T_\lambda} \sum_{r \in T_\rho} f_{\lambda j}(i) \delta_{kj} \delta_{mr}, \tag{4}$$

In practise, we store the taxonomies in JSON documents. The operation (1) and the loops (3) and (4) are performed in Python on top of an in memory database containing the document-oriented data in Apache Parquet. For the visualization of graphs we use the pyvis network library. One further aspect is the use of linkage between the taxonomies, which we model with graphs. We first start with the internal taxonomy edges, modelled by the mapping ϕ_λ

in which all entries that belong to project i are given by taxonomy set T_λ,

$$\phi_\lambda : E_\lambda(i) \to \{(f_{\lambda j}, \lambda) | f_{\lambda j} = \mathcal{T}_\lambda(i,j)\} \tag{5}$$

and where the node set contains all entries and a center point abbreviately written λ. This is a star-form graph that is very simple and helpful for evaluation. We connect all taxonomy entries for a project i on their top level, which means the trivial graph ϕ_T

$$\phi_T : E_T(i) \to \{(\lambda, \pi(i)) | \lambda \in (1,2,3) \wedge \pi(i) \in \Pi\}\,, \tag{6}$$

modeling a second star-form graph for the top-level edges. Here, $\pi_\lambda \forall \lambda$ and ϕ_T describe formally any project i.

3. ANALYSIS OF PROBLEM AND SOLUTION SPACE

3.1 Archetypes of control problems in steel processing

Of course, most projects circle around two archetypes of problems:

Type 1. Open or closed-loop control of a singular process where Jelali et al. (2002) is only one example of works that initiated a tremendous series of activities monitoring the health of control loops as shown in Jelali (2000), and Jelali (2005) and and consequently optimizing their performance presented in Jelali (2007).

Type 2. Control of (often complex) holistic systems, mainly characterised by considering a subbranch of the production street or the whole street. This type can be further divided into

(a) **Process chain optimisation** for the overall product flow, which was addressing reallocation problems in Neuer et al. (2016) and later led to yield improvement as shown by Iannino et al. (2021). Most recently, even cyber-attacks on control systems and the production route are considered, where attacks can be detected by appropriate machine learning techniques as shown by J. Ordieres-Mere (2022) and Neuer et al. (2019).

(b) **Network optimisation problems**, which are the foremost solution to the challenge of reducing emissions, waste, and to distribute available ressources effectively. Here the works of Matino et al. (2019), Dettori et al. (2022) and Wolff et al. (2019) demonstrated the successful application of mixed-integer programming to this class of problems.

3.2 Canonical control solutions

In its most canonical form, control theory uses feedforward and feedback loops to interfere with a given process, in order to achieve some desired system state. Reviewing all project reports, we see a technological advancement in the control methodology that was clearly sparked by the inclusion of model knowledge or otherwise prior knowledge on the system. While this is quite reasonable, we found the surprising fact of our analysis that stochastic approaches, especially those theoretical concepts that involve the inclusion of stochastic processes and uncertainty in the control process, are dramatically underrepresented. This coincides with a general observation among most considered projects, that elements of uncertainty quantification especially as it is common standard in natural sciences, was rarely adopted within these research projects. Even those projects that essentially relied on database techniques, often discarded inductive statistics or probabilistic treatments.

One clear example is the adoption of Bayesian statistics, especially Bayesian inference which was discussed for upstream steel processing and amongst others demonstrated by Klimes et al. (2011). In their influential paper, Leitão and Restivo (2006) showed how distributed problems like the process chain could be treated by autonomous agents - yet, without touching details of the individually involved control systems. It is important to note, that such type of meta-control approach requires in its most general treatment a complex interplay of different individual canonical process control loops. Such an complete treatment was not touched by any research project that we evaluated. Indeed, most projects that focused on distributed problem solving selected one or two central processes of most interest and built distributed use cases around those primary processes to study and research the dynamics. Summarizing, this class of solutions has not been utilized to full capacity yet. Consequently, we would expect that a true holonic control of the overall process chain should be focused by future research. We also reviewed most recent technologies of selected plant builders, which embraced the digitalisation trend and even prepared demonstration sites for new prototypes. Still, the full scope as envisioned by holonic manufacturing was not achieved so far.

