REVIEW

ABB

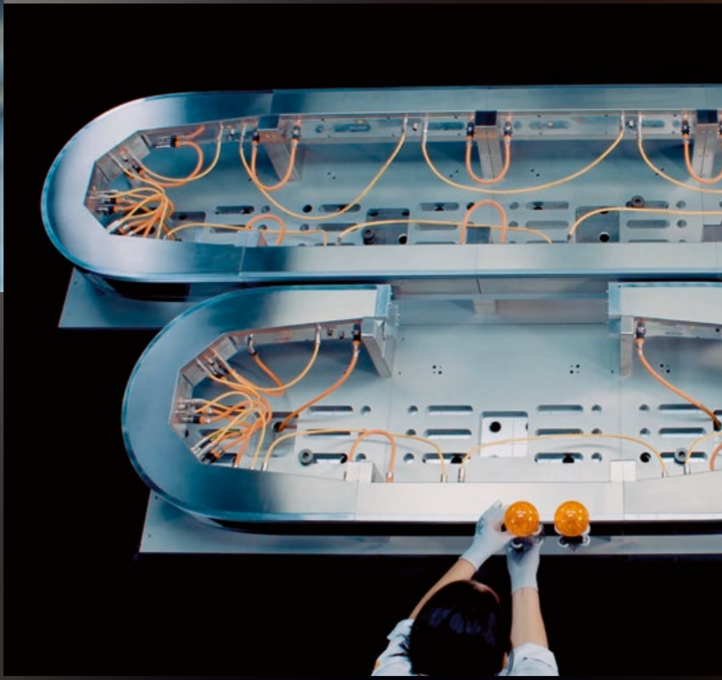
Women in technology special report

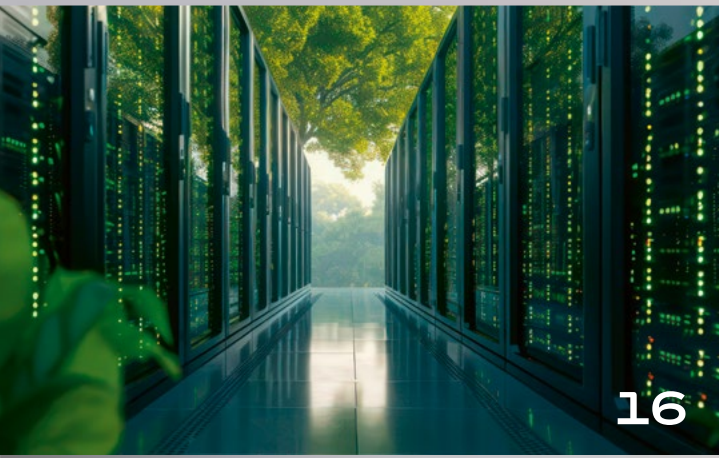
ONE OF THE WORLD'S LONGEST RUNNING TECHNICAL JOURNALS

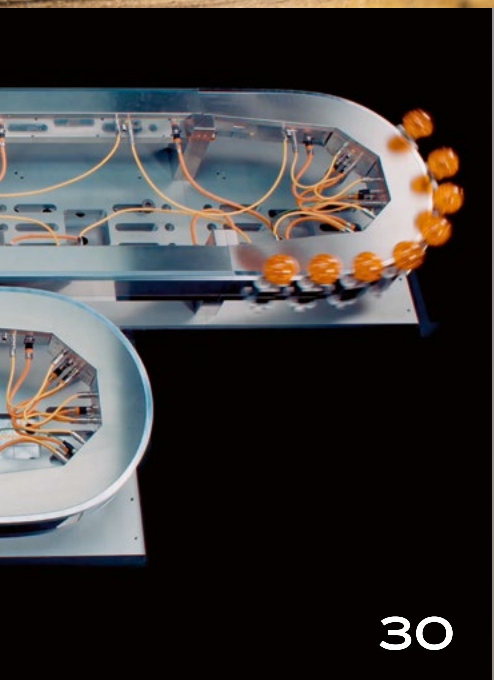
ENGINEERED TO OUTRUN


10


06







AI AND DIGITALIZATION

ENERGY AND SUSTAINABILITY

Share this issue of ABB Review

INTRODUCING THE AUTHORS

06
Stacey Jones
Global Asset Performance Management Portfolio Leader, Energy Industries, Digital
Bachelor of Science in Business

10
Ruomu Tan
Senior Scientist in Industrial AI, Tech Area Coordinator – AI
PhD in Chemical Engineering

16
Marie Platenius-Mohr
Research Team Manager, Automation & Software Architecture
PhD in Software Engieering

24
Rozenn Diehl
Project Manager, Process Automation Energy Industries
PhD in Physics, Master in Business Administration

30
Isabella Back
Global Product Marketing Manager, B&R Industrial Automation
MSc in Engineering – Information and System Management

38
Theresa Loss
Scientist in Optimization and Control
PhD in Electrical Engineering and Audio Engineering

38
Maryam Sharifi
Senior Scientist in Control and Optimization
PhD in Automatic Control

EDITORIAL

WOMEN IN TECHNOLOGY

Dear Reader,

Following the positive response to last year's Special Report, we're proud to bring you this new ABB Review Special Report: Women in Technology.

More women are entering science, technology, engineering and mathematics (STEM), taking on leadership roles, and driving innovation in industries around the world. Yet, there is still more to do – more barriers to break, more stories to tell, and more talent to elevate.

This edition features articles authored and co-authored by women across ABB. Their voices highlight that representation matters, that confidence grows through community, and that STEM is not only about solving problems – it's about imagining what's possible.

To anyone considering a career path in STEM: believe in your potential. The future needs your perspective, your curiosity, and your voice.

With pride in our progress and to keep reaching,

Amina Hamidi

Amina Hamidi
Managing Director, Instrumentation Business Line,
ABB Measurement & Analytics
Ph.D. in Electrical Engineering

Photo: ©Gorodenkoff, stock.adobe.com

ABB'S GENIX APM COPILOT TRANSFORMS ASSET PERFORMANCE WITH AI

GENERATIVE AI IN ASSET PERFORMANCE

ABB has expanded the generative AI capabilities of its Genix Asset Performance Management (APM) solution. The new component – Genix APM Copilot – offers intuitive functionality and streamlined contextualized data flow across processes and operations.

Stacey Jones
ABB Energy Industries, Digital
Lake Jackson, TX., United States

stacey.jones@us.abb.com

Key Facts

Effective management of assets and their performance is critical.

ABB uses LLMs and generative AI to make asset use more efficient.

Profitability up by 4–10%; maintenance costs down by 20–40%.

—
For further information, please visit https://new.abb.com/process-automation/genix/the-apm-you-built

As industrial operations become more extensive and more sophisticated, the effective management of assets and their performance becomes a critical task. APM is a data-driven approach that optimizes the reliability, efficiency and lifespan of physical assets such as turbines, compressors, electrical breakers and instrumentation. The methodology combines predictive analytics, condition monitoring and maintenance strategies to prevent failures, reduce downtime and improve operational performance. APM is crucial for industrial operations such as oil and gas, chemicals, power and water, where asset failures can lead to costly business disruptions and potential safety and environmental risks.

AI is transforming APM by enhancing predictive maintenance and optimizing asset utilization.

ABB Genix APM

ABB's APM offering is the Genix Asset Performance Management Suite. The comprehensive monitoring capability provided by this APM solution moves asset management beyond equipment condition monitoring and diagnostics to provide predictive and prescriptive analytics and optimization models, giving role-specific, actionable insights for quick strategic and tactical decision-making.

Boosting APM with AI

Currently, AI is transforming the world of APM by enhancing predictive maintenance, optimizing asset utilization, reducing downtime and improving overall operational efficiency, sustainability and profitability. By leveraging AI-driven analytics, machine learning models and real-time data processing, businesses can maximize asset performance while reducing costs and risks »01.

ABB Genix APM Copilot

To further leverage real-time data and AI-driven insights, ABB Genix APM now integrates large language models (LLMs) and generative AI capabilities in the cloud, using Microsoft's Azure Open AI Services. The new capability, ABB Ability Genix APM Copilot, enhances user experience by offering intuitive functionality and streamlining the flow of contextualized data across processes and operations. The companies are collaborating on this implementation of generative AI technology to help industrial customers unlock insights hidden in operational data and increase workflow efficiencies. Improved data insights and a more effective return of assets to optimal operating conditions will enable significant gains in efficiency, productivity, asset reliability and operational safety. Energy consumption and environmental impact can be reduced, too. The enhanced functionality

01 The power of AI can be harnessed to help tame the complexity of modern industrial plants and ensure the smooth running of operations.

Photo: ©zapp2photo, stock.adobe.com

Businesses running the Genix platform have already seen up to 40 percent cost savings in operations and maintenance.

and user experience delivered by the ongoing Microsoft/ABB industry and technology strategic partnership drives safer, more innovative and more sustainable operations across a wide range of industries.

ABB integrates generative AI by tuning LLMs, such as GPT-4, to the relevant platform and applications, enabling functionality such as code, image and text generation. The LLMs are trained on operational and maintenance data derived from industrial assets. In short, the potential of generative AI is used to improve data flow, provide real-time actionable insights and enhance decision-making processes for industry engineers, subject matter experts and executives [1]. Moving beyond the typical human-machine exchange of chat interfaces, Genix APM Copilot connects and takes actions within third-party solutions as part of integrated, agentic workflows.

The role of generative AI in industrial applications
Generative AI represents a significant leap forward in industrial digital solutions. By harnessing the power of LLMs to generate contextualized text, images and recommendations, ABB's Genix APM Copilot accelerates root-cause investigation and overall decision-making. This added capability ultimately lowers the cost for APM users of investigating an issue and enables the generation of an answer or course of action from within the same workflow. The Genix APM Copilot integration augments predictive maintenance, detailed asset analysis and the automation of complex tasks – all of which are pivotal for optimizing operations and reducing environmental impact.

Key features and benefits of Genix APM Copilot
Genix APM Copilot is designed to transform the way industries manage and maintain their assets by providing:

- Enhanced alert management by supplying intelligent support to help users prioritize and address critical alerts efficiently.
- Predictive and diagnostic insights that assist root-cause investigation based on embedded physics-based models and sensor rankings.
- An action support environment. For example, a work request in the maintenance management system can be created directly from APM workflows.
- Intuitive user interaction. Copilot offers a chat or voice experience that allows users to interact with the system through natural language queries. This feature simplifies data retrieval and analysis, making it accessible to users at all levels.

Transformation with Genix APM Copilot
The deployment of generative AI through ABB's Genix APM Copilot is poised to bring about transformative changes in industrial operations. As described above, improved data analysis and remediation will substantially improve efficiency, productivity, asset reliability and operational safety while reducing energy consumption. Insights into actions have never been faster or more intuitive. These advancements align with ABB's corporate commitment to enabling a more sustainable and resource-efficient future.

Businesses running the Genix platform have already seen up to 40 percent cost savings in operations and maintenance, up to 30 percent improvement in production efficiency and up to 25 percent improvements in energy and emission optimization [1]. The introduction of generative AI capabilities is expected to amplify these benefits further, ushering customers into a new era of AI-driven decision-making.

ABB's Genix APM Copilot, powered by generative AI, represents a significant milestone in the evolution of industrial digital solutions. By integrating cutting-edge AI technology with ABB's extensive domain expertise, this collaboration with Microsoft is set to revolutionize asset performance management, driving efficiency, sustainability and operational excellence across the industrial landscape. •

Reference

[1] ABB, "ABB and Microsoft collaborate to bring generative AI to industrial applications." Available: https://new.abb.com/news/detail/104829/abb-and-microsoft-collaborate-to-bring-generative-ai-to-industrial-applications. [Accessed February 13, 2025.]

—
Originally published in ABB Review 01/2025, pp. 30–33.

Read the original article here:
https://global.abb/group/en/innovation/news/generative-ai-in-apm

PLANTINSIGHT: HOW AI IS TRANSFORMING INDUSTRIAL PROCESSES

DATA-DRIVEN INSIGHTS

At the core of all industrial processes is the quest for ever more precise levels of control. Today, this quest is driven by increasing computer power and digitization in automation. However, although most of the underlying mathematics stems from the 1960s, only recently has it become feasible to apply some algorithms to real-time scenarios.

—
Heiko Petersen
Patrick Meade Vargas
ABB Process Automation
Mannheim, Germany

heiko.petersen@de.abb.com
patrick.meade@de.abb.com

Ruomu Tan
ABB Corporate Research
Process Automation
Ladenburg, Germany

ruomu.tan@de.abb.com

Key Facts

- PlantInsight is an ABB platform for machine-learning.
- Algorithms predict system responses several steps into the future.
- Control is more accurate.

01 The ability to manage continuous processes with precision is essential in the chemical industry.

ABB's Ability™ PlantInsight platform »01 is a case in point. The platform makes it possible to run a variety of machine-learning (ML) algorithms for detection, segmentation, and prediction of specific patterns in vast amounts of process data. This, in turn, enables the implementation of AI-based optimization solutions that help reduce pollutants, extend equipment lifespans, and lower production costs.

Nowadays, artificial intelligence (AI) seems omnipresent. It is a common topic of conversation, bookstores are flooded with literature about it, and few applications seem to do without it. But considering the sheer amount of hype, one may wonder why AI is still so seldom used in process industries. Or could it perhaps be that it is already used, but just not recognized as such?

Generally, a system is considered to be AI-driven when it performs tasks typically done by humans, such as visual perception, decision-making, speech recognition, and translation. As a matter of fact, AI-driven systems can outperform humans in a range of activities such as solving numerical problems, pattern recognition, and retrieving information from a massive number of sources.

Vision, hearing, speaking and motion are tasks very similar to those encountered by process control systems.

Nevertheless, such systems are still in their infancy when it comes to abstract reasoning or creatively turning information into eloquent texts, to say nothing of social interactions, consciousness, or self-awareness, all of which are routine for humans but out of reach for machines – at least so far.

In view of this, it is important to distinguish between different levels of AI. According to Kaplan and Haenlein [1], the evolution of AI can be divided into three stages:
1) artificial narrow intelligence – the application of AI to specified tasks.
2) artificial general intelligence – the application of AI to autonomously solve novel problems in multiple fields.
3) artificial super intelligence – the application of AI to any area that can benefit from scientific creativity, social skills, and general wisdom.

Most of today's AI solutions fall into the first category. In this sense, even James Watts' flyball governor, a speed regulator for his rotary steam engine of 1768, could be considered AI at stage one. However, it was never marketed as such – and the same can be said for the millions of control solutions operating in the power, refining, and chemical industries.

AI as a solution

Typically, AI systems not only consist of a brain, or in other words, a sophisticated algorithm; they also must be able to perceive and interact with the world. Vision, hearing, speaking and motion » complement the brain and allow AI-based systems to solve real-world problems – tasks very similar to those encountered by process control systems. While sensors measure process values (dependent variables), such as pressure, flow, temperature, etc., the controller takes these inputs and calculates the best way to adjust actuators such as valves, dampers, etc. (independent variables) to meet certain control objectives. In this scenario the controller's role is that of a brain, running algebraic calculations and making logical decisions.

Why now?

One of the most obvious reasons why AI is gaining ground now is the exponential increase of available computing power. Some of the constraints data scientists had in the past, such as a limited number of neurons in an artificial neural network (ANN), basically vanished, thus opening the door to leveraging the full potential of deep learning networks. Furthermore, anybody with a laptop and access to a cloud solution can run a training algorithm. This opens the market for new business models such as self-service model training and software as a service (SaaS). This not only democratizes AI but reduces engineering requirements on control solutions.

In the process industry, digitization began in the late '70s with the widespread introduction of programmable logic controllers (PLC) and distributed control systems (DCS), which replaced analog controllers. Adding new data points and control features became a programming, rather than a hardware installation and configuration task. This significantly increased the flexibility of the control process while reducing costs. However, adding more control features led to more complex control structures, which were often difficult to understand and maintain. They also required significant engineering effort and process know-how. The need for a leaner and more transparent control approach arose.

