Heat Pumping Technologies
MAGAZINE

Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions

Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

Contents

Technology Collaboration Programme
by iea

Heat Pumping Technologies

MAGAZINE

Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions

Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

In this Issue

Welcome to Heat Pumping Technologies Magazine – Issue 3, 2025

Welcome to this issue, where we turn our focus to one of the most critical frontiers in the energy transition: **the electrification and decarbonisation of industrial process heat**. This issue spotlights the rapidly evolving role of **high-temperature heat pumps** in industrial applications, examining how technological innovation is pushing performance boundaries and enabling solutions once considered out of reach.

Industry accounts for a substantial share of global final energy demand, with process heat representing the dominant portion of this consumption. While low- and medium-temperature applications have already seen significant uptake of heat pump technology, delivering reliable, efficient, and cost-effective heat at temperatures above 100 °C remains a central challenge. At the same time, it represents one of the greatest opportunities for reducing fossil fuel dependency and industrial greenhouse gas emissions. Against this backdrop, this issue explores how advances in refrigerants, system architectures, digital design methods, and integration concepts are accelerating the deployment of high-temperature heat pumps across diverse industrial sectors.

The **Foreword**, titled *"Industrial heat pumps – Pushing the boundaries to revolutionize industries"*, sets the stage by reflecting on the strategic importance of industrial heat pumps in meeting climate and competitiveness goals. It highlights the need to move beyond incremental improvements and embrace system-level innovation, new operating paradigms, and closer integration with industrial processes.

In the **Column**, *"Accelerating the Heat Pump Revolution: From Vision to Industrial Reality"*, the discussion shifts from ambition to implementation. The column addresses how policy frameworks, market conditions, and technological readiness must align to move high-temperature heat pumps from pilot projects to mainstream industrial solutions, emphasizing the importance of scalability, bankability, and industrial confidence.

The **Topical Articles** in this issue provide in-depth technical insights into key enablers for industrial heat pump deployment. Contributions examine refrigerant screening and advanced cycle configurations for high-temperature applications, automated and data-driven approaches to system design and integration, and innovative system concepts capable of overcoming large temperature lifts. Together, these articles illustrate how both fundamental thermodynamic research and applied engineering are converging to unlock new industrial use cases for heat pumps.

Complementing the topical content, this issue also includes a **Non-Topical Article** exploring AI-driven predictive control strategies for energy-intensive systems, highlighting how digitalisation and intelligent control can further enhance efficiency and operational performance in complex thermal systems.

In the **National Market** section, we spotlight **Sweden** with an exclusive **Heat Pump Market Report**. The Swedish market offers valuable insights into how strong policy support, industrial innovation, and system-integration expertise can create favourable conditions for the adoption of advanced heat pump technologies, including emerging industrial applications.

Taken together, **Issue 3/2025** presents a comprehensive view of how high-temperature heat pumps are transitioning from niche solutions to core components of industrial energy systems. We hope this edition informs, inspires, and supports stakeholders across industry, research, and policy as they work to accelerate the deployment of heat pumps and advance the global transition toward sustainable industrial heat.

Enjoy your reading!
Dr Metkel Yebiyo, Editor
Heat Pump Centre
The central communication activity of the Technology Collaboration
Programme on Heat Pumping Technologies (HPT TCP)

Heat Pumping Technologies
MAGAZINE
Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions
Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

Foreword

Industrial Heat Pumps – Pushing the Boundaries to Revolutionize Industries

By Benjamin Zühlsdorf, Innovation Director, PhD, Danish Technological Institute, Denmark

Industrial heat pumps have firmly established themselves as a game-changer for decarbonizing industrial process heating. By combining high energy efficiency with electrification, they are the optimal solution for a wide range of low- and medium-temperature applications, replacing fossil-fuel boilers while reducing emissions and operating costs.

Their relevance, however, extends far beyond climate policy. In an environment shaped by volatile fuel prices and geopolitical uncertainty, industrial heat pumps are increasingly seen as a strategic asset. They strengthen competitiveness through lower and more predictable energy costs and reduce dependence on imported fuels. These factors are rapidly becoming key decision drivers for companies and countries and are likely to further accelerate market uptake.

At the same time, the broader Electrotech-Revolution is accelerating deployment. The rapid build-out of renewable electricity, advances in power electronics and digitalization, and growing sector coupling all reinforce the business case for electrified process heat. Industrial heat pumps are at the heart of this development, converting increasingly green electricity and excess heat into reliable, high-value industrial heat.

Recent results from the IEA HPT Project 68 show that technologies are well underway. The technology portfolio is expanding, with more suppliers and solutions available, including systems delivering temperatures well above 200 °C. Commercialization and competition are intensifying, particularly in the range up to 140 °C, and market uptake is visible through a growing number of projects across diverse applications. Natural refrigerants are driving most new developments, while component supply chains are scaling as multiple suppliers build on proven compressors, heat exchangers, and other key components. Experience from real

projects also underscores that integration is key: the roles of integrators and process equipment suppliers are crucial to ensuring optimal design, performance, and reliability.
To fully unlock this potential, a clear shift in mindset is needed. End-users must move from simple fuel substitution to rethinking their process heat supply around efficiency, electrification, and smart integration of heat pumps into the overall process. This transformation can only succeed through close collaboration along the entire value chain – from technology providers, integrators, and process equipment suppliers to industrial end-users, researchers, utilities, and policymakers.

An ideal place to catch up on the latest developments and connect with this community is the High-Temperature Heat Pump Symposium in Copenhagen on 21 & 22 January 2026 (https://hthp-symposium.org/). I hope this issue of Heat Pumping Technologies Magazine inspires you to explore the opportunities of industrial heat pumps, and I look forward to continuing the discussion with many of you at the HTHP Symposium.

Name	Benjamin Zühlsdorf
Title	Innovation Director, PhD
Affiliation	DTI, Danish Technological Institute, Denmark
E-mail address	bez@dti.dk

Heat Pumping Technologies
MAGAZINE
Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions
Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

Column

Accelerating the Heat Pump Revolution: From Vision to Industrial Reality

By Rainer M. Jakobs, Alternate Delegate of Germany, IEA HPT TCP, PhD, Information Centre on Heat Pumps and Refrigeration IZW e.V

What is the current status of Heat Pumps (HPs)? Is the HP technology at a limit? Which temperatures and what capacities could HPs achieve? What technological progress has been made in recent years? Are major steps forward in increasing efficiency still possible?

Which findings did the European Heat Pump Summit provide us with on such topics? The Summit took place on October 28-29, 2025, and returned for its ninth edition at a pivotal moment in the global energy transition. The HP industry stands at an unprecedented inflection point where political commitment, technological innovation, and market momentum converge to reshape how we heat our buildings and power our industries.

The European HP market has undergone a remarkable transformation. Building on 2022's record-breaking 38 % sales increase to 3 million units, the market continues on a growth course. With over 25 million HPs (as of the end of 2024) installed across Europe, we're witnessing a fundamental shift in heating and cooling approaches. The European HP market is now showing a slight yet uneven recovery this year. Sales across 13 European countries increased by an average of 9% in the first six months of 2025, compared with the same period in 2024. Despite this slight rebound, overall market confidence is still lower than in the record sales year of 2022. Progress is constrained by policy uncertainty and reduced incentives, according to national experts. (Source EHPA)

While residential applications have driven initial growth, industrial HPs represent the next critical frontier. With process heat accounting for over two-thirds of industrial energy consumption and still predominantly using fossil fuels, the transformation potential is immense. Simultaneously, 20-50 % of industrial input energy is lost as waste heat, presenting an untapped circular energy resource.