4. IMPACT ANALYSIS

4.1 Impact dimensions

The analysis of the impact is the most crucial part of our work. It was performed as sketched in Sec. 2, primarily based on reviewing the project reports and some selected interviews. In the center of analysis, we put the definition of a suited set of impact dimensions to structure and quantify the impact. We identified following impact dimensions:

- Quality improvement
- Defect root cause (eliminated by better control)
- Cost reduction
- Yield improvement
- Power consumption (reduction)
- Waste reduction
- Emission reduction
- Worker safety
- Worker performance
- Customer satisfaction
- Enabling technology (iterative)
- Novel approach (breakthrough)
-

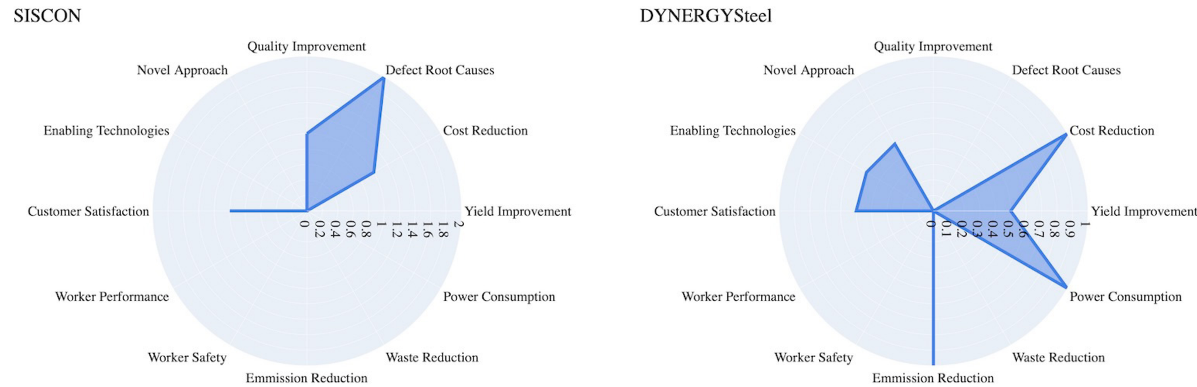

Fig. 4. Assorted projects and their impact evaluation charts plotted as radar charts as indicated abractly in Fig.1. Note, the projects cannot be rated against each other, as we normalised the maximum total impact - being interested in the distribution among the dimensions.

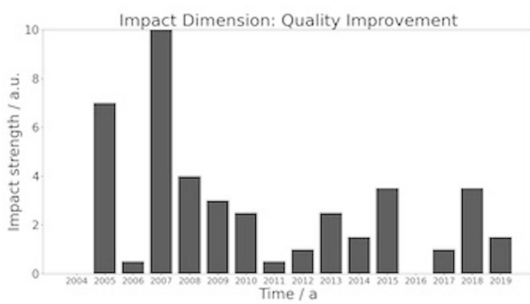

Fig. 5. Impact of quality improvement as a function of time. While being in the focus in the beginning, projects focusing this topic decreased over time.

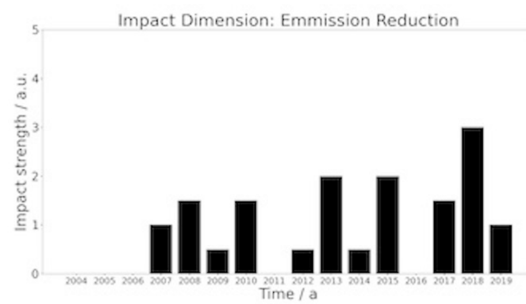

Fig. 6. Impact "emission reduction reduction" as a function of time. This slowly increases over the years.

where indeed some dimensions appear to be correlated, as quality improvement and customer satisfaction suggest. In fact, each dimension features an inherent uniqueness, which could not be covered by combinations of the other axis. The dimensions were choosen to optimize taxonometric orthogonality as good as possible. Figure 4 shows a selection of six projects and their corresponding radar plots. All projects are diverse in their impact distribution. Some results are straightforward, as SISCON concentrated on surface inspection, it is clear that it impacted root causes for defects and customer satisfaction. For others, like DynergySteel or PUC, emission reduction and waste reduction were achieved.

4.2 Discussion of results of impact scoring

We analysed all projects and mapped a fixed impact score of 5 to all projects, according to the appropriate distribution of impacts. Of course, this normalisation assumes that each project contributed the same maximum impact. Contrarily, one could have introduced a flexible scheme to judge the projects. This, however, would lead to subjective rating of projects that was not the intention of the analysis. With our approach, we extract just the distribution of impact. The resulting, overall impact distribution for all projects is shown in Figure 3. From this analysis it is clear, that the most impact was generated for quality improvement, closely followed by yield improvement. Neither emission reduction nor power consumption are among the most dominant topics.

A first obvious conclusion is to strengthen efforts for these latter impact fields. Without explicitly including newer projects by now, it can be seen that the political activities of the European Union, namely the Green Deal, are increasing the impact in these dimensions. The newest abstracts (2022) of recently started projects supports this assumption. Figure 3 also reveals missing efforts for worker safety and worker performance. Both fields could significantly benefit from advanced automation, yet both are dramatically underrepresented in steel research up to now. With respect to worker performance, it is difficult to comprehend, why so few projects actually covered robotics and robotic process automation (RPA) that supports the work of human staff. This field is at the heart of process control, see e.g. Edlich et al. (2019), having the capability to significantly improve worker performance. We see a clear need to improve in this field to keep steel production competitive. This need is emphasised by the fact that also Industry 4.0 and digitalisation requires a constant (re-)training of the people involved in the production process. In Bughin et al. (2018) this was also identified as key challenge for automation as such.