Advanced process steps

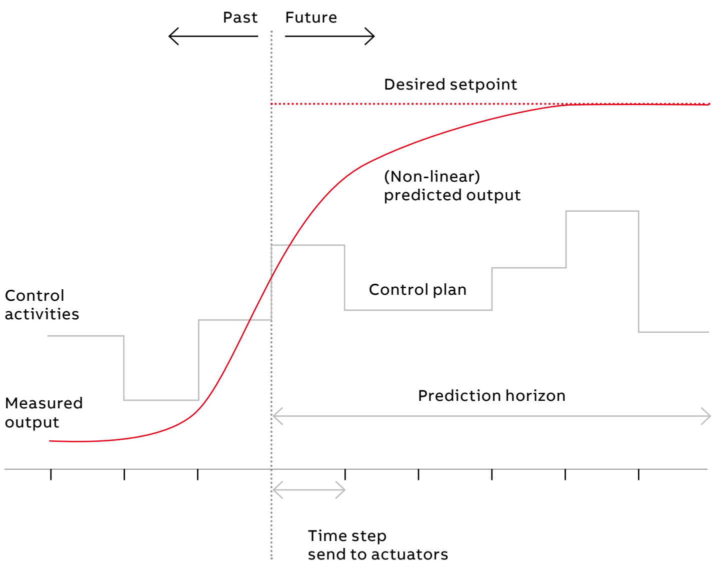
Advancements in mathematics and system theory, and the increasing availability of computer power, enabled the development of more advanced process controls. The mathematical fundaments behind this process can be traced back to the work of Rudolf Kalman et al. in the early 1960s [2]. While differential equations describe the dynamics of a physical system in a kind of 'cleanroom' scenario, Kalman added terms for state disturbance and measurement noise, something inevitable in any real-world application. Moreover, he directly formulated his equations using matrix representation, accounting for multiple differential equations with their respective inputs and outputs. This multi-inputs, multi-outputs (MIMO) approach made it possible to calculate an optimal control strategy not only for one actuator at a time, but for many simultaneously.

It turned out that Kalman's mathematical solution could be used to look into the future of a process.

Moreover, it turned out that Kalman's mathematical solution could also be used to look into the future of a process. In contrast to a simple controller, which only calculates the next optimal step for one variable, it was now possible to look multiple steps ahead into the future for multiple variables. The goal remained the same: to minimize the control error, which is the difference between desired and actual process values. But whereas a simple control is 'driving by sight', a forward-looking regulator creates a longer-term plan to act upon.

However, as things often do not go according to plan, it became evident that controllers must be able to adjust to changing situations based on feedback from a process. This led to the development of Model Predictive Control (MPC), which generates an optimal control path but triggers only the first step in each iteration. A moment later, once feedback is received, it repeats the process of calculating the optimal path until the desired operating point is reached »02.

Although these steps have significantly improved many processes, there are multiple areas where

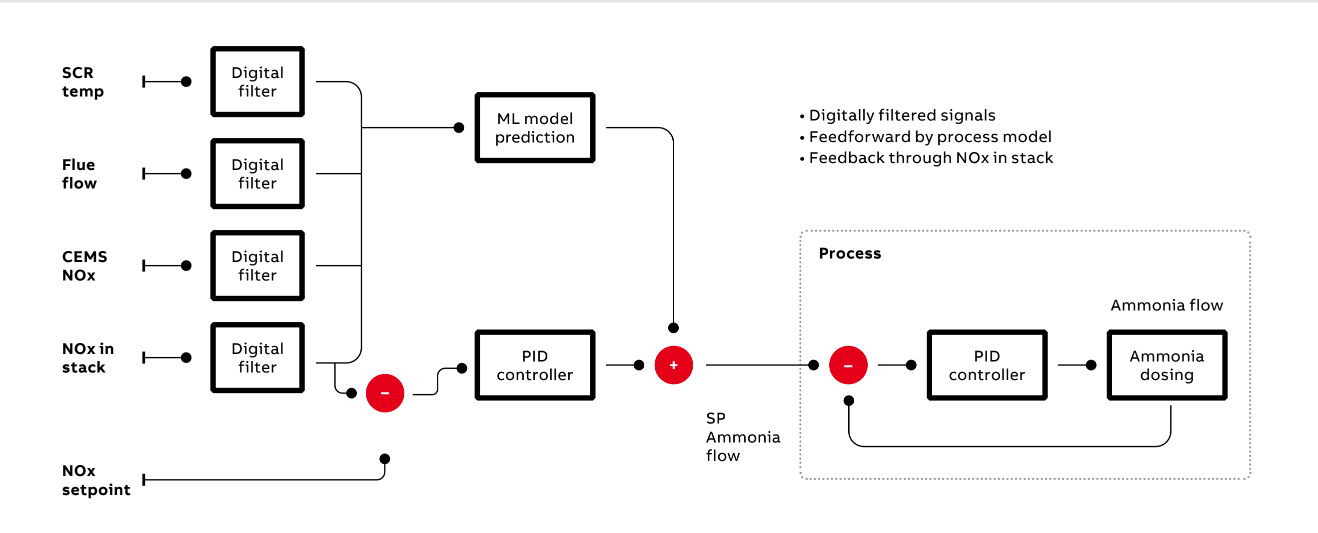


02 Principle of non-linear model predictive control (NMPC).

PLANTINSIGHT'S FIRST APPLICATION

Control solutions based on artificial intelligence not only keep a plant running at its economic optimum but can also help to better control and thus reduce its emissions. Indeed, as a first application, PlantInsight's AI capabilities were used in combination with conventional PID controls to reduce the emissions of a chemical residuals incineration facility. Compared to minimizing the emissions of a gas or coal-fired incineration plant, which is mainly a function of managing temperature, air and humidity, the emissions that chemical residuals emit during incineration are much harder to predict.

Based on one year of process data, PlantInsight analyzed the interdependency of multiple process variables that might contribute to emission levels. Deadtimes were considered by automatically shifting data and adding lagged process variables to the dataset. After feature selection of the most promising variable candidates, an artificial neural network was trained. The network was able to predict emission levels several minutes into the future with a high level of accuracy. This soft-sensor prediction was then used as feed-forward information for a conventional PID control system. As a result, cost savings due to reduced ammonia use were achieved. This in turn led to a reduced CO_2 footprint due to reduced ammonia consumption, while improving reliability and equipment lifetime due to reduced corrosive ammonia slip »03a.



03a Simplified representation of AI-supported control of an emission reduction system. PID: proportional integral derivative. CEMS: Continuous Emission Monitoring System

03 PlantInsight's first application.

process control is still limited. The following section describes some of these areas and how AI can contribute to overcoming the remaining limitations.

Real-time feedback
As described above, controllers require feedback from the process they are controlling, otherwise their performance may suffer. This problem intensifies the longer the delay between action and feedback. Specifically, data with large time gaps compared to the actual process might pose issues. This is typically the case for laboratory data covering product properties that cannot be measured continuously or in real-time, such as viscosity or flashpoint. Adjustments to a process can be performed only after receiving results from the lab, which, because of the inherent delay, might compromise product quality.

The accuracy of ML models, such as artificial neural networks, can be continuously improved with each new lab measurement.

One way to overcome this is to estimate the values of a product's qualities in real time using ML models, such as artificial neural networks (ANN). Here, the accuracy of the models can be continuously improved with each new lab measurement. Predicted qualities can then be used without delay by the control algorithm to adjust the process. In this configuration, conventional and AI-based control algorithms work hand in hand to achieve and » maintain desired production goals. This concept can also be applied to processes with long dead times or processes that use sensors that need to recalibrate regularly and are thus not continuously available. A practical use case pertaining to emission reductions illustrates the above concepts »03.

Adapting to non-linearities

Like most systems in the real world, industrial processes are often non-linear. This results in a systemic discrepancy between the real process and its linear process model. In the context of short time horizons and minor process alterations the resulting error may be neglectable. However, on a larger scale it may affect control performance. Although some non-linearities can be offset through transformation of their associated process data – for instance linearization of a control valve's characteristic curve – linearization is not always perfect and can be costly when dealing with many process variables.

AI techniques, on the other hand, can deal very well with non-linearities. ML models can basically adapt to any non-linear behavior. While most MPC implementations use a linear approach for modeling, the framework itself makes no assumption about the type of process model used or its linearity. Therefore, non-linear models trained with ML algorithms can also be used to reduce modeling errors. This leads to more accurate control and prevents the controller from getting trapped in minor optimizations.

Identifying the right process behavior

At the heart of any advanced process control system is a process model. However, the process of identifying the dynamics of a physical system is costly and requires domain know-how and experience.

Traditionally, there are two approaches to model design: a so-called first principles model, which is based on the design, mechanics, and fundamental physics of a system, and a so-called empirical model, which is based on observations of how a system reacts to stimuli, for example, by means of step-response experiments.

Both approaches can be highly complex, costly, and in some cases, due to the nature of the process, impossible to implement. However, in many cases, this burden can be avoided if adequate historical process data is available. During normal plant operation, setpoints are regularly changed and disturbances are continuously happening, both triggering reactions in the process, and thus revealing its dynamic behavior. These footprints can be used by ML algorithms to easily create accurate models »04. To accomplish this, the data must be representative and thus cannot be randomly picked. For instance, abnormal process behavior, or periods with missing data must be removed. Doing this manually would be costly, but for an algorithm this is the perfect task. Selecting, segmenting and clustering vast amounts of data is a home run for machine learning »05.

Non-linear models trained with ML algorithms can be used to reduce modeling errors, thus leading to more accurate control.

04 Machine learning models can help to overcome gaps in laboratory data.

05 Example of automatically identified operating states that could be used for process identification. (gray = not-applicable, pink = applicable).

Platform solution

Over the years, ABB has developed a suite of control and optimization solutions that have followed and often led technical developments in this area. From first principle modeling to proportional integral derivative (PID) loop monitoring, model predictive control, and dynamic optimization, ABB provides a wide range of solutions. Today, with the assistance of hardware and machine learning algorithms, it is now possible to complement this offering with the benefits and opportunities of artificial intelligence. With this in mind, ABB has developed ABB Ability™ PlantInsight, a platform that leverages the full potential of ML algorithms. This web-based application makes it possible to run a multitude of machine learning algorithms for prediction, segmentation, and detection of specific patterns in huge amounts of process data, while its modular concept makes it easy to embed proprietary python scripts and complement them with existing ones.

PlantInsight makes it possible to run a multitude of ML algorithms for prediction, segmentation and detection of patterns.

All in all, it can be said that joining the world of control to that of artificial intelligence can yield significantly improved results in terms of controlling industrial processes. Indeed, the more such hybrid control solutions spread, the more both worlds will converge – an apparently natural process since both share the same theoretical foundations. As this process evolves, continued progress is set to pave the way to the introduction of tomorrow's fully autonomous production facilities. •

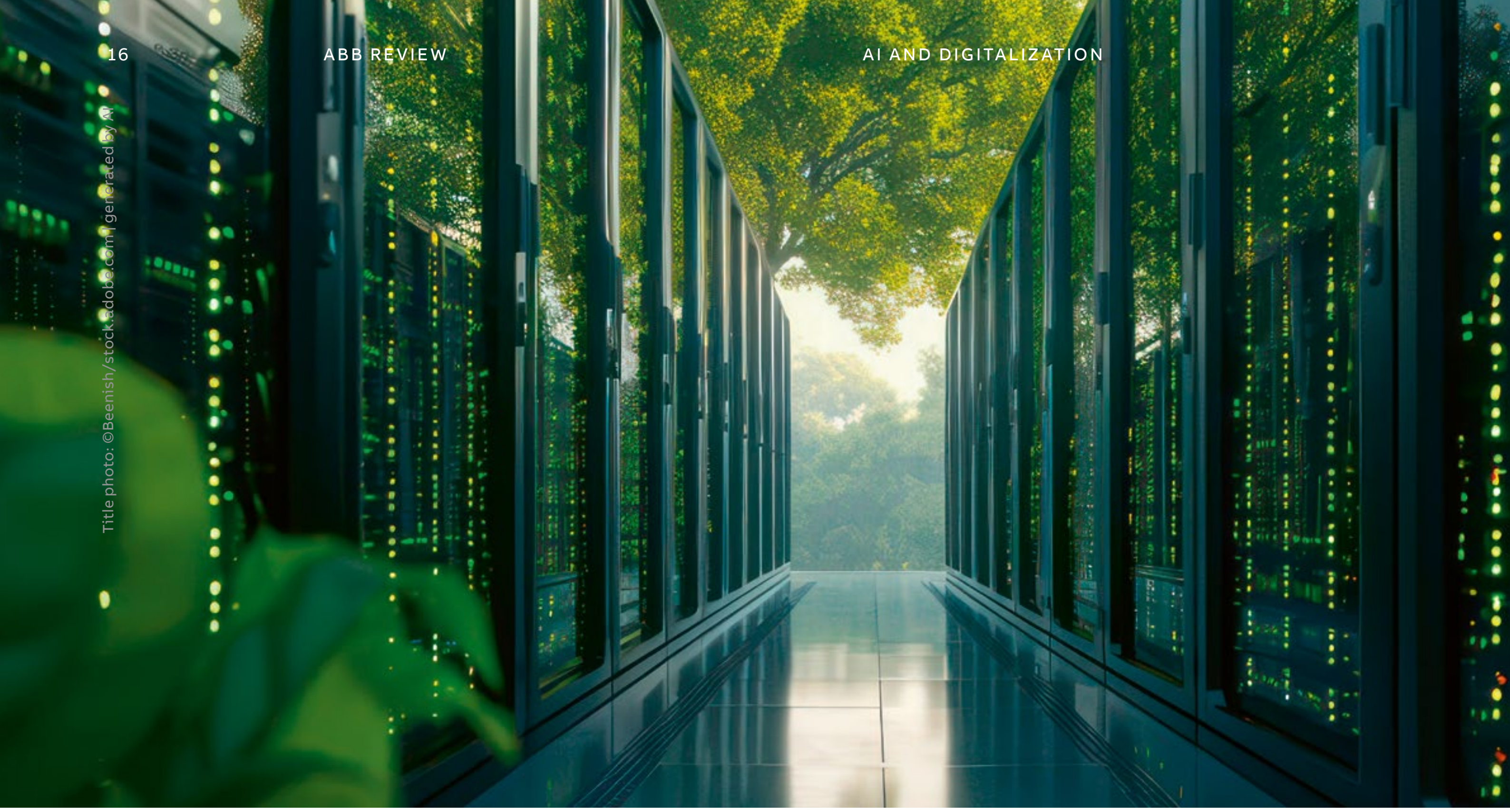
References

[1] Andreas Kaplan, Michael Haenlein, "Siri-Siri in my hand, who is the fairest in the Land?" in *Business Horizons*, vol. 62, no.1, 2019, pp. 15–25.

[2] Kalman, R. E. (1960). "A New Approach to Linear Filtering and Prediction Problems." *Journal of Basic Engineering.* 82, pp. 35–45.

—
Originally published in ABB Review 02/2024, pp. 94–99.

Read the original article here:
https://new.abb.com/news/detail/112771/data-driven-insights



SUSTAINABILITY OF ARTIFICIAL INTELLIGENCE MODELS

HOW GREEN IS THE MACHINE?

Understanding the factors that affect the carbon footprint of computation could help decision makers in the process industries to reduce their CO_2 emissions. Taking a theoretical and experimental approach, ABB explores this topic and provides advice on how to make AI models greener.

Key Facts

- What is the environmental impact of the computational needs of AI?
- ABB's framework explains the carbon drivers of models.
- The carbon footprint of process automation models was found to be negligible.

—
Ralf Gitzel
Marie Platenius-Mohr
ABB Corporate Research
Mannheim, Germany

ralf.gitzel@de.abb.com
marie.platenius-mohr@de.abb.com

Andreas Burger
Former ABB Employee

Artificial intelligence (AI), specifically machine learning (ML), has infiltrated everyday life. Neural networks (NN), which are multi-layered deep learning models, provide facial recognition capabilities for added mobile phone security or convert human speech into commands for smart home applications. This rapid progress is becoming increasingly relevant for process industries with applications ranging from the interpretation of infrared images of machinery [1] to the analysis of production-related data and more [2]. Clearly, the promise of this potential ignites competition to improve performance leading to ever larger AI models that are trained for longer, thereby generating worrisome secondary effects: more energy is consumed and more CO_2 is emitted [3,4,5] – less than laudable ramifications considering the current climate crisis.