The Summit featured groundbreaking presentations on industrial applications that push technical and economic boundaries: transforming waste heat into high-pressure steam at 215 °C, or delivering solutions with 50 MWth, e.g., with the flagship BASF Ludwigshafen installation. It demonstrates real-world viability at scale.

The programme reflected increasing sophistication in HP applications. Denmark's Esbjerg Mega HP showcases successful district heating integration, while presentations on chemical, petrochemical, food and beverage, and pulp and paper applications demonstrated sector-specific solutions. These real-world case studies offered practical insights for navigating complex industrial transformations.

The convergence of heat pump technology with digitalization and IoT capabilities opens new optimization frontiers. The comprehensive hardware-in-the-loop testing facility accelerates HP innovation from concept to deployment. Using versatile test benches with multiple cycle configurations, it optimizes refrigerant selection and system design for safety and performance across all operational ranges. The future of energy management with model predictive control (MPC) solutions will be thoroughly tested in holistic building energy systems. (Source RWTH Aachen)

The ongoing evolution of refrigerant regulations remains critical, demanding industry-wide collaboration and innovation. F-Gas Regulation amendments and PFAS discussions require thinking beyond traditional metrics. While Global Warming Potential remains important, comprehensive assessments using Total Equivalent Warming Impact (TEWI) and Life Cycle Climate Performance (LCCP) provide a more accurate picture of environmental impact.

The contents of the European Heat Pump Summit reflected the ongoing evolution of the HP sector and its growing relevance in various fields of application. Alongside market analyses and policy frameworks, the focus was on leading-edge topics, including digital simulation-based modelling methods, HP solutions for industry, and district heating at the megawatt scale. It is remarkable that even established companies in the engine manufacturing and machine construction sectors are increasingly engaging with HP technology, thereby contributing to the ongoing development of the sector.

Name	Dr.-Ing. Rainer M. Jakobs,
Title	Alternate Delegate of Germany, IEA HPT TCP
Affiliation	Information Centre on Heat Pumps and Refrigeration IZW e.V.
E-mail address	Dr.Rainer.Jakobs@t-online.de

Technology Collaboration Programme
by iea

Heat Pumping Technologies
MAGAZINE
Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions
Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

Topical Article

Fluid Screening for High-Temperature Heat Pumps

DOI: 10.23697/ghhy-qh89

Niclas Kalter, Cedric Kötting, Anna Halle, Matthias Mersch, Christian Vering, Dirk Müller

The paper industry is among the most energy-intensive sectors. High-temperature heat pumps are becoming increasingly viable at the required temperatures of about 130°C, offering the potential to efficiently decarbonise the process. However, the efficiency of heat pumps depends highly on both the cycle configuration as well as the refrigerant choice. Therefore, we present a screening of promising low-GWP refrigerants for subcritical and transcritical heat pump cycles, assessing both standard and enhanced vapour-injection cycle configurations. Our results show that most considered refrigerants outperform the commonly used conventional refrigerant R245fa in both Coefficient of Performance (COP) and volumetric heating capacity (VHC).

Introduction

Industrial manufacturing continues to be associated with significant CO_2 emissions, with the steel, chemical, and paper industries being among the most energy-intensive sectors [1]. Within the paper industry, many process steps require heat at temperatures attainable with high-temperature heat pumps (HTHPs), offering an opportunity to reduce fossil fuel dependence and emissions.

The paper production process consists of five main stages [2]:

1. Dilution section: The diluted pulp is evenly distributed onto the wire to form a uniform sheet.
2. Wire section: Water drains away while the fibers interlock, forming a continuous paper web.
3. Press section: The paper web is mechanically pressed to remove water and increase density.
4. Dryer section: Heated cylinders evaporate the remaining moisture from the paper.
5. Calendering section: The surface of the paper is smoothed before being wound into large rolls.

The drying and calendering stages are the most energy-intensive steps [3], accounting for approximately 60 - 70 % of the total energy demand of the paper manufacturing process [3]. During drying, the paper web passes over 50 to 100 steam-heated cylinders, depending on machine size. The specific energy consumption of this process is approximately 1245 kWh per ton of paper [4]. In the subsequent calendering process, the paper passes over rolls heated to 80-400 °C, depending on product requirements and paper type. Hot water or pressurized steam at temperatures of up to 120 °C is commonly used, with an energy demand of around 220 kWh per ton of paper [4].

At present, both processes are predominantly supplied by gas-fired boilers. High-temperature heat pumps represent a promising alternative for these applications. They can deliver heat at the required temperature levels while significantly reducing fossil fuel use and carbon emissions. Integrating such systems into paper manufacturing processes could therefore play an important role in the industry's transition towards climate-neutrality.

For the analysis presented here, we set the target heat sink temperature (T_{Sink}) for our HTHP to 130 °C, aiming to replace fossil-based steam generation at pressures below 5 bar, which is prevalent in European paper mills [5]. This selection is further motivated by the rapid advancements and existing market readiness of HTHPs, defined as systems providing heat sink temperatures above 100 °C [6], [7]:

- Commercial Availability: Several commercial products already operate in this space. Benchmarks demonstrate that specific HTHP units equipped with screw compressors are able to deliver temperatures up to 130 °C [6].
- Decarbonization Potential: By providing heat at 130 °C, the HTHP can directly substitute heat supplied by gas boilers for drying processes, which accounts for approximately 27 % of the total industrial process heat demand in the 100 °C to 200 °C range [7].

To operate HTHPs efficiently at this temperature, a sufficiently high-grade heat source (T_{Source}) is required. We propose utilising on-site low-temperature waste heat or District Heating (DH) at 70 °C as the heat source in our scenario. Higher source temperatures result in a lower temperature lift and thus higher COPs. In our case, the required temperature lift is relatively moderate at 60 K. This falls well within the high-efficiency operating window of modern HTHPs, whose COPs typically range from 2.4 to 5.8 for lifts between 95 K and 40 K [6]. However, reaching high efficiencies is key to the economic viability of industrial heat pumps. While achievable, it necessitates a sophisticated cycle design and refrigerant selection.

Cycle Configurations

The standard, single-stage vapor compression cycle is primarily used for low temperature lifts (around 30 K [7]). At higher temperature lifts, critical challenges arise:

1. Efficiency Degradation: The COP decreases significantly as the temperature lift increases [8]. However, multi-stage cycles can increase the overall efficiency [9], [10].
2. High Discharge Temperature: This yields a high lubricant temperature [8]. High temperatures threaten the thermal stability of the refrigerant-lubricant mixture [10]. This potentially leads to lubricant degradation, loss of lubrication, and material issues within the compressor [8], limiting its operating range and durability [11].

Although the maximum temperature lift for a standard cycle may reach up to 80 K [7], operation at a 60 K lift, as required for our application, places the system under substantial strain. Modifications to the simple heat pump flow sheet help overcome the performance and reliability limitations [10]. These modifications aim to reduce the compression work and mitigate the issue of high discharge temperatures, thereby extending the operating limits [12].

One of the most effective and widely adopted concepts for high-temperature lift applications is the quasi-two-stage cycle with vapor injection [10]. This technique fundamentally changes the compression process by injecting refrigerant vapor at an intermediate pressure to intercool the partially compressed vapour (see Figure 1) [11].

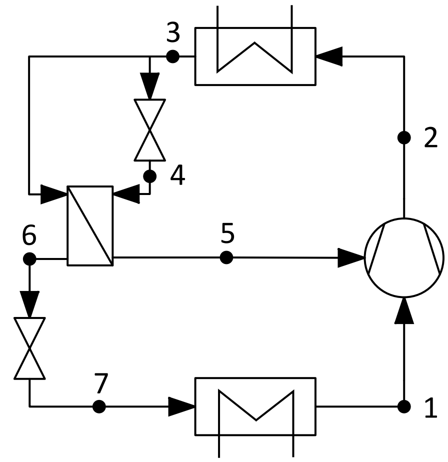

Figure 1: Flowsheet schematic of the vapor injection cycle with an economizer to control the condition (temperature and pressure) of the injected vapor.