Fig. 7. Autogenerated python graph visualisation of the CEFLA project - as one example of the graph approach given by combining the graphs of (6) and (5) for the different taxonomy sets.

Closely related to worker performance is the field of worker safety. Only few projects actually touched this topic. Many projects mentioned the objective to address worker safety, but rarely, automation and control concepts were dedicated to provide a solution. Mostly, the impact on worker safety was secondary in nature, rather a welcome side-product of some more important technological advancements.

4.3 Time development of impact

The accumulated picture is only one perspective. Impact can also be analysed as function of time. Formally, this means any project $i \in \Pi$ is linked to a specific time t_i. Both, start date or end date can be used here, where we decided to consider only the start year. This simplifies the treatment. In Figure 5 this is shown for the quality improvement. Quality improvement was a mayor aim in 2000 and following years. But it decreased over the following years. During this time, customer satisfaction, yield improvement and mostly production performance oriented topics dominated research projects. In contrast, Fig. 6 the impact on emission reduction is shown over the years. This impact, with exceptions, rises over time. It shows that the steering mechanism of the research funding and the work of involved evaluation experts, indeed helped to develop a strategy for steel research towards higher sustainability in production. Examples include complex control strategies for off-gas and steam distribution as shown e.g. by Colla et al. (2018).

5. RESULTS OF THE SEMANTIC DEDUCTION

5.1 Semantic model per project

As we stated in the introduction and in Sec. 2, one of our aims is to quantify the impact as function of the model. With the established amount of data produced by evaluating the projects we can now first derive a representation model for each project. In Figure 7 we show an example result for the one concise project, namely the CEFLA project. This project mainly applied model-predictive control and internal model control. The projection indicates, how these techniques impacted the steel production, in this case a cold-rolling problem regarding flatness. All project results would exceed the scope of this publication, but they can be found at our webpage www.controlinsteel.com, when opening the interactive project database.

5.2 Deduced impact per method

Figure 8 shows the accumulated methodological impact generated for two assorted methods IMC and Agent-based optimization, combined over all projects in the database using eq. (4). Nonetheless, Fig. 8 reflects the probability of achieving a certain impact dimension, when applying the selected method (at least for the research projects in steel industry).

Only those projects, that covered optimisation techniques like quadratic programming or mixed-integer solvers, were impacting emission reduction, waste reduction and power consumption.

5.3 Discussion on relevance of semantically stored knowledge for artificially intelligent control solutions

Artificial intelligence, meaning to depart from pure machine learning modelling towards truly mimicking human solution strategies, will require any AI system to have access to expert knowledge - here, the taxonomies and impact results we evaluated from the research projects are such an knowledge reservoir. Fig. 6 also highlights a success of the combination of optimization techniques and machine learning for control problems (here for emission reduction). Those techniques are clearly impacting ecologic improvement over the last years and should be treated as highly relevant for the field.

This knowledge is now available for an AI system. It can use the probability distributions to determine the best solution combination for maximising impact. To proof this claim, we trained a generic decision tree in Python using the Scikit-Learn toolbox and feeding 40 of the project models as input data. It was possible for the tree to predict the remaining 6 projects and their methodologies. Of course, an extension to predicting new combinations was not possible. The results of this investigation will be part of a separate and more detailed publication on this topic. Let us emphasise this point once again: it is clear, that humans can easily construct such solution paths. But in order to establish an AI procedure that autonomously decides which solution or solution combination is best, it requires access to those pieces of knowledge, which in parallel would be used by the human decider.

REFERENCES

Bughin, J., Hazan, E., Lund, S., Dahlström, P., Wiesinger, A., and Subramaniam, A. (2018). Skill shift: Automation and the future of the workforce. *McKinsey Global Institute.*

Colla, V., Matino, I., Dettori, S., Petrucciani, A., Zaccara, A., Weber, V., Salame, S., Zapata, N., Bastida, S., Wolff, A., Speets, R., and Romaniello, L. (2018). Assessing the efficiency of the off-gas network management in integrated steelwork. *Materiaux and Techniques 106(1), DOI:10.1051/mattech/2018068.*

Dettori, S., Matino, I., Iannino, V., Colla, V., Hauser, A., and P. Wolf-Zöllner, S.H. (2022). Optimizing methane and methanol production from integrated steelworks process off-gases through economic hybrid model predictive control. *FACPapersOnLine, 55(2) pp. 66-71.*