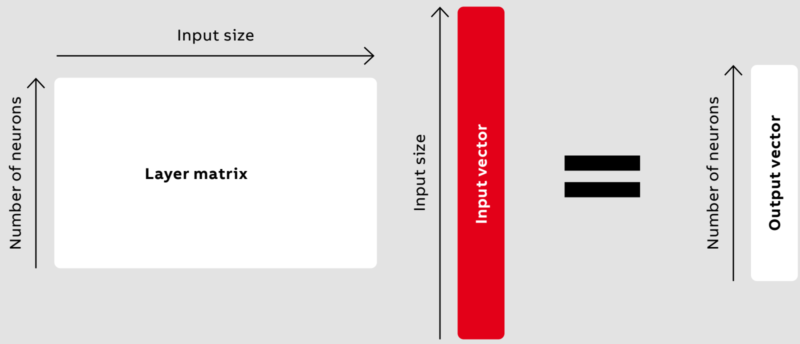
Dilemmas and goals

It would appear that carbon footprint goals must be sacrificed for the performance enhancement generated by AI models. But, is this necessarily so? At first glance, studies that have evaluated this tenet have focused on high-performance language- or image processing models such as GPT-3, a deep, a neural network model with 175B parameters that provides human-like texts. This large language model (LLM) required 1’287 MWh for training, which corresponds to 552 t of CO_2 – the annual emission of 276 medium-sized cars [4]. Though not yet disclosed, the footprint for GPT-4 will probably be much larger. Still, other high-performance AI models are associated with a smaller carbon footprint; some investigators have even suggested that the scope of the problem has been exaggerated [6]. Such discrepancies make AI providers unsure about the carbon footprint of their specific model or how to reduce it. Decision makers in the process industries face an additional challenge: models are typically much smaller in scale than the high-performance models discussed in the literature. Are emissions in these cases even relevant? In this paper, ABB aims to give engineers, managers, and others guidance so they understand the impact that individual AI models have on the environment. Specifically, ABB examines the literature to create a comprehensive framework to explain the various carbon drivers over the entire AI model lifecycle and offers advice about reducing those drivers [7]. Based on experimental data, ABB also tests the validity of literature recommendations to provide guidance for reducing the carbon footprint and energy consumption of AI models, eg, the use of transfer learning models. Moreover, the carbon footprint of process industry-relevant AI models is computed and discussed.

Ever larger AI models, trained for longer, generate worrisome secondary effects such as higher energy consumption.

State-of-the-art AI carbon footprint

While theoretical models from the literature estimate the carbon impact of a new AI model based on architecture (layer types and size), training and usage [4,5,8,9], it is difficult to determine which metric drives the footprint [5,10]. Contrastingly, the carbon footprint of ML models can easily be measured using software tools to record the impact of development or use via carbon accounting. Some tools use metrics, eg, training time, energy mix, and hardware information [8,11], while others, eg, energy usage or CodeCarbon, integrate directly with the ML code [6,12,13,14]. Other tools compute »



01 Inference of a single layer expressed as a matrix multiplication – a simple case. For a layer, a matrix multiplication and a simple activation function eg, (ReLU), are applied. Activation function omitted for clarity.

central processing unit (CPU) power usage, estimate graphics processing unit (GPU) runs, compare hardware type [12,15,16] and determine the carbon impact of image recognition models [17].

Despite the significance of this research, two study gaps loom: First, studies are either too generic or focus on specific unrelated domains, eg, images [17,16] or LLMs [3,4,5], with unknown relevance to process industry data since industry models are specific and use small data sets. Second, carbon calculation models are not standardized and specific lifecycle steps are often omitted [18]. To close these gaps ABB empirically evaluated the carbon footprint and created a model for all AI life cycle phases.

Deep model's carbon footprint

The carbon footprint (CO_2eq) of NNs depends on how much energy is used (in kWh) and the carbon intensity (in lbs/kWh) of the energy source. The carbon intensity of Chat GTP-3 (1'214'400 lbs CO_2eq) [4], Gopher (851'200 lbs CO_2eq) [18], and NAS (626'155 lbs CO_2eq) [9] are unsurprisingly high. Contrastingly, the carbon footprint of other high-performing models is much lower, such as BERTbase (1'438 lbs. CO_2eq) [9]. Such disparities suggest that the factors that impact a high-performance model's carbon footprint require more scrutiny.

Experiment	Energy use
MNIST-exception transfer learning	0.451 kWh
Dedicated MNIST model	0.005 kWh

02 The experimental comparison of transfer learning MNIST model results to the dedicated MNIST model results in relation to energy use.

Examining AI lifecycle phases – inference

To explain the impact of AI models on carbon footprint, ABB holistically modeled all life cycle phases [4,5]:

- model architecture search (MAS) – the design stage
- model training – for training the model with data
- inference – the usage stage

Because inference operations are executed during all phases, they are described first. Essentially, inference (which is estimated to cause between 80 and 90 percent of a model's total energy use [4]) can be defined as the computation of a mathematical formula expressed through a series of learned parameters that transforms an input vector into the correct output vector, eg, an image, time series, predicted value, etc. The mathematical operations for a standard NN dense layer consist of a matrix multiplication and application of a simple activation function to the result »01. Layer output acts as input for the next layer leading to a series of matrix multiplications, thereby consuming energy. The amount of inference energy used depends on: model architecture (M), ie, layer types, order, and

The impact of AI on carbon footprint is modeled for each life cycle phase: model architecture search, training and usage.

ABB's results confirm that smaller models use less energy than larger ones, especially if the model properties are similar.

size; and type and quantity of processing units (PT) eg, CPUs, GPUs, and tensor processing units (TPUs). The overhead imposed by the power usage effectiveness (PUE) of the datacenter also has an impact [19]. Thus, the energy cost I of an inference can be described as:

$$I = F(M, PT) \cdot PUE \quad (1)$$

Approximating f is challenging mainly due to different hardware implementations, memory access [13,9,21] and the use of specialized layers. Thus, simple substitutes for M such as the number of trainable parameters [20] are problematic [10]. Nonetheless, measurement-based estimates of I can be used to calculate the total life cycle carbon footprint of a model. Both PT and PUE can be optimized by choosing efficient data centers and/or hardware. For example, a GPU is 10 times more efficient than a CPU; a TPU is 4 to 8 times more efficient than a GPU [8]. Although the PUE of a datacenter might be unavailable, those centers located in colder regions generally consume less energy than those situated in warmer regions [22]. Selecting a low-carbon M can also reduce energy use without sacrificing performance [20,12,14]. To reduce model size suggested techniques are pruning, adding sparsity, quantization, or knowledge distillation [4, 23,24]. For DNNs [6], computation effort can be reduced by factors 5–10 [4]; for convolutional neural networks (CNNs) – a feature engineering NN – by a factor of 40 [20].

The training phase

Energy use during a model's training phase depends on training duration and number of processors used [4]. Three factors act as drivers: The energy cost of a single inference (I), the size of the training data set (D) and the number of epochs (E) used to optimize the model weights. Overhead[1] is expressed as a constant θ.

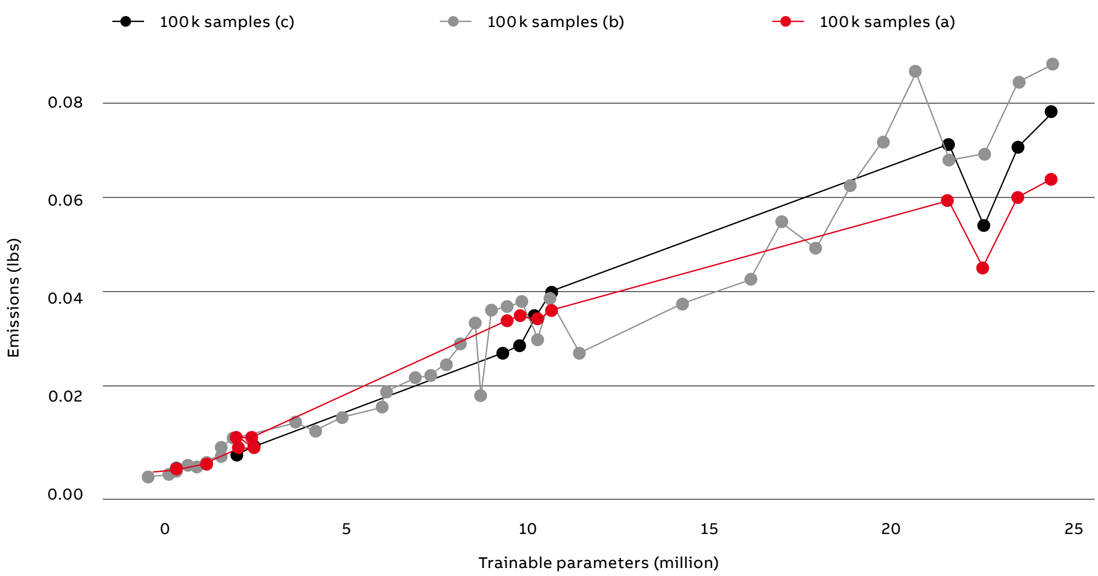
$$T \propto E \cdot D \cdot I \cdot \theta \quad (2)$$

Where, PUE, and number and type of processors are considered within the value for I. It follows that training energy can be reduced, theoretically, through transfer learning – the reuse of a pre-trained model on a new problem – as it reduces E and D [4,5,8].

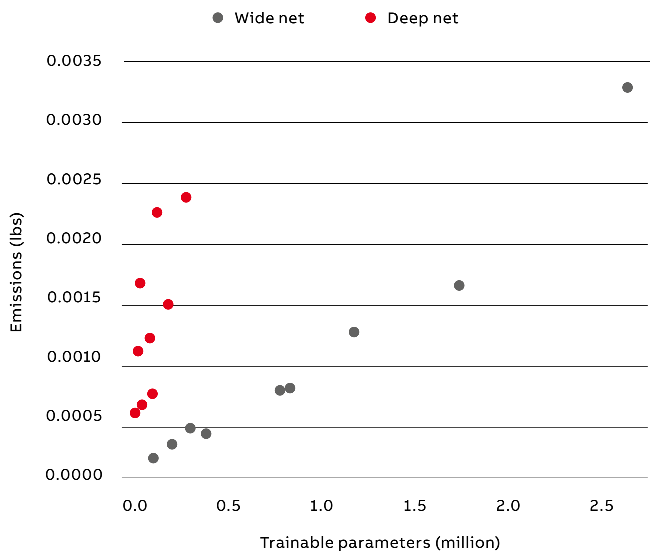
The MAS phase

Notably, different model architectures used for the same task can vary in accuracy. For this reason, many architectures are trained during MAS and »

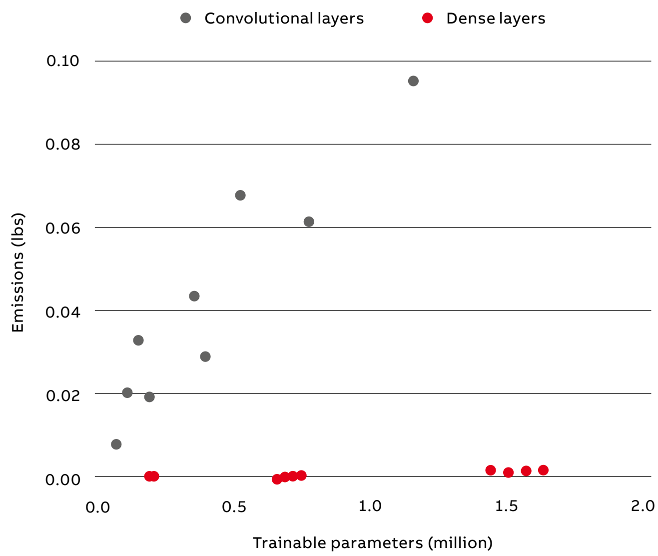
[1] There is significant overhead for the loss function and back propagation step, which is included in the calculation as its use has been validated in experiments.



03 Here, layer size was varied between 25 and 100 nodes and the number of layers ranged from 10 to 170. The models were trained for 10 epochs with 100'000 samples. Except for an odd dent at the end, the growth seems almost linear.



04 The energy footprint of varying shapes – 10 epochs with 100k samples with 100 values is displayed.



05 The energy consumption related to the trainable parameters of different layer types, ie, convolutional layers and dense layers (the simplest type) are displayed.

the best one is chosen for the final training phase. While performance is the optimization criterion of choice, energy consumption could be used as an additional criterion.

The cost at this stage (CT) is proportional to two factors [5]: The cost of training, T, and the number of times hyperparameters are tuned (H). Some of T's components, ie, I, E, and D, might vary for each tuning step resulting in different values of T for each tuning step.

$$CT \propto \sum_{n=1}^{H} T_h \tag{3}$$

Choice of MAS is critical because the more often Hs are tuned, the more energy is used. Interestingly, in terms of energy used, a random search is better than a systematic grid search, which compares many similar architectures [8]. It follows from equation iii that transfer learning could reduce MAS or even eliminate it [4].

Life cycle energy use and carbon footprint
The total life cycle energy use depends on the energy cost of all life cycle phases: CT, T, I; and the expected number of inference calls (e):

$$E_{life} = CT + T + I \cdot e \tag{4}$$

The CO_2eq is determined by multiplying (4) with the carbon emission factor (EF):

$$CO_2eq = E_{life} \cdot EF \tag{5}$$

EF varies greatly depending on the energy source used. For example, EF ranged from 20 g CO_2eq/kWh in Quebec, to 736.6 g CO_2eq/kWh in Iowa in 2019 [8]. Evidently, the easiest way to reduce CO_2eq is to choose the right location [4].

By taking into account the carbon footprint at each ML life cycle stage as determined in the previous sections, the resulting consolidated framework provides a reasonable estimate of a DNN model's carbon footprint[2].

Experimental conditions
To empirically test framework assumptions about the carbon footprint of models with different properties, ABB conducted a series of experiments[3].

ABB's tested process automation-relevant models have a negligible carbon footprint, yet perform well.

The code (Keras/ Python) was tested on a PC with a GeForce RTX 2080 Ti GPU and 32 GB RAM. ABB assumed the energy mix of Germany and used the CodeCarbon tool, which uses a carbon intensity of 365.5 g/kWh for its calculations.

[2] The framework ignores static energy consumption and original hardware production [13] and in contrast to some studies sacrifices accuracy to focus on ease-of-use.

[3] Model performance was not considered in the experiments. Carbon optimization and performance optimization interfere with each other but are not a direct trade-off.

As large deep models enter the industrial domain, ABB's results will help engineers and managers to make better decisions.

Testing set size, epochs, and pretrained model use
To test if set size and epoch number increase energy use, ABB conducted two experiments:
1. H was fixed for a model and the number of training samples was increased.
2. The number of epochs was increased for two models of different sizes: 100 and 50 layers of size 25.

The results confirm that energy use linearly increases with increasing number of training samples. Similarly, increasing epoch size results in a linear growth of emissions, thereby demonstrating the vital importance of both factors. While these results seem to confirm the value of pretrained models to reduce energy consumption [4,5,8], ABB's experiments indicate that pretrained models bear the risk of using oversized and therefore inefficient models. A pretrained model can be fine-tuned for a fraction of the cost that training the same architecture would require when trained from scratch [26]. However, using a dedicated (smaller) architecture for a problem can be even more energy-efficient. For example, in an experiment with MNIST classification, a dedicated model needed only a tiny fraction of the energy used for a fine-tuned Xception model of comparable performance »02.

Testing the impact of model size
Generally, larger models require more energy than smaller ones, especially if the model properties are similar. ABB's experimental results confirm this »

	ECOD [28]	DeepSVDD [29]	BLOOM (Benchmark) [18]
Trainable parameters	N/A	3,202,048	1.79 x 10^9
Training data	60 MB	60 MB	1.6 TB
Training time	<1 min	11 min	1 x 10^6 h
Training carbon footprint	0.000027 lbs	0,00012 lbs	661'387 lbs

06 An overview of typical industrial models. Data about a less intensive LLM (BLOOM) is added for comparison. Good performance levels are indicated. The industrial models seen in the literature are on a similar scale.

statement »03. They also support the literature in rejecting the number of trainable parameters as a carbon driver.

In one test, ABB compared wide and narrow models with the same number of trainable parameters and observed great divergence in energy used »04. In deep learning, it is notable that the energy consumption of nets with many small layers, "deep nets", is much higher than for "wide" nets of the same size (fewer layers yet more nodes per layer). Nevertheless, depth is far better than width at increasing expressive power of an NN [27].