We therefore compare the performance of the simple, single-stage standard cycle, which serves as a baseline, with that of the vapour injection cycle. This comparison allows us to quantitatively demonstrate how modifications to the flow sheet address the core challenges of achieving high COPs and VHCs depending on the selected refrigerant.

Refrigerant Choice Matters for Efficiency and Volumetric Heating Capacity

The fluid screening is performed on refrigerants preselected from the REFPROP database to conform to the EU regulation on fluorinated greenhouse gases [13], which requires the fluids to have an Ozone depletion potential (ODP) of 0 and a Global Warming Potential (GWP) of less than 150 [14]. Flammability and toxicity are not restricted during the refrigerant selection, as we consider systems for industrial applications that will be tightly monitored and operated by professional staff.

Two screening processes are implemented: one for subcritical and one for transcritical processes. To avoid air leaking into the system, the minimum allowable evaporation pressure is set to 1.1 bar. For subcritical processes, the condensation pressure is limited to 90 % of the critical pressure, while for transcritical processes, the critical temperature must lie between the source and sink temperatures.

When applying these criteria, 20 suitable refrigerants remain. Among these, 13 refrigerants allow a subcritical process while 7 refrigerants result in transcritical operation. Figure 2 compares them in terms of their *COP* and *VHC* for the standard cycle and the vapor injection cycle.

The COP is a measure of cycle efficiency and, thus, of electricity usage and operating costs. The VHC*,* on the other hand, quantifies the useful heat output per unit volume. Thus, a higher VHC implies smaller components and lower investment costs [15]. For high-temperature heat pumps, operating costs typically dominate the levelised cost of heat [16], suggesting a higher importance of the COP for optimal refrigerant selection.

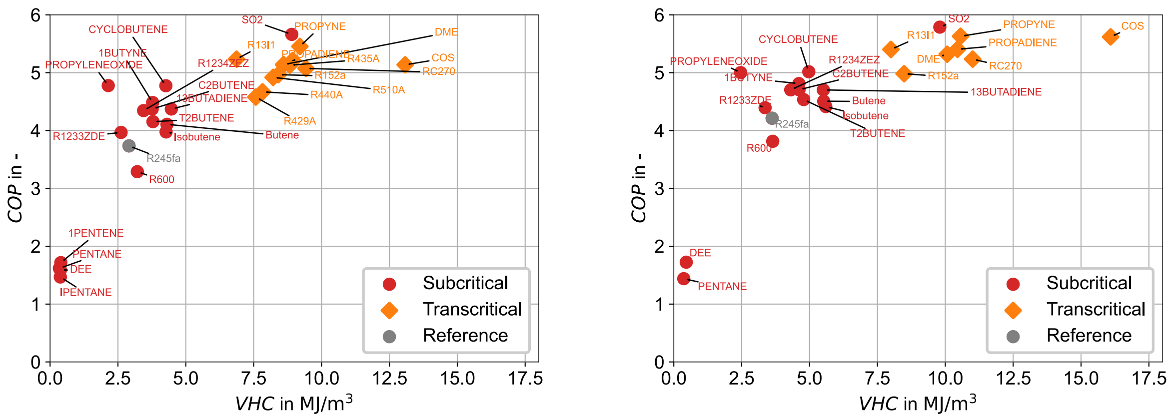

Figure 2: Comparison of volumetric heating capacity (VHC) and coefficient of performance (COP) for simple (left) and vapour injection (right) cycles.

In both simulations, three clusters of refrigerants emerge:

1. One cluster in the lower left-hand corner with *COP*s near 1.5 and $VHCs$ lower than 1 MJ/m^3, which are not suitable for our application due to their low efficiency.
2. A second cluster predominantly containing refrigerants in subcritical operation at *COP*s of 3.5 to 5 and *VHC*s of 2.4 MJ/m^3 to 5.5 MJ/m^3. Notably, natural refrigerants with four carbon atoms such as butane, butene and butyne offer a similar performance to HFOs like R1234ze(Z) in terms of both efficiency and volumetric heating capacity.
3. A third cluster with *COP*s over 5. These refrigerants require transcritical operation under the investigated conditions. Transcritical cycles feature a gas cooler instead of a condenser and pose larger difficulties in designing control architectures, as pressure and temperature are no longer coupled by the phase change within the two-phase region and must thus be controlled independently from one another.

Almost all of the investigated refrigerants in the second and third cluster outperform the reference fluid, R245fa, a HFC with a GWP of 858 [18] commonly used in high-temperature heat pumps [6].

The majority of refrigerants in this third cluster contain a double (SO_2) or triple bond (Propyne) or a carbon ring (RC270). These compounds have a high energy content but may be chemically unstable [18]. This is especially of concern at higher temperatures. Sulfur dioxide (SO_2) is an outlier, as it performs well while remaining subcritical under the investigated operating conditions. However, it is corrosive to copper and therefore necessitates more expensive stainless steel piping.

Advanced Flowsheets Allow even Further Improvement

More complex cycle configurations can potentially offer efficiency improvements [18]. Figure 3 compares the results for the vapour injection cycle with the baseline standard cycle results.

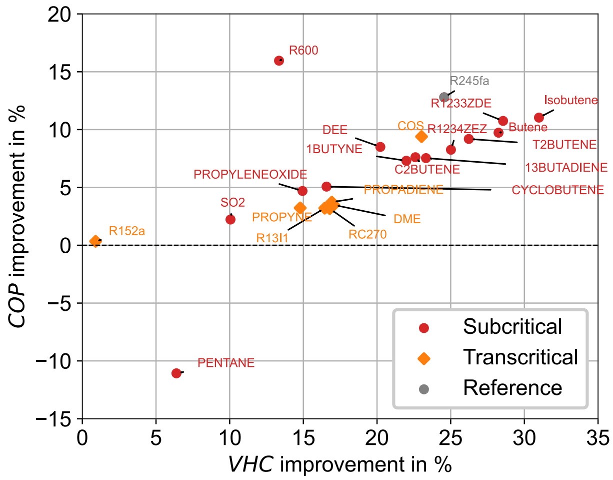

Figure 3: Relative improvement in COP and VHC when switching from a simple refrigeration cycle to a vapor injection cycle with an economizer.

One of the largest *VHC* improvements is achieved for R1233zd(E), which was also identified as a promising refrigerant by Frate et al. [15], who screened refrigerants considering both COP and VHC under operating conditions similar to ours. The largest *COP* improvement is observed for R600, which is consistent with the findings of Höges et al. [14].

The results clearly indicate that vapour injection refrigerant cycles are superior in terms of efficiency and energy density for high-temperature heat pumps. The advanced cycle configuration enables improvements of over 10% in COP and over 25% in *VHC*. The large improvement in VHC is due to the injection of high-density fluid during the compression, limiting the maximum compressor temperature and increasing mass throughput.

Limitations

This study focuses on pure fluids. An alternative is the use of zeotropic mixtures to achieve temperature glide matching. The mixture composition is chosen so that the slope of the non-isothermal phase change matches that of the secondary fluid. This minimises the temperature difference in the heat exchanger and thus exergy losses. However, as Roskosch et al. [19] point out, the utilised compressor model must accurately reflect the fluid-dependent efficiencies of the zeotropic mixture to realize the theoretical *COP* improvements in practice. Widmaier et al. [20] have used Roskosch's compressor model to explore the possibilities of using refrigerant mixtures to increase standardisation in the HTHP sector.