Edlich, A., Phalin, G., Jogani, R., and Kaniyar, S. (2019). Driving impact at scale from automation and ai. *McKinsey Digital.*

Iannino, V., Colla, V., Maddaloni, A., Brandenburgerc, J., Rajabi, A., Wolff, A., Ordieres, J., Gutierrez, M.,

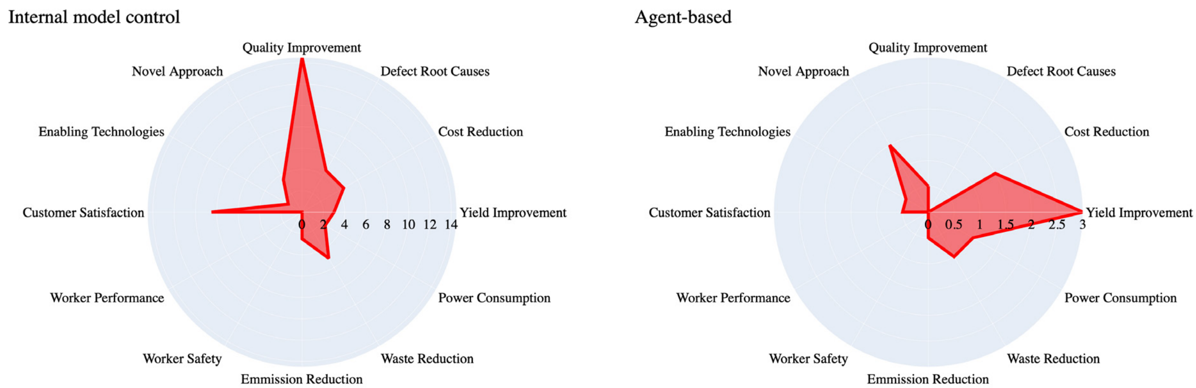

Fig. 8. Assorted elements of the solution taxonomy T_2 mapped via the deduction (4).

Sirovnik, E., Mueller, D., and Schirm, C. (2021). A hybrid approach for improving the flexibility of production scheduling in flat steel industry. *Integrated Computer-Aided Engineering*.

J. Ordieres-Mere, A. Arteaga, N.M.A.W.V.I. (2022). Process supervision of long products hot rolling mill. attack or failure identification. *IFACPapersOnLine, 55(2) pp. 72-77*.

Jelali, M. (2000). Explicit models of thickness profile and tension stress distribution applications. *Steel Research 6+7*. CEFLA.

Jelali, M. (2005). Regelkreis berwachung in der metallindustrie: Anforderungen, stand der technik und anwendungen. In *Proc. GMA-Kongress, Baden-Baden, Germany*. AutoCheck.

Jelali, M. (2007). Performance assessment of control systems in rolling mills - application to strip thickness and flatness control. *Journal of Process Control*. doi: 10.1016/ j.jprocont.2007.03.001. AutoCheck.

Jelali, M., Müller, M., Wolff, A., Ungerer, W., and Thiemann, G. (2002). Advanced measurement and flatness control for hot strip mills. *La Revue de Métallurgie-CIT*, 517–522. CEFLA.

Klimes, L., Mauder, T., and Stetina, J. (2011). Stochastic approach and optimal control of continuous steel casting process by using progressive hedging algorithm. *Proceedings of the 20th International Conference on Metallurgy and Materials (METAL) 2011*.

Leitão, P. and Restivo, F. (2006). Adacor: a holonic architecture for agile and adaptive manufacturing control. *Comput. Ind.*, 57(2), 121–130. doi: 10.1016/j.compind.2005.05.005.

Matino, I., Dettori, S., Colla, V., Weber, V., and Salame, S. (2019). Forecasting blast furnace gas production and demand through echo state neural network-based models: Pave the way to off-gas optimized management. *Applied Energy, 253, art. no. 113578*.

Neuer, M.J., Marchiori, F., Ebel, A., Matskanis, N., Piedimonti, L., Wolff, A., and Mathis, G. (2016). Dynamic reallocation and rescheduling of steel products using agents with strategical anticipation and virtual market-structures. *IFAC-PapersOnline*, 49, 232–237.

Neuer, M.J., Quick, A., George, T., and Link, N. (2019). Anomaly and causality analysis in process data streams using machine learning with specialized eigenspace topologies. In *Proceedings of the European Steel Application and Technology Days*.

Wolff, A., Mintus, F., Bialek, S., Dettori, S., and Colla, V. (2019). Economical mixed-integer model predictive controller for optimizing the sub-network of the bof gas. *Proceedings of the European Steel Application and Technology Days (ESTAD)*.

Zhu, W., Li, S., and Zhang, H. (2022). Evaluation of scientific research projects on the basis of evidential reasoning approach under the perspective of expert reliability. *Scientometrics 127, 275–298 (2022)*.