To assess the impact of layer type [20], ABB compared two groups of models: a series of wide models with dense layers, and a series of similarly shaped convolutional layers (where the "width" is represented by the number of filters). The results show that purely convolutional models consume significantly more energy than do dense models »05. Thus, trainable parameters are a basic indicator of energy use, but only if the models compared share many properties ie, shape and type of layers.

Implications for AI models in process industry
While the consolidated framework presented and the experimental results can help users and decision-makers to reduce model carbon footprints, the question arises: Are these findings even applicable to process industry-relevant models? Certainly, vast

The explosive growth of LLMs, eg, GPT 4.0, strongly indicates that large deep models will enter the industrial domain soon.

quantities of industrial data are produced, thanks to distributed control systems (DCS) indicating that NN models could be useful. Unfortunately, significantly less data is available for training because most of this data is unlabeled. Less available data means less training time and lower energy costs. But, is this positive, or not? Crucially, such a scenario implies lower performance and, yet smaller AI models with a specific use case and good feature engineering do perform well indicating that performance might not need to be sacrificed.

To evaluate the carbon footprint of small models, ABB chose to evaluate two literature examples »06: a non-deep anomaly detection algorithm (ECOD) [28] and a deep anomaly detection model, Deep Support Vector Data Description (DeepSVDD) [29]. Both models were trained on data from an angular sensor used for condition monitoring. Not only do both models perform well in the test, their carbon footprints are negligible even when compared to an efficient LLM such as BLOOM »06.

These results suggest that further actions to reduce the carbon footprint of such process automation-relevant models is currently unnecessary. Nonetheless, the explosive growth of LLMs, eg, GPT 4.0, strongly indicates that large deep models will enter the industrial domain soon. When this happens, the consolidated framework and experimental findings discussed in this paper will help engineers and managers make better decisions about the design, deployment, and use of their models in terms of carbon footprint. •

—
Originally published in ABB Review 04/2024, pp. 252–258.

Read the original article here:
https://global.abb/group/en/innovation/news/how-green-is-the-machine

References

[1] R. Gitzel, et., "Maps of Infrared Images to Detect Equipment Faults", *IEEE Eighth International Conference on Big Data Computing Service and Applications (BigDataService)*, 2022, pp. 167–172.

[2] M. Gaertler, et al., 2021, "The machine learning life cycle in chemical operations–status and open challenges", *Chemie Ingenieur Technik*, vol.93, no.12, 2021, pp. 2,063–2,080.

[3] O.Y. Al-Jarrah, et al., "Efficient machine learning for big data: A review", *Big Data Research*, vol. no.3, 2015, pp. 87–93.

[4] D. Patterson, et al., "Carbon emissions and large neural network training" *arXiv preprint arXiv:2104.10,350, 2021*.

[5] R. Schwartz, et al., "Green ai" *Communications of the ACM*, vol.63, no. 12, 2020, pp. 54–63.

[6] D. Patterson, et al., "The carbon footprint of machine learning training will plateau, then shrink". *Computer*, vol.55 no.7, 2022, pp. 18–28.

[7] D.H. Fisher, "Recent advances in AI for computational sustainability" *IEEE Intelligent Systems*, 31(04), 2016, pp.56–61.

[8] A. Lacoste, et al., "Quantifying the carbon emissions of machine learning", *arXiv preprint arXiv:1910.09700*, Available online at doi:10.48550/ARXIV.1910.09700, 2019.

[9] E. Strubell, et al., "Energy and policy considerations for deep learning in NLP" *arXiv preprint arXiv:1906.02243*. Available online at doi:10.48550/ARXIV.1906.02243, 2019.

[10] L. Lai, et al., "Not all ops are created equal!", *arXiv preprint arXiv:1801.04326*, 2018.

[11] L. Lannelongue, et al., "Green Algorithms: Quantifying the Carbon Footprint of Computation". *Advanced science*, vol. 12 no.8, 2021, pp. 1–10.

[12] K. Lottick, K., et al., "Energy Usage Reports: Environmental awareness as part of algorithmic accountability", *arXiv preprint arXiv:1911.08354*, 2019.

[13] P. Henderson, et al., "Towards the systematic reporting of the energy and carbon footprints of machine learning", *The Journal of Machine Learning Research*, vol.21 no.1, 2020, pp. 10,039–10,081.

[14] M. Kumar, et al., "Energy-efficient machine learning on the edges", in *IEEE international parallel and distributed processing symposium Workshops*, 2020, pp. 912–921.

[15] Y. Wang, et al., "Benchmarking the performance and energy efficiency of AI accelerators for AI training", in *20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing*, 2020, pp. 744–751.

[16] R. Selvan, et al., "Carbon footprint of selecting and training deep learning models for medical image analysis" in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Cham, CH, Springer Nature 2022, pp. 506–516.

[17] L. Heguerte, et al., "How to estimate carbon footprint when training deep learning models? A guide and review". *arXiv preprint arXiv:2306.08323*, 2023.

[18] A. S. Luccioni, et al., "Estimating the carbon footprint of bloom, a 176b parameter language model" *Journal of Machine Learning Research*, vol.24 no.253, 2023, pp. 1–15.

[19] E. Jaureguialzo, "PUE: The Green Grid metric for evaluating the energy efficiency in DC (Data Center). Measurement method using the power demand", *IEEE 33rd International Telecommunications Energy Conference 2011*, pp. 1–8.

[20] E. Cai, et al., "Neuralpower: Predict and deploy energy-efficient convolutional neural networks", in *Asian Conference on Machine Learning*, 2017, pp. 622–637.

[21] M. Horowitz, "1.1 computing's energy problem (and what we can do about it", in *IEEE international solid-state circuits conference digest of technical papers*, 2014, pp. 10–14.

[22] M. Sharma, et al., "Analyzing the data center efficiency by using PUE to make data centers more energy efficient by reducing the electrical consumption and exploring new strategies" *Procedia Computer Science*, 48, 2015, pp. 142–148.

[23] D. Blalock, et al., "What is the state of neural network pruning?", *Proceedings of machine learning and systems*, 2, 2020 , pp. 129–146.

[24] G. Hinton, et al., "Distilling the knowledge in a neural network", *arXiv preprint arXiv:1503.02531*, 2015.

[25] L. Heim, et al., "Measuring what really matters: Optimizing neural networks for tiny ml", *arXiv preprint arXiv:2104.10645*, 2021.

[26] P. Walsh, et al., "Sustainable AI in the Cloud: Exploring machine learning energy use in the cloud", in *36th IEEE/ACM International Conference on Automated Software Engineering Workshops*, 2021 pp. 265–266 doi:10.1109/ASEW52652.2021.00058

[27] Z. Lu, et al., "The expressive power of neural networks: A view from the width", *Advances in neural information processing systems*, 30, 2017.

[28] Z. Li, et al., "Ecod: Unsupervised outlier detection using empirical cumulative distribution functions", *IEEE Transactions on Knowledge and Data Engineering*, 2022.

[29] L. Ruff, et al., "Deep One-Class Classification" *Proceedings of the 35th International Conference on Machine Learning*, 80, 2018, pp. 4,393–4,402.

—
This article is a condensed and edited version of: "Estimating the sustainability of AI Models: Based on theoretical models and experimental data" in *ATP magazine*, vol. 66 no. 3 March 2024.

SENSITIVE DETECTION BY REINJECTION

HIGH-PRECISION GAS MEASUREMENTS

An ABB research project has successfully employed an optical reinjection technique to construct a high-precision analyzer for the simultaneous measurement of methane and ethane. The method also significantly reduces false positive errors and improves the ethane-peak detection rate.

Hsiang-Yu Lo
ABB Research Center
Baden-Dättwil, Switzerland

hsiang-yu.lo@ch.abb.com

Rozenn Diehl
has changed her role in ABB, please direct enquiries to the other authors.

Douglas S. Baer
Kyle Owen
Tharindu Jayasinghe
Julio D. Lobo Neto
ABB Measurement & Analytics
San Jose, United States

doug.s.baer@us.abb.com
kyle.owen@us.abb.com
tharindu.jayasinghe@us.abb.com
julio.d.lobo-neto@us.abb.com

Oil and gas infrastructure is a significant source of emissions of methane and ethane, both of which are greenhouse gases. These emissions occur throughout oil and gas production processes, from extraction to transportation, due to unintentional leaks and intentional venting. Such releases not only intensify climate change – methane has a global warming potential more than 80 times that of carbon dioxide (CO_2) over 20 years (GWP-20) – but also represent a significant loss of valuable resources. These considerations make efficient

Methane's GWP-20 is over 80 times that of CO_2, making efficient methane and ethane leak detection indispensable.

Key Facts

- Interaction between laser light and gas sample in analyzers can be low.
- Optical reinjection improves analyzer performance.
- Better precision, better thermogenic/biogenic source discrimination.

01 ABB has a range of products that detect gas leaks, such as the ABB Ability™ MobileGuard™ shown here. An optical reinjection technique improves methane and ethane detection precision.

methane and ethane leak detection systems indispensable. Such systems must be capable of rapidly identifying and locating leaks and providing operators with high-accuracy real-time data. By meeting these requirements, the risks associated with gas transport and storage can be mitigated, the environment and public safety protected and regulatory compliance ensured.

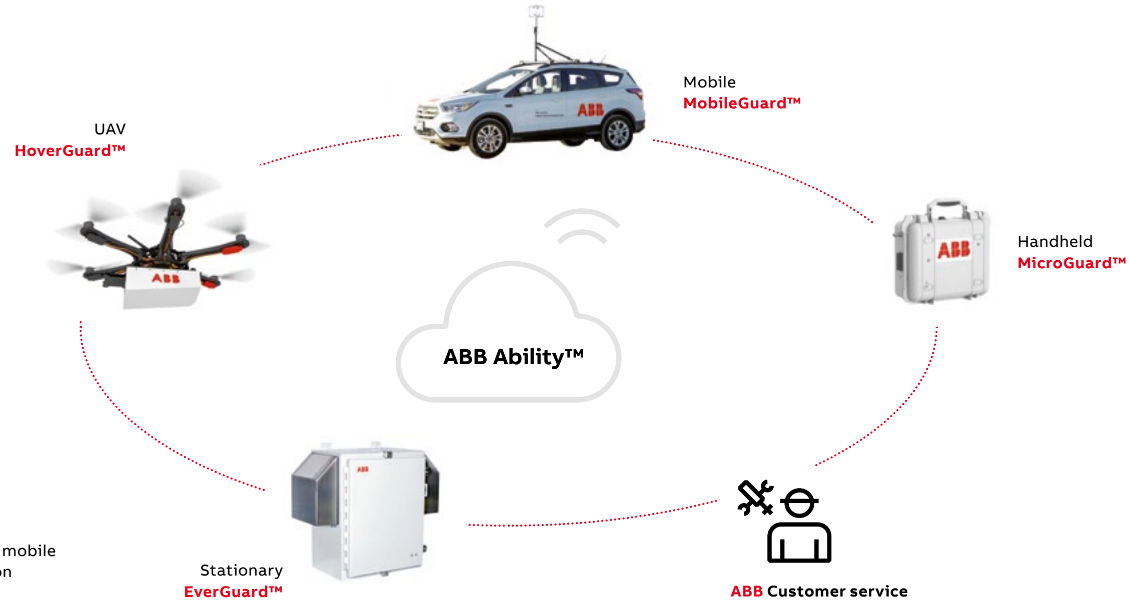
Why simultaneous detection of methane and ethane is important

The simultaneous detection and quantification of methane and ethane is essential in determining the origin of these gases. Biogenic methane is derived from natural sources such as wetlands, while thermogenic methane is produced from fossil fuels. In contrast, ethane is almost exclusively generated from fossil fuels. Therefore, a higher-than-expected ratio of ethane to methane can indicate that the methane emissions primarily originate from fossil-fuel sources rather than natural sources »01. A simultaneous measurement of both gases can determine this ratio.

ABB is the first and only company offering four different solutions for detecting, finding, quantifying, and mapping leaks.

Off-axis but on-target

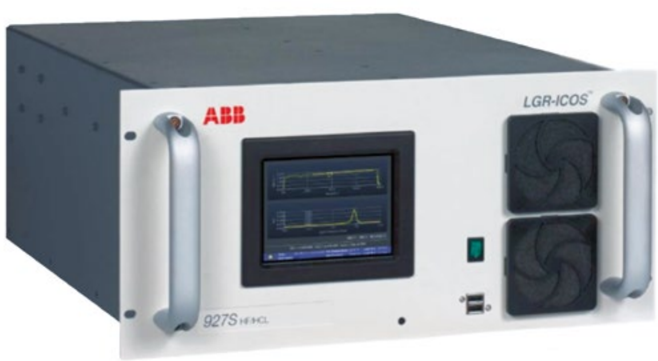
ABB provides a wide range of gas analyzers based on unique and proprietary off-axis integrated cavity output spectroscopy (OA-ICOS) technology. OA-ICOS represents the fourth generation of cavity-enhanced tunable diode laser absorption spectroscopy (TDLAS). This technology has revolutionized the detection and quantification of gases. More specifically for natural gas leak detection, the laser-based analyzers offered by ABB can be deployed quickly to identify and quantify gas leaks in the field. ABB is the first and only company offering four different solutions for detecting, finding, quantifying and mapping leaks of natural gas while driving, walking, flying or stationary »02. »



02 ABB Ability™ mobile gas leak detection system family.

These ABB gas analyzers are based on so-called non-mode-matched optical cavities, in which only a very small amount of the laser light enters the cavity to interact with the gas sample. The research project described here aims to increase the amount of light entering the cavity – and thus available for measurement purposes – by using an optical reinjection technique. The reinjection arrangement will:

- Replace costly high-power laser sources with cost-effective and low-power alternatives
- Enhance the signal-to-noise ratio, improving instrument precision to expand the sales volume and available market
- Enable use of higher reflectivity cavity mirrors, thus increasing the optical path length and thus measurement sensitivity
- Tolerate a wider range of mirror reflectivity values, thus increasing allowable mirror manufacturing margins.



03 The ABB GLA231 series of analyzers provides highly sensitive and accurate measurements of gases such as HF, HCl and NH_3.

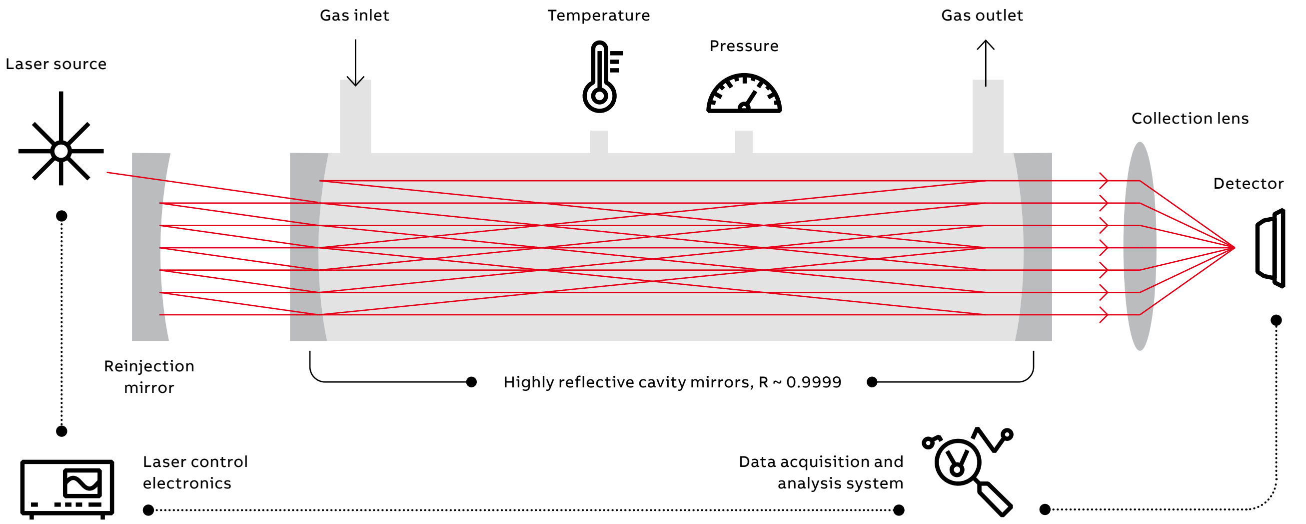
Simulate, design, build and test

The development and realization of optical reinjection in an OA-ICOS analyzer was carried out in several steps. Firstly, ABB developed an optical simulation tool to establish optimal and feasible configurations. Secondly, the team designed and set up an interband cascade laser (ICL)-based OA-ICOS analyzer that demonstrated optical reinjection. The

The optical reinjection technique can be applied to the whole ABB OA-ICOS platform for various applications.