The methodology outlined in this study may be extended in similar ways in future work. Optimisation algorithms similar to those developed by Mersch et al. [21] may be used to include cost correlations rather than screening purely based on *COP* and *VHC*. Using group contribution theories such as PC-SAFT, the optimization can even go down to the molecular level in order to identify entirely novel refrigerants [22].

Conclusions

In this study, we investigated the decarbonisation of the drying and calendering processes in a paper mill using a high-temperature heat pump. We screened the REFPROP database for refrigerants that comply with current EU regulations. The analysis considers subcritical and transcritical processes and compares the results of a standard refrigerant cycle with those of a vapor-injection cycle using an economizer. The key findings are:

- The choice of refrigerant has a large impact on both efficiency and energy density, thus operating and investment costs.
- Low-GWP natural and synthetic refrigerants can outperform conventional refrigerants such as R245fa even using a standard cycle.
- Using a vapour-injection cycle improves the attainable *COP* by up to 15 % and the *VHC* by up to 32 %.
- The work can serve as a basis for future optimisation studies on optimal combinations of refrigerant and flow sheet with respect to economic and ecological aspects.

Author contact information

Name	Niclas Kalter
Affiliation	Institute for Energy Efficient Buildings and Indoor Climate, RWTH Aachen University
Postal address	Mathieustrasse 10, 52074 Aachen
E-mail address	niclas.kalter@eonerc.rwth-aachen.de
Phone number	+49 241 80 49 793

References

[1] A. von Rebnitz, M. Hantke, and K. Brabender, “DIE PAPIERINDUSTRIE - Leistungsbericht PAPIER 2024,” DIE PAPIERINDUSTRIE e.V., Berlin, Leistungsbericht, Apr. 2024.

[2] F. Veitengruber, “Energetische Hybridisierung von energieintensiven Prozessen – Leitfaden für die Erschließung der Potenziale im Unternehmen.” Accessed: Oct. 30,

2025. [Online]. Available: https://www.ffe.de/wp-content/uploads/2020/02/Energetische-Hybridisierung.pdf

[3] J. Laurijssen, A. Faaij, and E. Worrell, “Benchmarking energy use in the paper industry: a benchmarking study on process unit level,” *Energy Effic.*, vol. 6, no. 1, pp. 49–63, Feb. 2013, doi: 10.1007/s12053-012-9163-9.

[4] M. Bons, “Energiewende in der Industrie,” Executive Summary to the Bundesministerium für Wirtschaft und Energie Vorhaben: I C 4 – 80 14 38/42; Projekt-Nr. 42/17. Accessed: Oct. 30, 2025. [Online]. Available: https://www.bundeswirtschaftsministerium.de/Redaktion/DE/Downloads/E/energiewende-in-der-industrie-ap2b-executive-summary.pdf?__blob=publicationFile&v=4

[5] T. Nowak and A. Westenbroek, “Cepi-x-EHPA.” Feb. 2023. Accessed: Oct. 30, 2025. [Online]. Available: https://www.cepi.org/wp-content/uploads/2023/02/Cepi-x-EHPA.pdf

[6] C. Arpagaus, F. Bless, M. Uhlmann, J. Schiffmann, and S. S. Bertsch, “High temperature heat pumps: Market overview, state of the art, research status, refrigerants, and application potentials,” *Energy*, vol. 152, pp. 985–1010, Jun. 2018, doi: 10.1016/j.energy.2018.03.166.

[7] K.-M. Adamson *et al.*, “High-temperature and transcritical heat pump cycles and advancements: A review,” *Renew. Sustain. Energy Rev.*, vol. 167, p. 112798, Oct. 2022, doi: 10.1016/j.rser.2022.112798.

[8] A. Redón, E. Navarro-Peris, M. Pitarch, J. Gonzálvez-Macia, and J. M. Corberán, “Analysis and optimization of subcritical two-stage vapor injection heat pump systems,” *Appl. Energy*, vol. 124, pp. 231–240, Jul. 2014, doi: 10/f54xsz.

[9] M. E. Konrad and B. D. MacDonald, “Cold climate air source heat pumps: Industry progress and thermodynamic analysis of market-available residential units,” *Renew. Sustain. Energy Rev.*, vol. 188, p. 113739, Dec. 2023, doi: 10.1016/j.rser.2023.113739.

[10] M. Zehnder, “Efficient air-water heat pumps for high temperature lift residential heating, including oil migration aspects,” Lausanne, EPFL, 2004. doi: 10.5075/EPFL-THESIS-2998.

[11] F. M. Tello Oquendo, “Study of scroll compressors with vapor-injection for heat pumps operating in cold climates or in high-temperature water heating applications,” Universitat Politècnica de València, Valencia (Spain), 2019. doi: 10.4995/Thesis/10251/120473.

[12] H. Ceylan, “Review on the two-stage vapor injection heat pump with a flash tank,” *Eng. Sci. Technol. Int. J.*, vol. 48, p. 101583, Dec. 2023, doi: 10.1016/j.jestch.2023.101583.

[13] European Union, *Regulation (EU) 2024/573 of the European Parliament and of the Council of 7 February 2024 on fluorinated greenhouse gases*. 2024, p. 67. Accessed: Oct. 30, 2025. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202400573

[14] C. Höges, D. Roskosch, J. Brach, C. Vering, V. Venzik, and D. Müller, "Investigation of the Interactions between Refrigerant, Flowsheet, and Compressor in Residential Heat Pumps," *Energy Technol.*, vol. 11, no. 2, Feb. 2023, doi: 10.1002/ente.202201295.

[15] G. F. Frate, L. Ferrari, and U. Desideri, "Analysis of suitability ranges of high temperature heat pump working fluids," *Appl. Therm. Eng.*, vol. 150, pp. 628–640, Mar. 2019, doi: 10.1016/j.applthermaleng.2019.01.034.

[16] F. Schlosser, S. Zysk, T. G. Walmsley, L. Kong, B. Zühlsdorf, and H. Meschede, "Break-even of high-temperature heat pump integration for milk spray drying," *Energy Convers. Manag.*, vol. 291, p. 117304, Sep. 2023, doi: 10.1016/j.enconman.2023.117304.

[17] A. Okasha, N. Müller, and K. Deb, "Bi-objective optimization of transcritical CO2 heat pump systems," *Energy*, vol. 247, p. 123469, May 2022, doi: 10.1016/j.energy.2022.123469.

[18] C. Höges, V. Venzik, C. Vering, and D. Müller, "Decarbonizing Steam Generation with High Temperature Heat Pumps: Refrigerant Selection and Flowsheet Evaluation," in *Proceedings of the 14th IEA Heat Pump Conference 2023*, Chicago, IL, USA, May 2023. [Online]. Available: https://www.researchgate.net/publication/371314235_Decarbonizing_Steam_Generation_with_High_Temperature_Heat_Pumps_Refrigerant_Selection_and_Flowsheet_Evaluation

[19] D. Roskosch, V. Venzik, J. Schilling, A. Bardow, and B. Atakan, "Beyond Temperature Glide: The Compressor is Key to Realizing Benefits of Zeotropic Mixtures in Heat Pumps," *Energy Technol.*, vol. 9, no. 4, p. 2000955, Apr. 2021, doi: 10.1002/ente.202000955.

[20] P. Widmaier, L. P. M. Brendel, S. S. Bertsch, A. Bardow, and D. Roskosch, "One Mixture to Rule Them All: Enhancing Efficiency and Standardization of Industrial High-Temperature Heat Pumps," *ACS Eng. Au*, vol. 5, no. 4, pp. 359–369, Aug. 2025, doi: 10.1021/acsengineeringau.4c00060.