ICL produces the wavelengths of interest for the spectroscopic examination of the gases of interest. The demonstrator is based on a commercially available ABB GLA231 series analyzer »03, in which key elements such as the laser source, the photodetector, the cavity mirrors, the collection lens and the reinjection optics, were replaced by custom versions appropriate to the aims of the project.

Various optical reinjection scenarios suggested by the simulations were experimentally investigated and verified. Measurements of precision, accuracy, linearity and cross-interferences for methane, ethane and water were performed with and without optical reinjection. Although the optical reinjection technique is implemented in the methane/ethane OA-ICOS gas analyzer in this project, it can be applied to the whole ABB OA-ICOS platform for various applications.



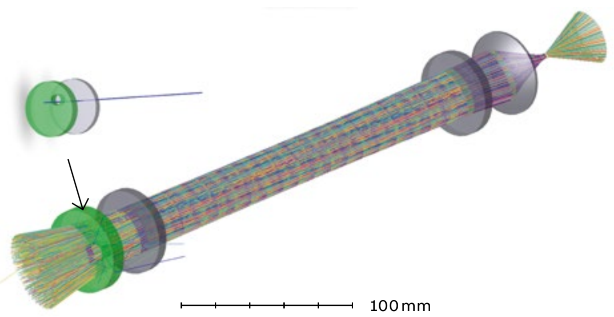
04 OA-ICOS analyzer schematic.

Three mirrors and a cavity
The OA-ICOS analyzer with an optical reinjection system consists of an ICL laser source coupled to an optical cavity with two high-reflectivity mirrors, a photodetector to measure and monitor the signal, the laser control electronics, the data acquisition and processing system and a reinjection mirror »04.

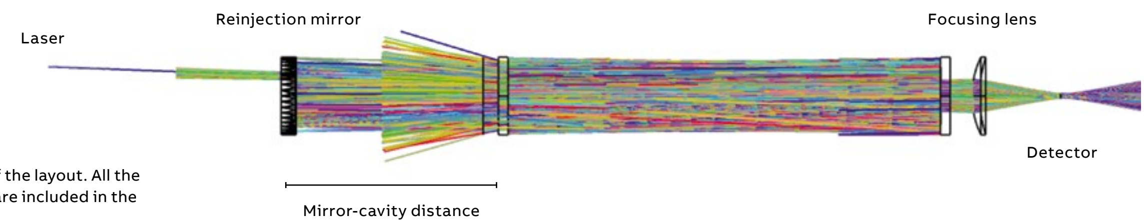
Initially, the laser beam passes through a small hole in the reinjection mirror and enters the first cavity mirror. The interior of the "exit" side of this mirror is highly reflective (99.99 percent) so most of the light is reflected internally, back towards the reinjection mirror, where it is reflected back to the cavity mirror. Since this three-mirror configuration is optically stable, this process repeats itself continuously. Only about 0.01 percent of the available light enters the cavity at each pass. The light that does enter the cavity 'bounces' between the cavity mirrors in an off-axis manner. A small fraction of the light leaks out through the rear cavity mirror, giving rise to the ICOS signal that indicates the concentrations of the various gases present in the cell. The output light is then focused onto a photodetector with a suitable collection lens. The gas sample is pumped continuously through the cavity using a vacuum pump and the pressure inside the cavity is controlled and measured. The temperature inside the cavity is measured with a temperature sensor.

The optical system must be carefully designed, particularly with regard to the reinjection mirror.

The optical system must be carefully designed, particularly with regard to the reinjection mirror. Further, given the space limitations of the final product, the dimensions of the reinjection setup are critical. To design an appropriate optical reinjection setup, ABB used an optical simulation software called Zemax that integrates all the features required to conceptualize, design, optimize, analyze and »



05a Optical configuration: 3-D layout in Zemax.



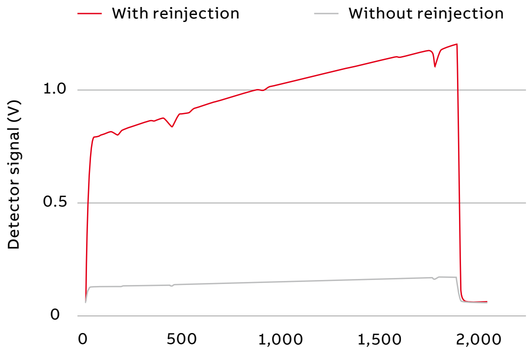
05b Side-view of the layout. All the key components are included in the simulation.

06 Measurement of ambient air in the laboratory.

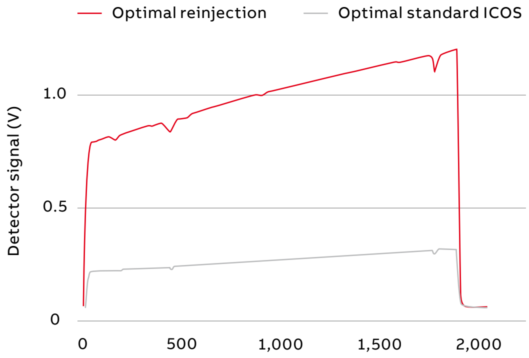
06a Comparison of detector signal with and without optical reinjection.

06b Comparison of the detector signal of optimized optical reinjection and optimized standard ICOS.

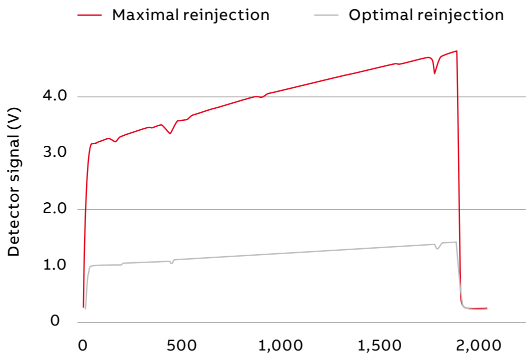
06c Comparison of the detector signal of optimized optical reinjection and maximized optical reinjection.



06a



06b



06c

document any optical system. Zemax is widely used in the optics industry as a standard design tool.

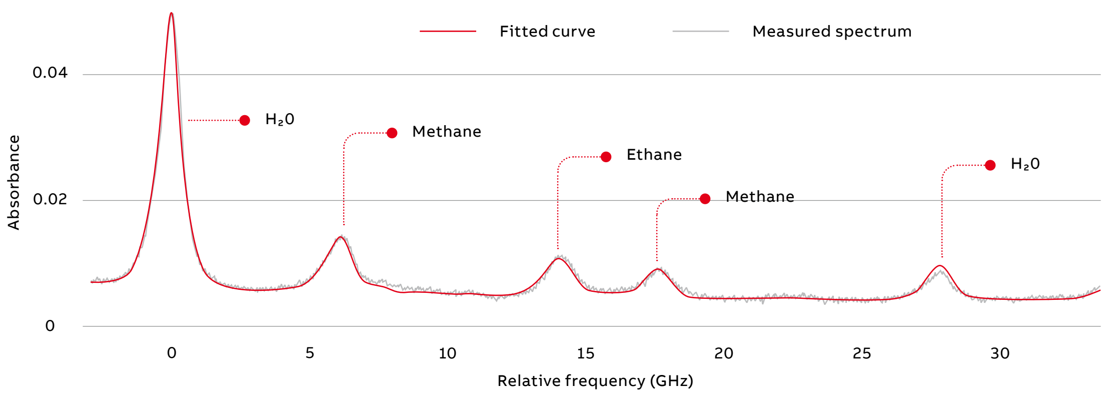
The simulation and experimental setup are configured identically, having the same mirror curvatures, clear apertures, reinjection mirror hole size and offset, beam diameter and divergence, collection lens surfaces and detector size »05. The efficiency of optical reinjection is related to the position of the hole, the distance between the reinjection mirror and the cavity, the curvature of the reinjection

The enhancement of the signal amplitude is around a factor of four with optical reinjection.

mirror and the incident angles of the laser beam. In the end, ABB experimentally tested and verified six reinjection mirrors with various cavity-mirror distances.

A spectrum of results

»06a compares the signal with and without optical reinjection (by simply removing the reinjection mirror and keeping the laser alignment unchanged). Here, the alignment of the standard ICOS system may not be optimal in terms of signal amplitude and the signal-to-noise ratio. The comparison results after re-optimizing the standard ICOS (without the reinjection mirror) by adjusting the laser beam direction are shown in »06b. The enhancement of the signal amplitude is around a factor of four with optical reinjection. However, as shown in the optical simulation, more reinjection power is expected and can be obtained by further tweaking the incident angle of the laser beam »06c, though this significantly increases optical noise.



07 Absorbance spectra from the measurement in ambient air. The absorption peaks from three species are indicated. The absorbance spectra act as a spectral fingerprint, unique for each gas species.

The concentration of the gas samples and the measurement precision for a given period can be obtained by fitting the spectra. First, the measured time-dependent transmission spectra must be turned into frequency-dependent absorbance spectra. A physics model is then fitted to the recorded spectra, using species-specific spectral data from the HITRAN database (a molecular spectroscopic library) as input. The proportionality between the area of the fitted line shape and the directly measured parameters (ie, gas temperature and pressure) allows direct inference of the gas concentration. »07 shows the absorbance spectra measured in ambient air, where the absorption from methane, ethane and water vapor can be observed. In the case of optimized optical reinjection (upper curve in »06b), the demonstrator is capable of reporting methane and ethane concentrations in ambient air continuously with a precision of 10 parts-per-billion (ppb) for methane and 15 parts-per-trillion (ppt) for ethane over one second of measurement time.

Reinjection perfection?

Various optical reinjection configurations have been simulated, set up, optimized and characterized to construct a mid-IR OA-ICOS analyzer for the simultaneous measurement of trace-level methane and ethane. Compared to incumbent methods, the precision achieved for the ethane measurement is enhanced by three orders of magnitude as a result of employing optical reinjection, a new laser source and optimization of the mirror and detector. With this high precision, the discrimination between thermogenic and biogenic sources in the field can be improved significantly. Based on gas dispersion simulations, the leak attribution accuracy can be improved to nearly 98 percent across all the simulated leak rates and conditions. This improvement reduces the time surveyors spend investigating biogenic emissions, allowing them to focus on actual natural gas emissions, thus enhancing the safety of natural gas grids. In addition, optical reinjection can be leveraged for the entire ABB ICOS product platform to improve overall performance on applications requiring high gas measurement sensitivity. •

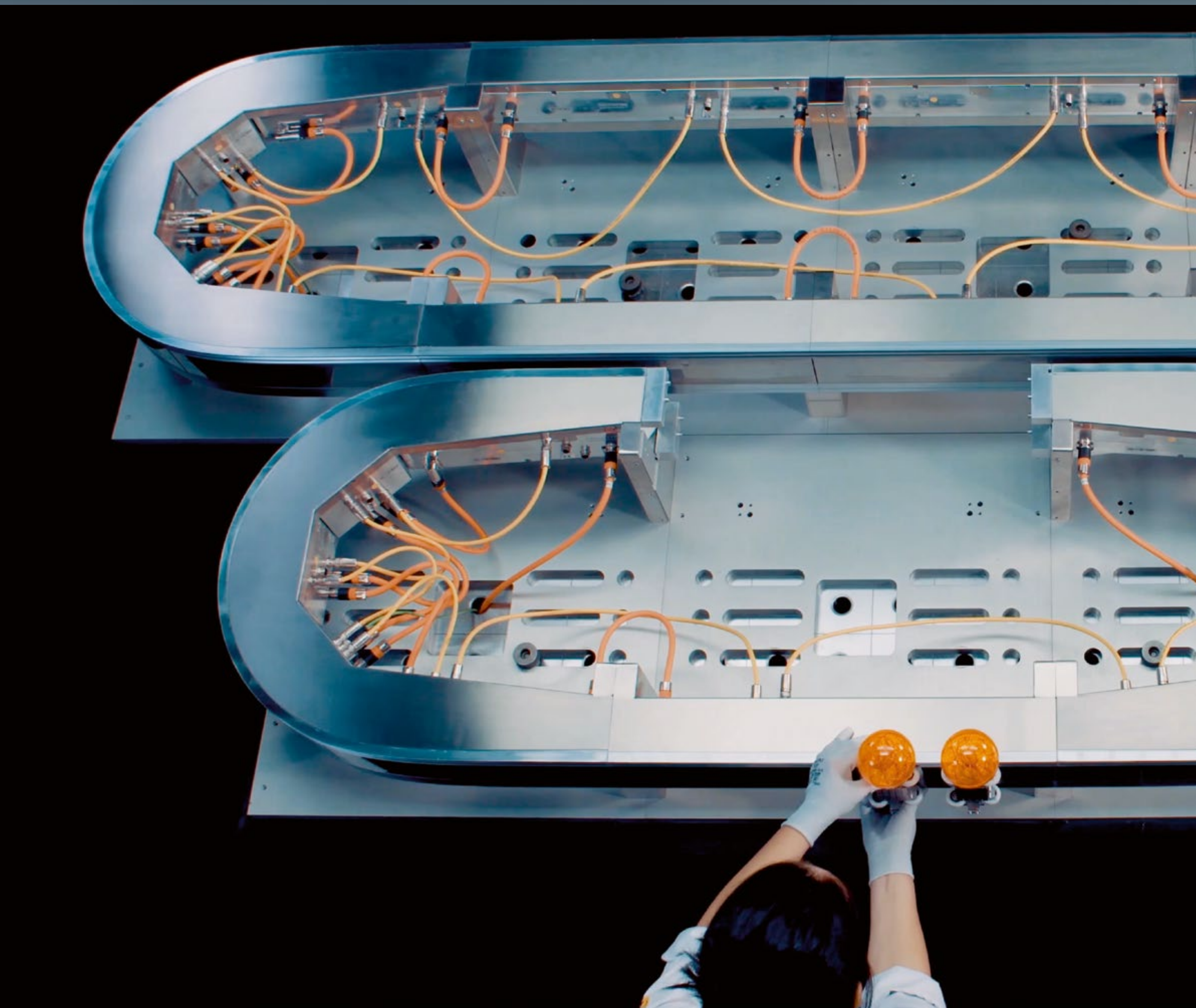
Compared to incumbent methods, the precision achieved for ethane measurement is three orders of magnitude better.

—
Originally published in ABB Review 04/2024, pp. 278–283.

Read the original article here:
https://global.abb/group/en/innovation/news/high-precision-gas-measurements

—
ACOPOSTRAK: HEADING FOR HIGH SPEED FLEXIBLE MECHATRONIC PRODUCTION LINES

BUILDING ADAPTIVE SOLUTIONS



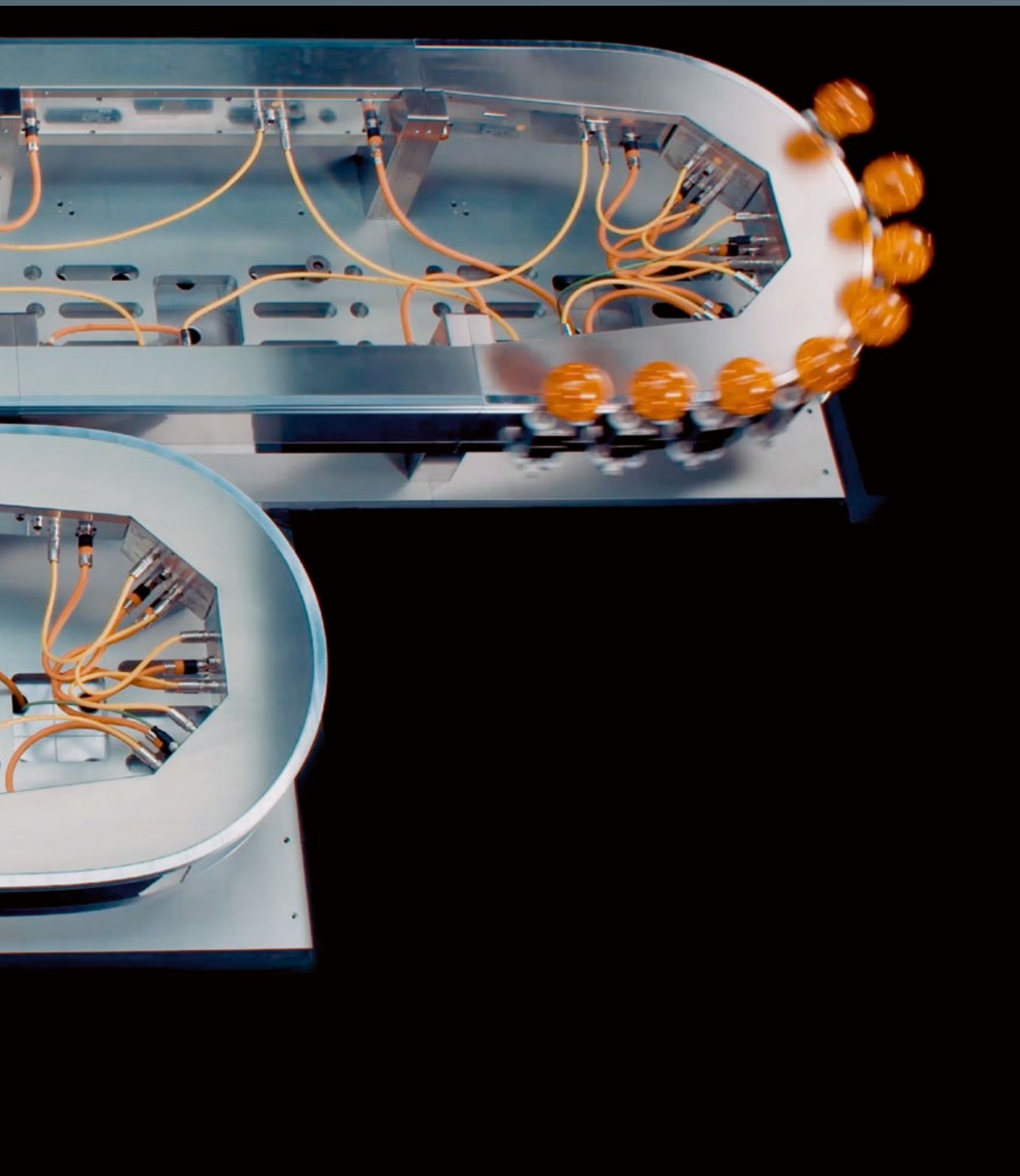
Key Facts

Modern manufacturing requires flexible solutions beyond traditional conveyors.

ABB's ACOPOStrak is an induction-based transportation system within factories.

Individual items are whisked along different paths for high customizability.

To produce economically under increasingly variable conditions, manufacturers are building adaptive solutions based on mechatronic product transport systems like ACOPOStrak from B&R, the Machine Automation division of ABB.



01 Changes in consumer behavior and expectations are among the driving factors behind a major shift in how products are made, assembled and packaged.

—
Johannes Vitzthum
ABB Robotics & Machine Automation
Eggelsberg, Austria

johannes.vitzthum@br-automation.com

Isabella Back
has changed her role in ABB, please direct enquiries to the other authors.

Armin Steinhauser
ABB Robotics & Machine Automation
Linz, Austria

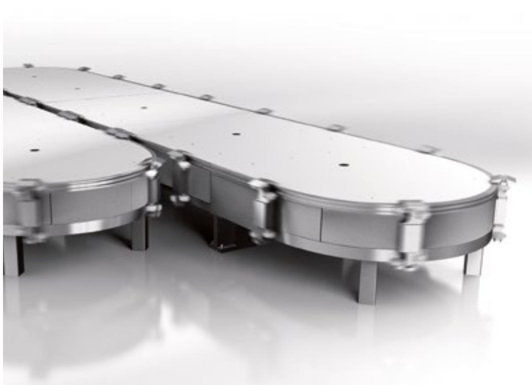
armin.steinhauser@br-automation.com

When a social media sensation sparks an overnight spike in demand, adaptive manufacturing lines are quick to pivot and catch the first wave. When digital consumers expect personalized products to be delivered quickly, adaptive systems adjust on the fly to run mixed batches and package custom sets »01. When production moves closer to consumers, adaptive solutions make small batches – even batches of one – with the efficiency of mass production.

Adaptive manufacturing solutions get their special blend of intelligence, flexibility and efficiency from today's most advanced automation: smart software and open communication standards, digital twins and edge computing, perfectly coordinated robotics and AI-enhanced machine vision »02. Fully adaptive manufacturing might not be possible, however, if it were not for a major evolution in one particular area: product transport.

Unlocking adaptability with mechatronic transport
When people imagine a production line, the first thing that comes to mind is often a conveyor belt. Yet, because of the rigid sequence of processes they impose, these mechanical movers often limit flexibility.

Mechatronic systems, based on linear or planar motor technology, work much differently. Here, the motor itself essentially becomes the conveyor. In a linear motion system like ACOPOStrak, the active part of the motor – the

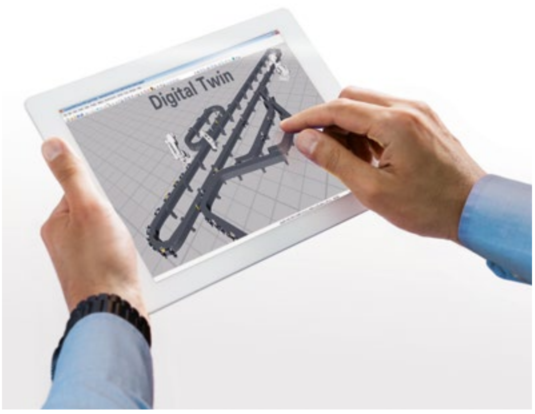


stator – extends along the entire transport path in the form of a track. The passive part of the motor – the rotor – takes the form of shuttles that carry products along the track. This evolution elevates product transport into a new role. What before may have been considered auxiliary equipment – a means to an end – now becomes the intelligent, value-adding backbone of the entire operation. By eliminating mechanical

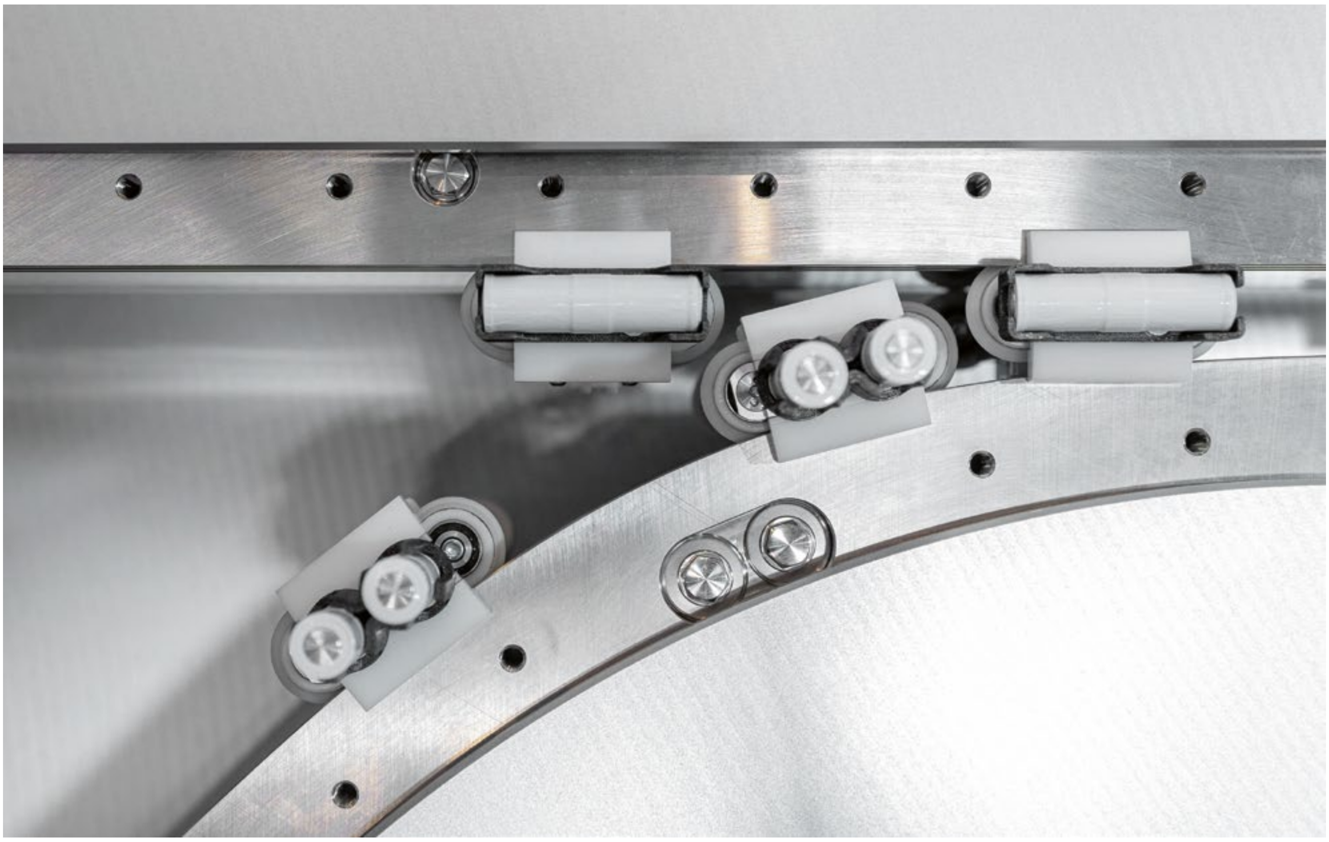
The active part of the motor – the stator – extends along the transport path in the form of a track.

power transmission, mechatronic systems avoid slippage and wear, enabling movements with both high dynamic forces and extremely repeatable precision. Whether the product is an electric vehicle battery cell, an insulin pen or a package of chips – such systems carry each item on a unique high-speed journey through production.

In these dynamic processes, shuttles must often switch from one track to another – whether to sort out defects, balance flow among parallel processing stations, or merge products from different lines to be packaged in custom sets. While other linear motion systems do this using mechanical »

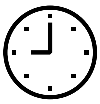


02 Adaptive manufacturing solutions combine today's most advanced automation technologies to meet evolving production demands.



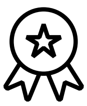
03 Fully contactless ACOPOStrak diverters allow high-speed track switching with no mechanical wear – leading to higher productivity with less downtime.

Adaptive manufacturing: The OEE multiplier



Availability
- Little to no downtime for changeover
- Individual control for fault tolerance
- Automatically skip malfunctioning stations

Performance
- Non-stop productivity: On-track and in-motion
- Processing density: No wasted conveyance or handling
- Parallel processing: Multiply performance, not footprint



Quality
- Tighter synchronization for tighter tolerances
- In-line quality control and early defect removal
- Robust fault tolerance for reduced waste production

04 Adaptive manufacturing solutions combine flexibility with productivity by targeting all three factors that define overall equipment effectiveness (OEE).

rails or gates, B&R's ACOPOStrak diverter technology »03 is fully contactless. In the absence of wear-prone moving parts, switching can occur at full production speed (up to 5 m/s) with no spacing needed between the 50-mm shuttles.

This patented technique is based on the principle of field weakening. In field-oriented control, the motor's magnetic field comprises two perpendicular components: the d-axis (direct axis) produces the force that holds the shuttle on the track, while the q-axis (quadrature axis) produces the torque that propels the shuttle along the track. When a shuttle is moving between two opposing tracks, reducing its d-axis current weakens the magnetic field generated by the stator and allows the shuttle to switch sides.

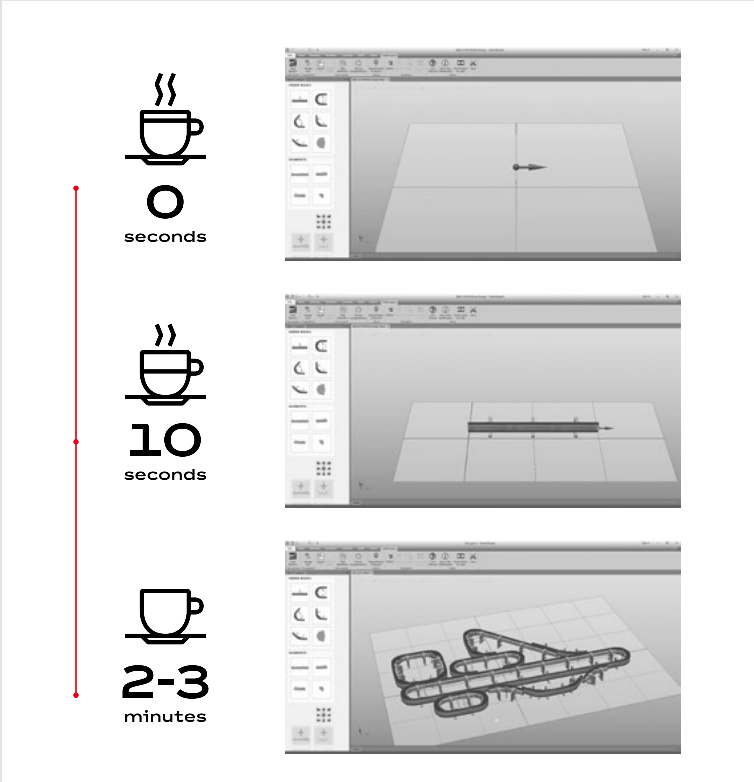
Adaptability at scale: High speed and high density
With highly dynamic control over the speed, spacing, and routing of each individual shuttle, it becomes possible to perform processing steps all along the track, even while in motion. Square meters otherwise consumed by buffer zones and empty stretches of conveyor instead host high-value processing. Different products – with different sizes, shapes, contents, closures or labels – each get exactly the treatment they need with minimal intervention or downtime. The ability to split and merge product flows at full speed makes it possible to optimize throughput with asynchronous workflows and robust fault tolerance. Together, these and other capabilities target all three factors that define overall equipment effectiveness (OEE) »04, thus helping manufacturers maintain high availability, performance, and quality as they adapt on the fly.

»

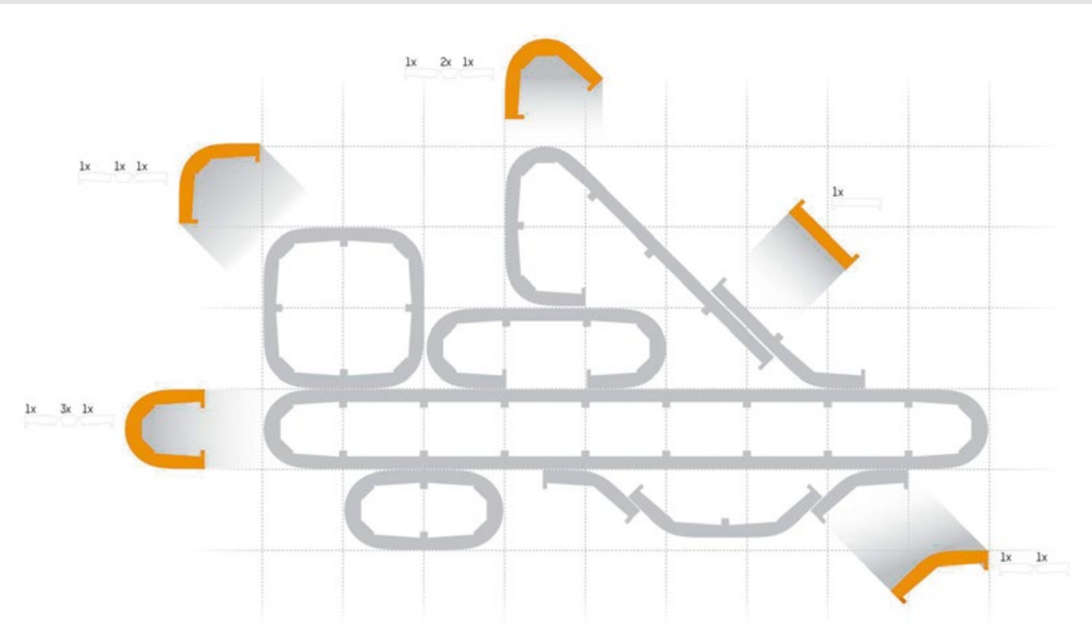
05 Accelerating system design with a no-code 3D graphical simulator

Building adaptive solutions is a multidisciplinary exercise that often requires a fundamental shift in thinking about system design. By making it possible to visualize and simulate the performance of these systems, ACOPOStrak Designer provides easy access for all stakeholders, enables more efficient decision making and helps ensure predictable results. Once all are satisfied, software is generated at the push of a button with no coding required. As an add-on for RobotStudio®, ACOPOStrak Designer makes it possible to configure a track layout in minutes using drag-and-drop 3D graphics »05a. Clear visualization makes it easy to design around obstacles and simulate production flows – even for mixed batches of different products. Right from the outset, users can assess the feasibility of their ACOPOStrak system, accurately predict its future throughput and ensure safe, energy-efficient operation.