[21] M. Mersch, D. Tillmanns, P. Sapin, J. Schilling, A. Bardow, and C. N. Markides, "Integrated thermo-economic organic Rankine cycle and working fluid design – On the accuracy of molecular-based computer-aided methodologies," *Comput. Chem. Eng.*, vol. 199, p. 109151, Aug. 2025, doi: 10.1016/j.compchemeng.2025.109151.

[22] J. Schilling, D. Tillmanns, M. Lampe, M. Hopp, J. Gross, and A. Bardow, "From molecules to dollars: integrating molecular design into thermo-economic process design using consistent thermodynamic modeling," *Mol. Syst. Des. Eng.*, vol. 2, no. 3, pp. 301–320, 2017, doi: 10.1039/C7ME00026J.

Heat Pumping Technologies
MAGAZINE
Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions
Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

Topical Article

From Expert Decisions to Automation: Designing and Integrating Industrial Heat Pumps

DOI: 10.23697/bn59-5q86

Mostafa Babaei, Institute for Environmental Sciences, University of Geneva, Switzerland

High-temperature heat pumps are becoming a cornerstone of industrial decarbonization, offering a pathway to replace fossil-fuel-based heating with clean, efficient alternatives. Yet determining how and where to integrate them effectively remains a major challenge for engineers and planners. Traditionally, system design has relied on expert judgment and manual iteration, an approach that is too slow to keep pace with industrial transformation and the diversity of process conditions. This article examines how combining physics-based reasoning with data-driven algorithms can automate design and integration, enabling faster and more efficient electrification of industrial heat.

Introduction

Industry accounts for more than one-third of global final energy use, and nearly two-thirds of that energy is consumed to produce heat. Despite steady efficiency improvements, about 80 percent of industrial heat worldwide is still generated by burning fossil fuels. Replacing this heat with renewable or electricity-based alternatives is therefore one of the most direct and impactful steps toward reducing industrial emissions.

High-temperature heat pumps (HTHPs) offer a practical solution. They can recover low-grade waste heat from cooling streams or exhaust gases and upgrade it to provide process heat at the required temperature levels. For many applications, particularly those operating below approximately 200 °C, HTHPs can already achieve high efficiency and competitive operating costs. Their adoption, however, remains limited, not because of technology immaturity but because each industrial site presents a unique combination of heat sources, sinks, and process constraints.

Designing an effective heat-pump system requires selecting suitable temperature lifts, identifying integration points, and determining the number and size of units. Today, these decisions still rely heavily on expert judgment and manual assessment of process heat profiles. While this hands-on approach can yield robust solutions for single facilities, it becomes time-consuming and inconsistent when scaling up to many sites or exploring multiple design scenarios.

To enable large-scale deployment, heat-pump integration must evolve from manual, experience-driven design toward systematic and automated workflows. Such automation can use thermodynamic reasoning to identify technically sound operating conditions while algorithmic methods handle the complex task of configuration and sizing. By combining engineering knowledge with computational intelligence, designers can rapidly evaluate a wide range of options within a single framework. This transition does not replace human expertise; it amplifies it, allowing engineers to focus on strategic decisions while algorithms perform the heavy analytical work.

From Expert Design to Automated Integration

The integration of industrial heat pumps is a complex engineering task that requires balancing thermodynamic feasibility, process compatibility, and economic performance. Traditionally, this process has relied on expert-driven studies in which engineers manually analyze process data to identify suitable heat sources, sinks, and temperature levels for heat-pump operation. Such assessments typically make use of Grand Composite Curves (GCCs) to visualize process heating and cooling demands and to identify opportunities for heat recovery. While these methods provide valuable insights, they are highly dependent on engineering experience and can be time-consuming when applied across multiple sites or scenarios.

As industrial decarbonization progresses, there is a growing need for design approaches that can systematically evaluate a large number of integration opportunities. Applying conventional methods at scale would require considerable effort and may still overlook viable options due to the diversity of industrial processes and boundary conditions. Automated frameworks can help address this challenge by combining process data with thermodynamic reasoning and algorithmic tools to identify feasible and cost-effective heat-pump solutions. The key idea is to move from a sequential, human-led workflow to a structured system that can automatically evaluate many potential matches. In this context, each possible combination of waste-heat sources and process-heat demands is tested using simplified thermodynamic models. These models estimate whether the required temperature lift and expected performance fall within practical operating ranges. Configurations that are technically unrealistic or inefficient can be excluded early in the process, leaving only those that warrant detailed investigation.

This thermodynamic screening serves as a transparent filter, reducing the design space and ensuring that subsequent optimization focuses on meaningful cases. Figure 1 illustrates an example of a Grand Composite Curve from a paper mill, where the regions below and above the pinch point represent potential heat-source and heat-sink zones for a high-temperature heat pump. Such profiles form the foundation for automated screening by clearly outlining feasible temperature intervals and helping identify the most promising integration levels.

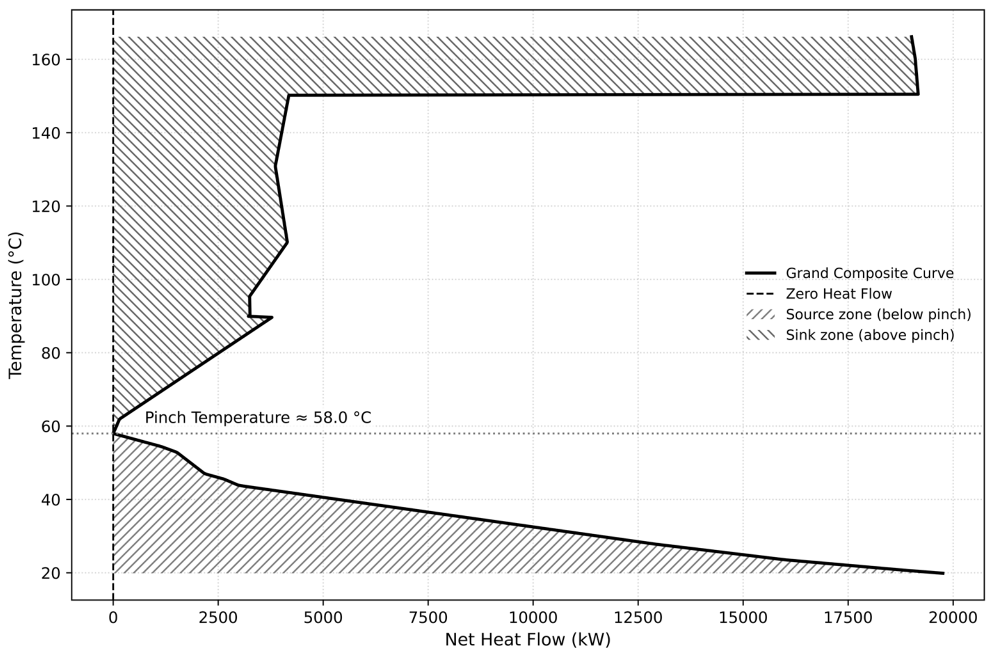

Figure 1. Grand Composite Curve showing potential heat-source and heat-sink regions for heat-pump integration.

Thermodynamic Feasibility and Process Matching

A central element of automated design is the assessment of thermodynamic feasibility. Instead of testing every possible configuration in detail, simplified models use temperature data from process streams to estimate whether a given source-sink pair is compatible with heat-pump operation. Using the concept of an ideal Carnot cycle as a reference, these models approximate the maximum theoretical performance and identify the range of temperature lifts that remain realistic within a given efficiency threshold.