With no complex programming, ACOPOStrak Designer realistically simulates performance under real-world operating conditions. This includes the thermal behavior of each track segment, allowing strategic placement of water-cooled segments to enhance throughput while planning the power supply and design to optimize energy consumption. Analysis functions help determine the optimum number of shuttles to use, as well as their maximum speeds and payloads. Users can quickly test and evaluate different layouts and scenarios »05b to maximize both throughput and energy efficiency, then fine tune shuttle movements and other parameters before quickly exporting their design to B&R's engineering software, Automation Studio.



05a Adaptive manufacturing lines can be laid out in minutes by dragging and dropping 3D graphics in ACOPOStrak Designer, an add-on for RobotStudio®.



05b ACOPOStrak's flexible layout adapts to the needs of each site and application.

A B&R customer in the cosmetics industry illustrates the order-of-magnitude impact this can have: Replacing three of four existing lines with one adaptive line, they reduced their footprint by around 75 percent while maintaining nearly the same overall output. The changeover time to a new bottle with a pump closure went from three hours to 40 minutes. A batch of shampoo and conditioner with a new bottle design that would have taken around three weeks was produced in five days. Similar results can be seen from customers in many other sectors as well, such as one in medical device assembly who achieved a double-digit boost in productivity while at the same time reducing the footprint of the machine – avoiding the cost of building a new cleanroom.

The integral role of software

Bringing the inherent flexibility of mechatronic hardware to life, software and AI add value throughout the lifecycle of an adaptive manufacturing solution. This process begins at the very first steps of conceptual design, where ACOPOStrak Designer »05 makes it easy to plan a layout graphically, simulate throughput and energy consumption, and then generate the machine code at the push of a button. During operation, advanced control software dynamically routes

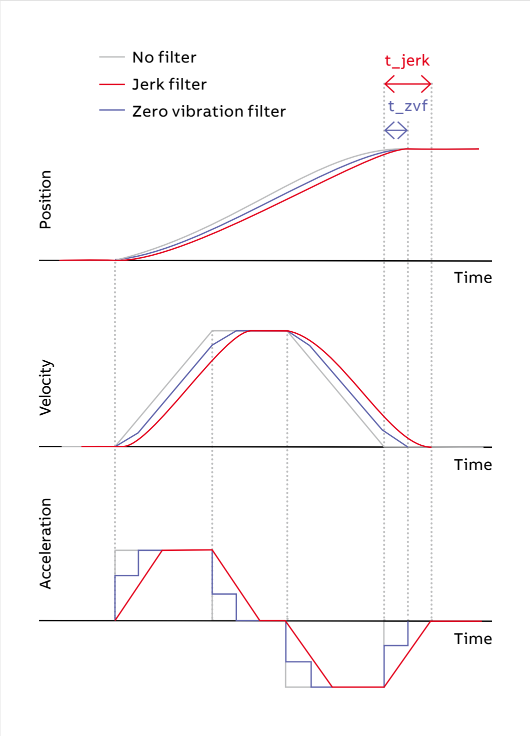
One customer in medical device assembly achieved a double-digit boost in productivity.

products through an easily configurable network of processing stations, adapts movement parameters to changing conditions, synchronizes with pick-and-place robots and other systems, and makes subtle adjustments to ensure gentle product handling and minimize sloshing of liquid contents during high-speed operation »06. Each shuttle is uniquely identified and inherently provides an array of data, which can be visualized in the ACOPOStrak Monitor app »07, and used on site or remotely for data-driven, AI-enhanced performance optimization and predictive maintenance. •

06 Sophisticated trajectory planning for gentle handling at top speed

During high-speed production, ACOPOStrak generates up to 85 newtons of propulsion force and accelerates its shuttles at up to 10 g – similar to what a fighter jet pilot experiences while pulling a sharp aerial maneuver. If applied without filtering (see the blue line in the graph), that would be enough to exceed the physical limits of products or their holders. Sophisticated trajectory planning integrated in ACOPOStrak software ensures that highly dynamic processing steps are completed exceptionally quickly to maximize throughput without damaging products or spilling liquid contents.

A jerk filter (see yellow line) would be one way to eliminate the jump in the acceleration's derivative and reduce oscillations, but it would extend the overall movement duration by the filter time. Instead, zero vibration filtering (see purple line »06a) minimizes oscillations much faster (typically by around 50 percent) by targeting critical frequencies – those that stimulate a resonance frequency and the first two harmonics. If the system in question is a fluid-filled container with known dimensions, the frequency is calculated by the system; if not, the measured eigenfrequency can be used as input.



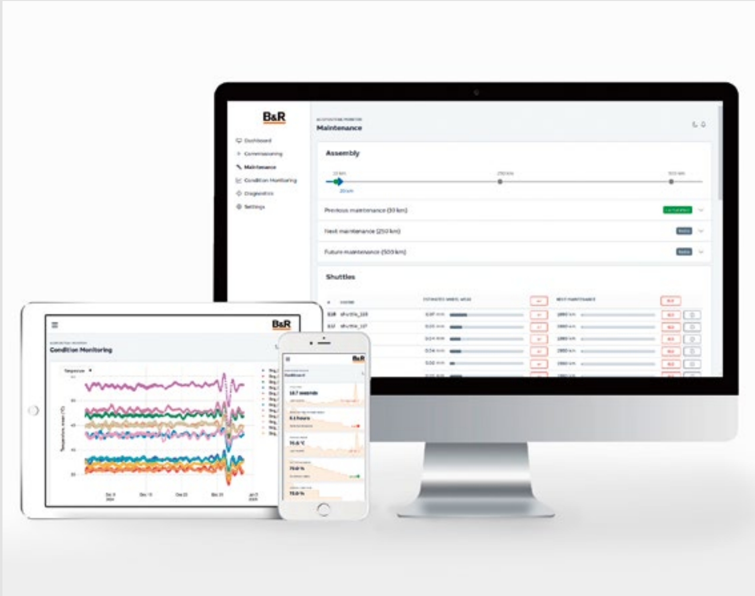
06a Zero vibration filtering ensures gentle product handling while allowing more dynamic movements for higher productivity.

—
Originally published in ABB Review 01/2025, pp. 34–41.

Read the original article here:
https://global.abb/group/en/innovation/news/high-speed-flexible-mechatronic-production-lines

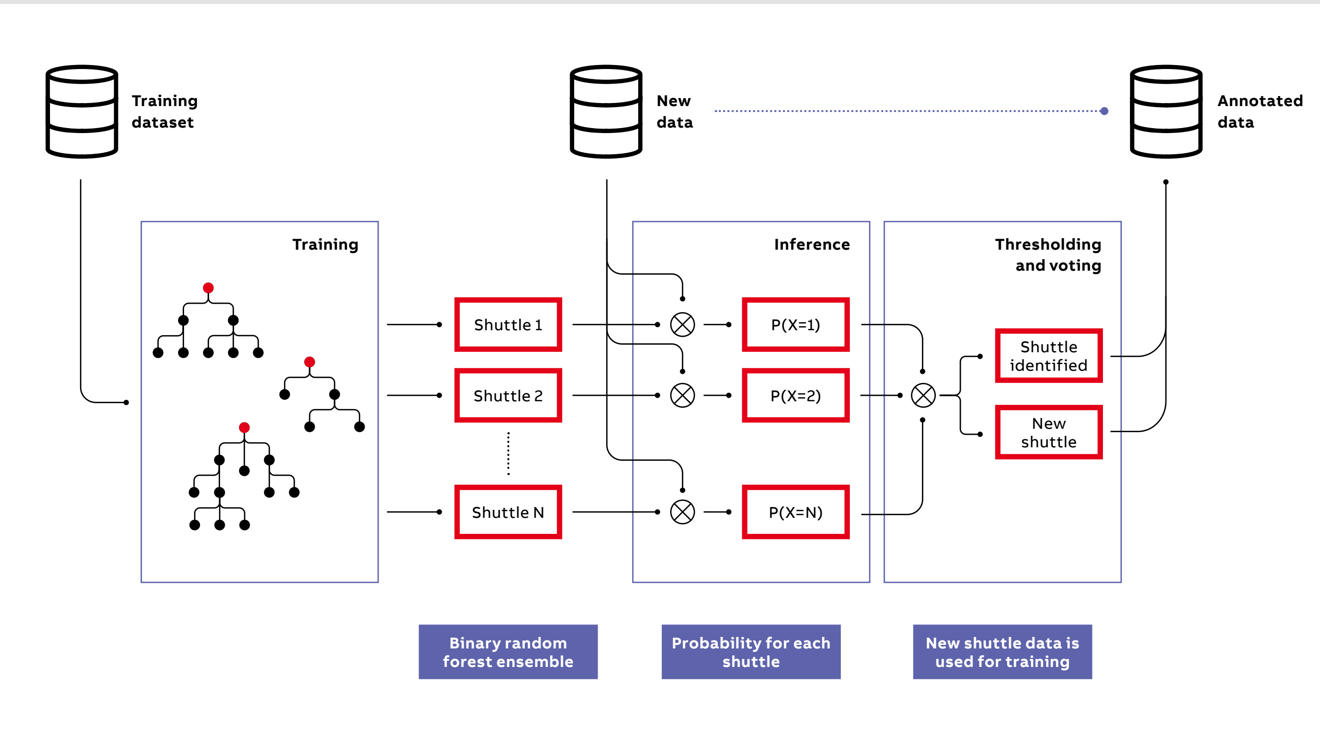
07 AI-enhanced insight for uninterrupted productivity

When machine builders are able to visualize the real-time status of their systems, they can identify issues proactively and implement predictive interventions to streamline maintenance. As an Industrial IoT app with optional cloud connectivity, ACOPOStrak Monitor provides data-driven, AI-enhanced services for collecting, analyzing and visualizing production »07a. These features are crucial for optimizing performance and maintenance to ensure uninterrupted high productivity throughout the system's entire lifecycle. An API allows machine builders and operators to access system data remotely and integrate it in custom solutions for further analytical or troubleshooting tasks.



07a The ACOPOStrak Monitor dashboard provides operators with a clear and immediate overview of machine status. Condition monitoring offers unique insights through machine data collected over time.

State-of-the-art AI algorithms provide unprecedented opportunities to monitor, analyze and ultimately optimize machine operation. Examples include regression-based estimation of shuttle condition and anomaly detection through unsupervised learning. Continuous monitoring of machine data facilitates more efficient maintenance planning, reduces downtime and extends equipment life. Importantly, the ability to derive shuttle condition from ordinary sensor data means this can be done without having to install any additional hardware. This is assisted by the classification of physical shuttle entities using an ensemble-based random forest approach – a type of machine learning – to ensure correct attribution of collected data »07b.



07b Shuttles are identified using ensemble-based random forest classification, which helps support shuttle condition monitoring without needing additional sensors.

MASTERING COMPLEXITY WITH AUTOMATED GRID DESIGN AND ANALYSIS

GREENER GRIDS

Boda Li
Chen Song
Theresa Loss
Matthias Biskoping
Jan Schlake
ABB Corporate Research
Mannheim, Germany

boda.li@de.abb.com
chen.song@de.abb.com
theresa.loss@de.abb.com
matthias.biskoping@de.abb.com
jan-christoph.schlake@de.abb.com

Maryam Sharifi
Fredrik Ljungberg
Stefan Thorburn
ABB Corporate Research
Vaesteras, Sweden

maryam.sharifi@se.abb.com
fredrik.ljungberg@se.abb.com
stefan.thorburn@se.abb.com

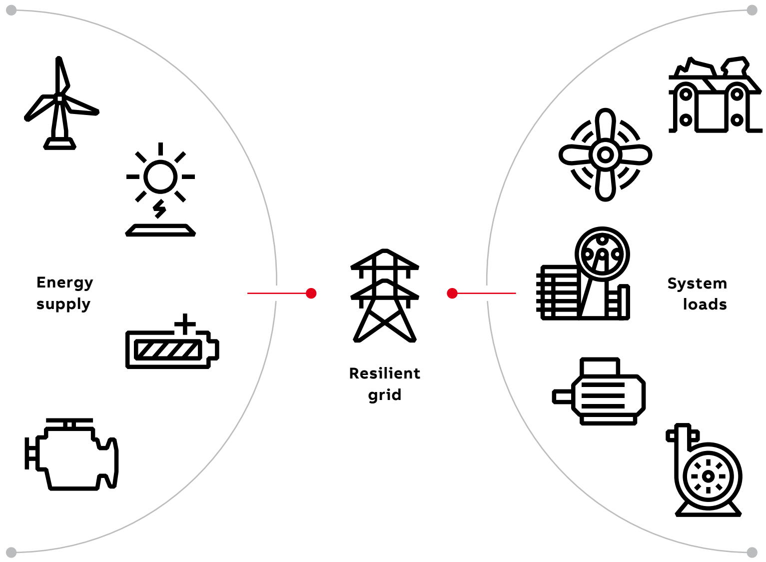
Nic Beutler
Process Industries,
Baden-Daettwil, Switzerland

nic.beutler@ch.abb.com

How can energy intensive industries, such as mining, meet sustainability goals? ABB is evolving its industrial grid design toolbox to enable automatic analysis of various load, supply, and potential fault scenarios in the industrial power grid, thereby assisting engineers, designers, and operators in making design and operational decisions.

Key Facts

- Industrial power grids integrate renewable and conventional generation.
- Use of renewable energies must be optimized while assuring stability of operations.
- ABB provides frameworks and tools such as ABB eMine™ that guide decisons.



01 Industrial system grid layout based on renewable energy integration.

By merging detailed models of consumption, equipment and operations with historical data on renewable energy, ABB's mine electrification framework, ABB eMine™, efficiently optimizes the integration of renewable power sources into industrial settings – particularly in the mining industry.

To reach worldwide sustainability targets, the integration of renewables is of paramount importance. This is particularly critical in industries such as manufacturing, mining, hydrogen production, and data centers, which are known for their substantial energy needs. Transitioning these industries to operations based on renewable energy is not only about meeting their high energy demands but also about aligning these efforts with further targets such as carbon emission reductions and

ABB's approach improves the grid layout design process by incorporating an adaptive feedback loop.

sustainability. The availability of renewables at any given industrial plant varies significantly based on daily and seasonal variabilities. This means that each installation must be individually designed to meet the specific needs of its industry. Current estimates foresee renewable sources supplying 45 to 50 percent of global electricity generation by 2030, and between 65 and 85 percent by 2050 [1].

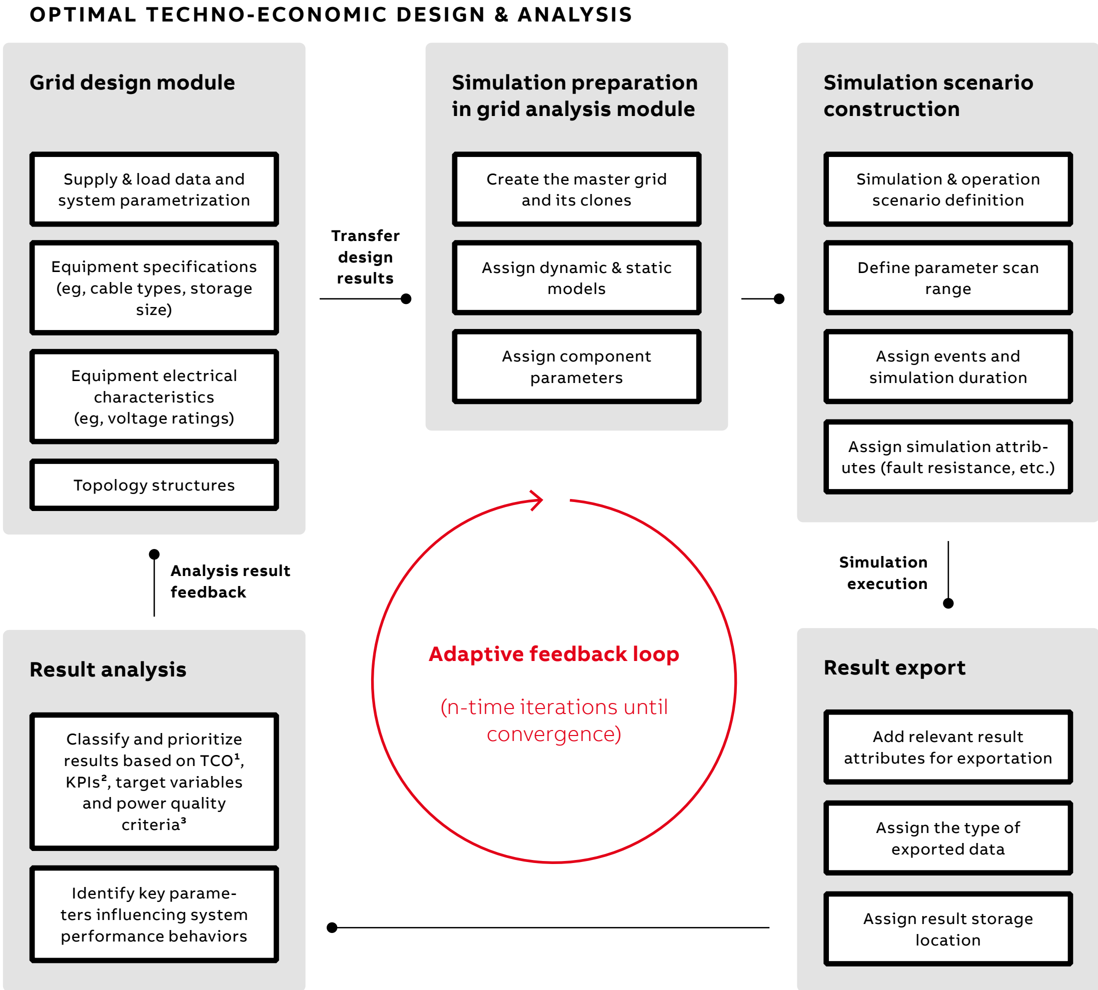
However, integrating renewable energy into the industrial power grids of these industries presents a set of unique challenges, because each industry exhibits complex operational dynamics that lead to intricate load changes. The variabilities of renewable energy sources further escalate this complexity, making the analysis and design of power grids challenging. Moreover, to ensure the smooth operation of any given industrial site, there must always be a balance between multiple energy sources and the overall energy-consuming system »01. In almost all cases this requires a continuous power supply all year round. In addition to the challenges on supplies and loads posed by intermittency, an increasing number of applications utilize non-linear equipment, which adversely affects power quality through harmonics and power factor issues. As a result, poor power quality impacts asset health and performance, resulting in asset failure, diminished or halted production, and eventually plant financial losses.

To address these challenges, industrial power grids must efficiently and reliably handle a mix of conventional and renewable energy sources and accommodate diverse operational scenarios ranging from steady state to transient conditions. Furthermore, the scarcity of skilled professionals in many industrials domains, combined with the imperative to integrate user requirements early in industrial power grid development, emphasizes the urgent need for innovative and efficient design solutions. This need is further compounded by the absence of automated solutions and the lack of accessible and useful data, which significantly delays the incorporation of renewables into systems. The resulting prolonged integration periods can lead to substantial additional costs, delays and intensified challenges associated with renewable energy adoption.

Detailed simulation models and advanced framework

In today's industrial power grids, understanding the intricate dynamics of equipment energy consumption is essential for renewable energy integration, increasing operational efficiency, and reducing costs. To better support the grid design and corresponding analysis, ABB sets out to provide an in-depth analysis of energy consumption in industrial settings by precisely modelling different assets. These models can capture the unique profiles from diverse stationary systems, such as motors and charging stations, as well as from mobile assets such as trucks.

ABB also considers how dynamic changes in plant operations affect models. For this purpose, a sufficiently generic modeling framework is used to capture physical effects across different time frames. Additionally, industrial plants often operate assets from various manufacturers, and because data and expertise are valuable, it is crucial to keep the models' implementation details confidential. To »

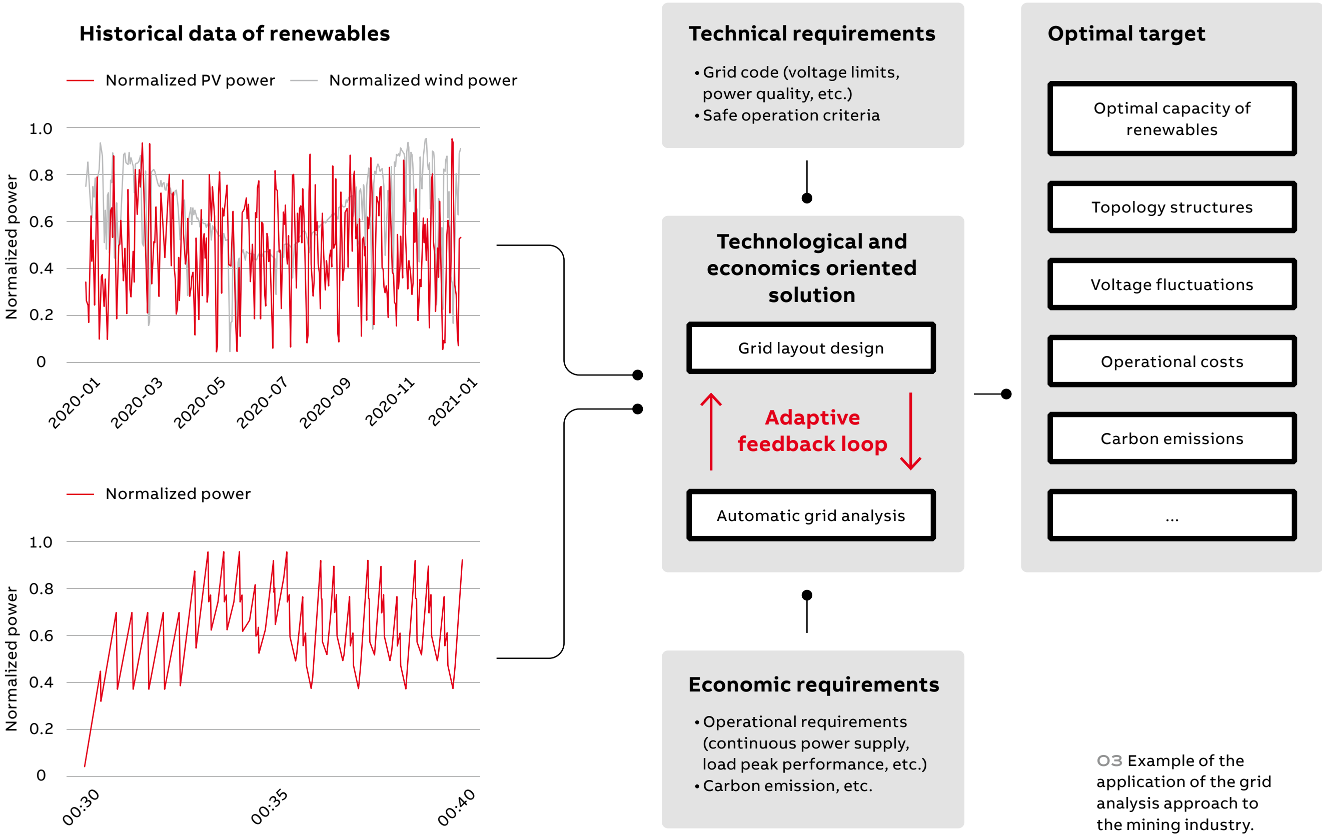


02 Workflow in an optimized techno-economic grid design & analysis solution

[1] TCO: Total cost of ownership

[2] KPI: Key Performance Indicator

[3] Power quality criteria: Power quality criteria refer to the set of standards and guidelines used to evaluate the stability and reliability of the electrical power. These criteria cover various aspects of power supply, including voltage/frequency stability, power supply continuity, harmonic distortion, power factor, etc. High-quality power is essential for the smooth operation of industrial equipment and for reducing the risk of operational disruptions.



03 Example of the application of the grid analysis approach to the mining industry.

tackle these issues, ABB uses the interoperability standardized Functional Mockup Interface (FMI) concept, allowing for the combination of models from different manufacturers. Functional Mockup Units are simulation model files conforming to the FMI standard, comprising only compiled code. This allows the protection of model integrity as well as the intellectual property of sources. The models and simulation framework facilitate the creation of load profiles for grid design and seamless integration into the grid analysis module.

Automated industrial power grid design

With these challenges in mind, ABB has developed an automated power grid design and analysis framework that is specifically tailored for modern industrial power grids and is focused on the challenges encountered during renewable energy integration. As illustrated in »02, this solution leverages advanced optimization theory and thorough analysis of power grid conditions to streamline the design process, enhance operational efficiency and reduce the costs associated with integrating renewable energy.

Indeed, the integration of renewable energy is the essential ingredient in achieving truly sustainable systems. However, as many users have discovered, integrating renewable energy into industrial energy systems often necessitates upgrades or modifications to the existing power grid to accommodate complicated load dynamics and renewable energy fluctuations. ABB addresses this challenge by developing a comprehensive automation of grid design solutions tool suite. This innovative approach improves the grid layout design process by incorporating an adaptive feedback loop informed by dynamic analysis results.

The process begins with collecting data and user-specific requirements, including key performance indicators such as building costs and carbon emissions. This are used to generate an optimal power grid layout that takes elements such as voltage ratings and cable selection into account. This layout is achieved by solving an optimization model that adheres to grid codes and static operational safety margins. Optimized grid design results can then be read by current mainstream power system analysis software. The system automatically generates various simulation and operational scenarios for transient behavior analysis, such as switching and faults. In addition, the feedback loop adjusts the design through iterations based on transient analysis, thus refining its layout solutions.

The method's advantages are manifold and significant. Compared to the traditional approach, it can dramatically reduce design time, thus streamlining processes. In addition, the combination of steady-state and dynamic transient analyses ensures high safety and accuracy, guaranteeing that the grid layout meets diverse operational requirements under various scenarios. The adaptive feedback loop further reinforces reliability, satisfying both static and dynamic safety criteria. Importantly, the method's design is reusable and scalable, making it applicable to a range of industries such as mining, marine and port operations, as well as hydrogen production plants.

Industrial power grid performance optimization with automatic transient analysis

Today, the prevailing methodologies employed for the dynamic analysis of power grids predominantly entail the construction of a single-line diagram (SLD) within the relevant simulation tool, followed by the execution of simulations to examine grid behavior. However, there are significant drawbacks to this approach, as it takes a long time to manually set up grids – a process that can lead to human error – especially in cases with many grid components.

To avoid such drawbacks, a comprehensive approach for executing grid analysis is needed – one that includes automatic model creation and modifications to the grid setup. With this in mind, the analytical approach developed by ABB for power grid analysis delves into transient grid dynamics. Initially, the grid layout design results are imported as inputs of the simulation. All pertinent grid information, encompassing the type and electrical characteristics of each grid component, as well as the topological structure, resides within a data exchange universal format file as a bidirectional interface designed specifically for data transfer among applications. This file format is universally supported by both the grid layout design and grid analysis modules, facilitating a seamless transfer of data without necessitating manual implementations of the grid. This aspect is particularly important when dealing with extensive grids or those undergoing structural changes due to topology modifications.

A comprehensive approach to grid analyisis with automatic model creation and modifications to the grid setup is needed.

This initial phase marks the inception of the first step of automatic grid analysis, which offers notable advantages. After importing grid attributes, a master grid aligned with an SLD is automatically generated. Clone grids, modified from the master grid by altering component parameters or » setups, are created with various dynamic asset models for transient behavior studies. In addition, the required dynamic simulation events, including load energization via switching and fault events, are generated and executed automatically.

Analyses based on simulation results are then looped back into the grid design module, enabling the generation of a revised design if needed. This iterative process highlights the integration of analytical insights and automation, fostering an efficient and informed approach to power grid stability analysis.

This automated open interface opens the door to optimizing grid operation and analysis. This process involves sweeping through component parameter ranges and different operation scenarios to verify grid stability and optimize for a target variable, such as operational costs.

Most importantly, the techno-economical calculation functionality in this proposed simulation module enables investigation into grid expansion strategies, power quality assessments, and service interruptions, all of which further incorporate optimization decision criteria for resource allocation and profitability.

Applying grid design & analysis solutions to mining industries

The mining and mineral processing industry offers an ideal application area for ABB's solution. Facing high energy demand and significant pushback due to its environmental impact, this industry urgently needs to integrate renewable energy sources such as solar and wind power to mitigate greenhouse gas emissions. Given the complex and dynamic power requirements of mining equipment, particularly the accelerated adoption of electric trucks and the associated infrastructure for stationary (battery charging), and dynamic energy transfer (vehicle propulsion), the complexity of elerictrical grid design and analysis in this area cannot be underestimated.

Yet, the absence of effective automated solutions hampers the smooth integration of renewable energy, thus slowing down the journey toward sustainability. ABB eMine™ answers this challenge by evolving its approach for mining grid design in such a way as to be able to efficiently address the related complexities and significantly contribute to more sustainable and effective industrial practices. Additionally, by utilizing the automated grid analysis method, various load, supply, and potential fault scenarios can automatically be analyzed in the power grid, thereby assisting engineers, designers, and operators in making design and operational decisions. This is accomplished by constructing detailed energy consumption and simulation models for a range of mining equipment. By merging this with historical data on renewable energy data »03, ABB can efficiently optimize power grid solutions for the integration of renewable energy.

ABB eMine™ supports the integration of renewable energy thanks to advanced mining equipment models.

Advanced analysis of this type offers optimal configuration solutions for mining grids. These solutions, informed by the renewable capability design and equipment energy consumption analyses, serve

Photo: ©rh2010/stock.adobe.com

as essential inputs for the subsequent phase of dynamic behavior verification, thus ensuring system reliably without compromising efficiency.

In addition, the grid analysis approach automatically evaluates various operating scenarios, empowering mine operators with critical insights for informed decision-making. The solution will further form the basis and provide potential for developing advanced features in other mining areas, such as optimizing the sizing of battery storage systems and electric truck fleets.

All in all, by leveraging ABB's process automation mining expertise, its industrial grid and analysis solutions enhance the offering around ABB eMine™ and support informed investment decisions, thus pointing the way to improved decarbonization and cost-efficient operations. •

Reference

[1] Mckinsey & Co. Global Energy Perspective 2023. Available: https://www.mckinsey.com/industries/oil-and-gas/our-insights/global-energy-perspective-2023 [Accessed February 17, 2024].

Originally published in ABB Review 03/2024, pp. 193–197.

Read the original article here:
https://new.abb.com/news/detail/115039/greener-grids

ABOUT ABB REVIEW

Publication
In continuous publication since 1914, ABB Review is published twice a year in both print and electronic formats. ABB Review is free of charge to those with an interest in ABB's technology and objectives.

Our charter
The objective of ABB Review is to showcase to ABB's customers the latest research results of ABB's >$1 billion annual R&D investment, innovative solutions and achievements, in an objective, permanent, trustworthy and precise way.

ABB Review

SUBSCRIBE

How to subscribe
Subscription is available in both print and digital formats. For a subscription, please contact your nearest ABB representative or subscribe online to ABB's Innovation and Technology Newsletter at innovationnews.abb.com/registration or using the subscribe QR code on the right.

Subscribers to the Innovation and Technology Newsletter will receive regular updates on innovation and technology news as well as a notification of the publication of new editions of ABB Review.

Subscribers will receive a form in which they can express a preference to receive the print edition. Print production is limited and not all requests can be honored.

Subscription

Please note that if you subscribed to either the print publication or the email notification prior to 1st August 2024, you need to re-subscribe if you wish to continue to receive ABB Review.

Reprint request

ISSN: 1013-3119

abb.com/abbreview

Layout
Publik. Agentur für Kommunikation GmbH
Mannheim/Germany

Publisher and copyright

ABB Review is published by
ABB Switzerland Ltd
Group Technology Management
Bruggerstrasse 66
5400 Baden
Switzerland
abb.review@ch.abb.com

Make sure you never miss an issue of ABB Review. Subscribe to ABB's Innovation and Technology Newsletter for regular email updates from ABB, including notifications of new editions of ABB Review.
innovationnews.abb.com/registration

Readers desiring to receive the print edition of ABB Review should sign up for the email newsletter. A registration form will be emailed to new subscribers.