By mapping this information across all process streams, the framework generates a systematic picture of which heat sources and sinks can potentially be connected. Pairs that would require excessive temperature lift or lead to poor performance can be filtered out automatically. The result is a reduced set of technically sound options that form the basis for further system-level evaluation. This approach accelerates the assessment and improves consistency, enabling fair comparison of multiple sites or process scenarios using a common set of thermodynamic criteria.

Superstructure and Integration Framework

Once the feasible temperature levels and source-sink combinations are known, the next challenge is to determine how one or several heat pumps can best be configured to meet the process requirements. Industrial facilities often include multiple heating and cooling streams with different temperature levels and loads, making manual evaluation of all possible arrangements difficult.

A superstructure-based framework addresses this challenge by representing all potential configurations within a single model. Each configuration describes a specific way of linking evaporators and condensers, possibly with intermediate heat exchangers or different working fluids. Optimization algorithms explore these configurations to find the combination that minimizes total annual cost while satisfying thermodynamic and operational constraints.

The concept is illustrated in Figure 2, in which multiple heat sources and sinks are evaluated under alternative heat-pump configurations. The algorithm then selects the most cost-effective configuration that meets process requirements. This approach enables consistent evaluation of both technical and economic performance, providing a unified view of feasibility, efficiency, and cost within the same analytical framework. By embedding thermodynamic screening and cost evaluation into a single structure, the superstructure framework bridges the gap between early-stage feasibility and detailed design. It enables engineers to rapidly and transparently explore multiple integration options through a reproducible process that complements, rather than replaces, engineering expertise.

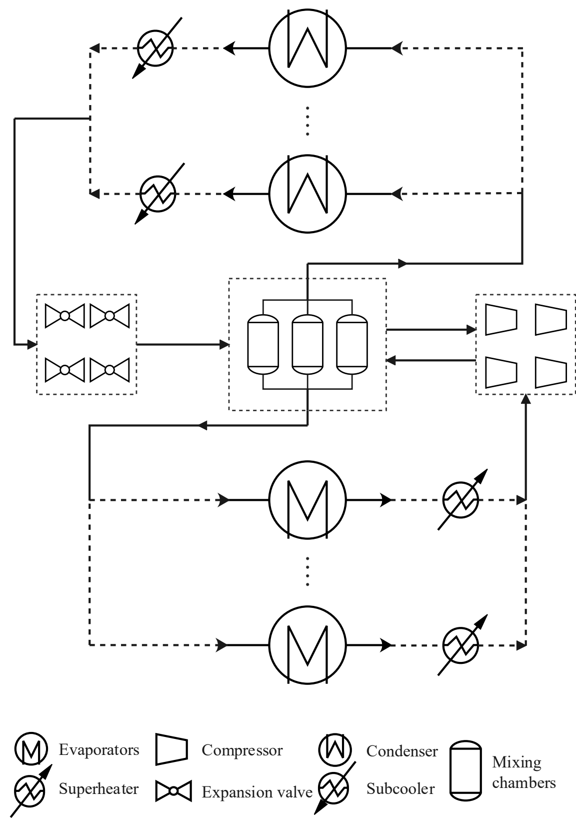

Figure 2. Heat pump superstructure linking multiple sources/sinks through alternative heat-pump configurations.

Conclusions

Automating the design and integration of industrial heat pumps can significantly accelerate their deployment and improve decision-making consistency. By combining thermodynamic reasoning with algorithmic exploration, feasible temperature levels and optimal system configurations can be identified efficiently across diverse industrial processes. The superstructure framework provides a unified platform to compare alternatives, balancing technical feasibility and economic performance within a single tool.

Beyond individual projects, such approaches can support broader energy-transition strategies by identifying the most promising opportunities for electrifying industrial heat. As computational methods become more accessible, automation will increasingly complement engineering expertise, enabling faster, more transparent, and data-driven design of high-temperature heat-pump systems for industry.

Author contact information

Name	Mostafa Babaei
Title	PhD Researcher
Affiliation	Group for Energy Efficiency, Institute for Environmental Sciences, University of Geneva
Postal address	University of Geneva, Institute for Environmental Sciences Boulevard Carl-Vogt 66, 1205 Geneva, Switzerland
E-mail address	Mostafa.babaei@unige.ch
Phone number	+41 76 645 41 30

References

[1] International Energy Agency (IEA). Energy System – Industry. Available at: https://www.iea.org/energy-system/industry (accessed October 2025).

[2] International Energy Agency (IEA). Energy System – Buildings. Available at: https://www.iea.org/energy-system/buildings (accessed October 2025).

[3] European Heat Pump Association (EHPA). Market and Policy Information on Heat Pumps. Available at: https://www.ehpa.org (accessed October 2025).

[4] Babaei, S., et al. (2025). Title of Your Paper. Journal of Cleaner Production. https://doi.org/10.1016/j.jclepro.2025.146109

[5] Wallerand, A. S., Kermani, M., Kantor, I., & Maréchal, F. (2018). Optimal heat pump integration in industrial processes. Applied Energy, 219, 68-92. https://doi.org/10.1016/j.apenergy.2018.02.114

Heat Pumping Technologies
MAGAZINE
Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions
Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

Topical Article

Large Temperature Lifts: A Major Challenge? The Future of Industrial Process Heat

DOI: 10.23697/8d20-xk89

Christoph Rau, SPH Sustainable Process Heat GmbH, Germany

The industry faces the challenge of providing process heat as efficiently and sustainably as possible. Industrial heat pumps that are able to utilize typical waste heat sources offer significant potential. However, the high temperatures required and the associated temperature lifts often pose a challenge for conventional heat pump technologies. Innovative high-temperature heat pumps extend the boundaries of conventional systems and enable efficient utilization of waste heat.

Introduction

Every year, more than 400 terawatt-hours (TWh) of energy are used to generate process heat, accounting for about two-thirds of Germany's annual industrial energy demand [1]. Decarbonizing these processes is essential for achieving European climate goals. However, the high temperatures required in many processes and the necessary temperature lifts of more than 80 Kelvin (K) present a technical challenge for conventional industrial heat pumps. Specialized systems are needed to combine future-proofing, efficiency, and economic viability.

Electric process heat generation is a promising approach: According to a study by the Niederrhein University of Applied Sciences, the German industry sector could save up to 21 billion Euros (23 billion USD) annually in energy costs for process heat generation [2]. The largest savings potential arises from electrification, heat recovery, and waste heat utilization—12.8 billion Euros (14 billion USD) could be saved through immediately implementable measures alone. Overall, this represents a potential reduction of up to 33% in the industry's final energy demand, equivalent to approximately 226 TWh per year, with a significant portion amortized within 3 years.

At the EU level, the annual energy demand for process heat is estimated at 2,950 TWh, with a share of 25% in the 100-200 °C temperature range [3][4]. This makes high-temperature heat pumps such as the ThermBooster™, developed by SPH Sustainable Process Heat GmbH, located in Overath, Germany, a key technology for the European industrial sector. A McKinsey analysis projects annual market growth exceeding 15% for this sector by 2030, with global investments reaching approximately 12 billion US dollars [5].

Large Temperature Lifts: A Challenge for Heat Pump Systems?

Despite the potential of electrifying process heat generation, technical challenges remain: industrial waste heat sources typically range from 40 to 80 °C. For temperature requirements between 100 and 200 °C, this necessitates a temperature increase exceeding 100 K in some cases.

Industrial heat pumps face significant challenges when the temperature lift exceeds 80 K. The reason being that, for maximum efficiency, the system must transfer as much heat energy as possible from the heat source to the heat sink. This typically requires a high evaporation pressure and high suction density. At low waste-heat temperatures, which result in low evaporation pressures, high pressure ratios are required to achieve significant temperature lifts.

For example, in a piston-based heat pump compressor, residual gas remains in the cylinder after each compression cycle (dead volume). This residual gas expands during the subsequent intake phase (re-expansion), thereby reducing the fresh-gas intake. At a pressure ratio of 9.5 between the high- and low-pressure sides of the heat pump cycle and a dead volume of only 3% of the total volume, up to 28% of the actual suction volume is blocked by residual gases. The larger the dead volumes and the higher the pressure ratio, the lower the overall system efficiency.

A low evaporation pressure of only 2 bar and a suction density of around 10 kg/m^3 result in low specific thermal performance, as the system can evaporate less refrigerant and thus transfer less heat per cycle. This low performance requires high displacement capacity, which can only be achieved with a larger, more powerful compressor or multiple compressors. However, this significantly increases the investment costs for such systems.

Refrigerants with specific performance characteristics, unfavourable pressure ratios, and limited volumetric efficiency—these obstacles ultimately lead to a fundamental problem: traditional single-stage refrigeration cycles often struggle to meet the requirements of industrial heat generation and efficient waste heat utilisation due to their design limitations.

Innovative Heat Pump Technologies Enable Temperature Lifts Beyond 100 Kelvin

Solutions such as the modular high-temperature heat pump ThermBooster™ offer various approaches to overcoming these challenges. One possible solution is two-stage cascading: The ThermBooster™ can use two separate refrigerant cycles in series to efficiently, reliably, and economically handle temperature lifts beyond 80 K. Unlike other compressor types, the specially developed piston compressor can achieve the required pressure ratio regardless of its speed, making it ideal for applications with high temperature lifts.

In the lower cycle, refrigerants such as R-515 B, propane, or R-1234ze can be used for temperature lift at lower temperatures, whereas the upper cycle can utilize R-1233zd, butane, or isobutane. For even higher temperatures, R1336mzz(Z) or pentane can be used as refrigerants. Each of the two stages provides a temperature lift exceeding 50 K, yielding a total temperature lift exceeding 100 K.

Combining the system with a steam compressor (mechanical vapour recompression) offers a second possible solution: By integrating a steam compressor, the ThermBooster™ can economically bridge even greater temperature differentials.

Achieving Nearly Any Temperature Lift with Cascading

Sterilization and drying processes, two key application areas for process heat in the food and chemical industries, benefit significantly from the efficiency of two-stage systems such as the ThermBooster™. A project example:

- Heat source: 50/36 °C (evaporation at 33 °C)
- Steam: 3.5 bar(a), 139 °C (condensation at 144 °C)
- Capacity: 646 kW

Figure 1 illustrates a cascaded high-temperature heat pump system that uses two refrigerants, arranged in two stages, to achieve a substantial overall temperature lift. In the first stage, the refrigerant evaporates at approximately 33 °C and condenses at approximately 85 °C, thereby raising the temperature to an intermediate level. In the second stage, the intermediate heat is further boosted, with evaporation occurring at approximately 80 °C and condensation at approximately 144 °C, enabling the system to supply high-temperature process heat or steam up to approximately 139 °C.

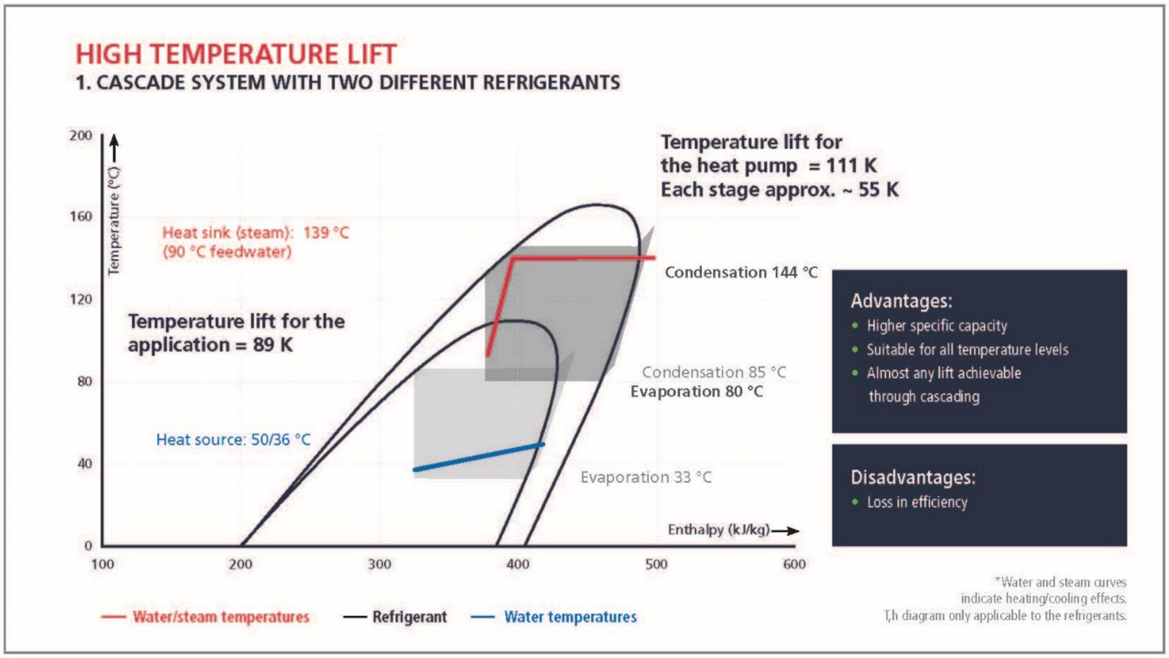

Figure 1: Temperature–enthalpy diagram of a two-stage cascade high-temperature heat pump system using two different refrigerants for process heat generation.

The use of highly efficient piston compressors in both stages, subcoolers for feedwater preheating, and parallel intermediate heat exchangers to minimize the pinch point between the hot and cold sides of the heat exchanger increased the COP in this project to 2.3. This means that 1 unit of electrical energy provides 2.3 units of heat energy, corresponding to a Carnot efficiency of 61% relative to the evaporation and condensation temperatures. Given the efficiency limitations imposed by heat transfer between the first and second stages, this value is high.

Use of Steam Compressors Expands Performance Ranges

The combination of one- or two-stage heat pump systems with steam compressors significantly extends the range of possible applications of this technology, particularly in industries with high temperature and pressure requirements. In an innovative project with the pharmaceutical company Takeda in Vienna and the AIT (Austrian Institute of Technology), high-temperature lift and high steam pressures have been applied [6]. Here, the ThermBooster™ is combined with a steam compressor to achieve higher pressures and temperatures. In this project, waste heat from a chiller used to provide cold is being reused in winter by a heat pump to generate hot water for building heating. In the remaining months, this potential had not yet been used. The innovative new system uses this hot water to generate high-quality steam. To achieve this, the efficient one-stage ThermBooster™ system uses two compressors in parallel, combined with two evaporators in series on the source side. This configuration maximizes thermodynamic performance and ensures optimum use of the available energy. The steam generated by the ThermBooster™ is compressed by a mechanical vapour recompressor (MVR) based on piston-compressor technology, which offers good part-load behaviour. Both existing and new systems use natural refrigerants:

- **ThermBooster™:**
    - Heat source: 70 °C/65 °C hot water
    - Heat sink: 1.7 bar(a) saturated steam
    - Refrigerant: Butane
    - COP: 4.3
- **MVR Spilling Project Partner GmbH:**
    - Heat sink: 11 bar(a) at 195 °C
- **Total:**
    - Heating output: 1,561 kW
    - Electrical power consumption: 672 kW
    - COP: 2.3

Figure 2 shows the cascaded high-temperature heat pump system in which a standard refrigerant cycle is combined with a steam compressor to reach very high process-steam temperatures. The heat pump first lifts the temperature from a heat source of about 65 to 70 °C to around 115 °C, corresponding to a heat-pump temperature lift of about 58 K. From this intermediate level, a mechanical vapour recompressor further compresses and superheats the steam from 115 °C up to about 195 °C, enabling the system to supply high-temperature process steam.

This configuration maximizes thermodynamic performance and optimizes the use of available energy, enabling efficient conversion of low-temperature waste heat into high-quality process heat. Thanks to their low GWP values, synthetic and natural refrigerants are possible options for significantly reducing CO_2 emissions. The system's flexibility enables its use for steam production and for specific applications, such as sterilization or drying.

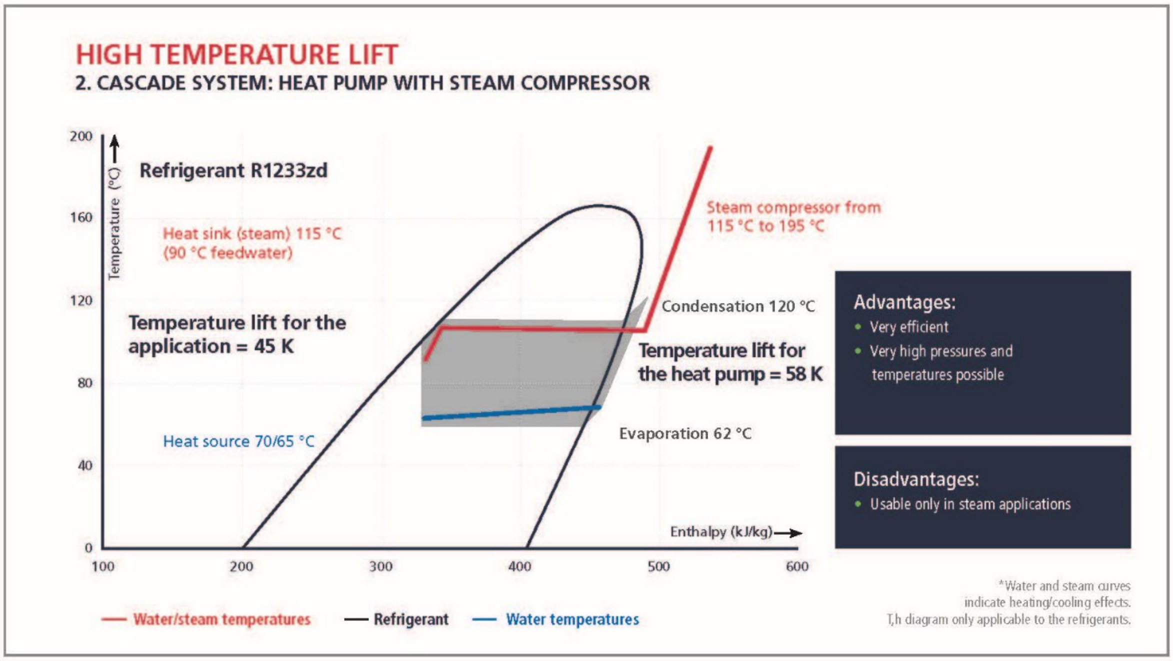

Figure 2: Temperature–enthalpy diagram of a high-temperature heat pump combined with a mechanical vapour recompressor based on piston-compressor technology.

Unused Potential in Numerous Branches of Industry: Investments that Pay Off

In addition to drying and sterilization processes, two-stage systems also show their potential in other branches of industry:

- Textile industry (washing, dyeing, pressing)
- Paper industry (bleaching, drying)
- Metal industry (electroplating, phosphating)
- Food industry (pasteurization, evaporation)

Precise temperature control, reduced energy losses, and improved efficiency are key benefits for these sectors. Especially in energy-intensive industries with high process-heat requirements, the savings in operating costs and waste heat recovery from the efficient operation of the ThermBooster™ can more than offset the higher initial investment. Modular industrial heat pumps, such as the ThermBooster™, therefore position themselves as the first choice for processes that require high temperatures with the greatest possible precision and maximum efficiency.

ThermBooster™ Technology Offers a Number of Advantages:

1. **Energy savings:** The primary energy requirement for process heat could be reduced by over 50 % by 2040 compared to 2022 [2].

2. **Cost-effectiveness:** measures that can be implemented close to the market pay for themselves quickly.

3. **Sustainability:** Electrification of process heat generation minimizes CO_2 emissions.

4. **Flexibility:** the two-stage cascading and modular design can be adapted to different requirements.

Conclusions

The decarbonization of industrial process heat is a central building block on the path to climate neutrality. Two-stage high-temperature heat pumps overcome the technical hurdles of conventional heat pump systems and offer a powerful, efficient, and economically attractive solution today. Companies that rely on this technology are therefore not only making an important contribution to the sustainable transformation of industry but are also benefiting from increased efficiency and significant cost savings.

Author contact information

Name	Christoph Rau
Title	Sales Manager
Affiliation	SPH Sustainable Process Heat GmbH
Postal address	Zur Kaule 1, 51491 Overath
E-mail address	christoph.rau@spheat.de
Phone number	+49 2206 90506-79

References

[1] Greenhouse gas emissions data for the year 2023 according to the Federal Climate Protection Act (2024): https://www.umweltbundesamt.de/sites/default/files/medien/361/dokumente/2024_03_13_em_entwicklung_in_d_ksg-sektoren_thg_v1.0.xlsx

[2] Meyer, J., Zaubitzer, L., Alsmeyer, F., and Madsen, M. "Short Study: Energy-efficient and CO_2-free Process Heat". SWK E^2 – Institute for Energy Technology and Energy Management, Niederrhein University of Applied Sciences, Krefeld, Germany. 2024.

[3] Eurostat. "Energy Balances". 2019.

[4] Fleiter T, Elsland R, Rehfeldt M, Steinbach J, Reiter U, Catenazzi G, et al. "Heat Roadmap Europe". 2017.

[5] Bauer, H., Ehrmaier, J., Gigliotti, L., Liebach, F., Schleyer, T., and Simoncini, A., "Industrial heat pumps: Five considerations for future growth". McKinsey & Company Insights. 2024.

[6] https://www.takeda.com/de-at/news/2025/ahead-opening/

Technology Collaboration Programme
by iea

Heat Pumping Technologies
MAGAZINE
Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions
Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

Non-Topical Article

AI-Driven Predictive Control for Data Center HVAC Systems

DOI: 10.23697/w508-4y55

Mamtakumari Chauhan, Jones Lang LaSalle Inc., USA

Data centers consume a significant portion of their energy in cooling (often 30–40%), making HVAC optimization critical for efficiency. Traditional rule-based HVAC controls cannot readily adapt to dynamic server workloads and changing ambient conditions, resulting in energy waste. This article proposes an AI-driven predictive control framework for data center cooling that integrates IoT sensor data (temperature, humidity, IT load) with machine learning models, specifically a reinforcement learning (RL) agent augmented with time-series forecasting (e.g., Long Short-Term Memory (LSTM) neural networks). The RL agent learns optimal cooling strategies (such as adjusting airflow and temperature setpoints) by anticipating cooling demand and continuously optimizing HVAC operations. A simulation case study and a pilot deployment demonstrate that the AI-based approach can reduce cooling energy use by approximately 15–25% relative to conventional controls, thereby improving the facility's Power Usage Effectiveness (PUE) and maintaining safe thermal conditions for IT equipment. These results highlight the potential of intelligent, predictive HVAC control to improve data center sustainability and operational reliability.

Introduction

Modern hyperscale data centers are extremely energy-intensive facilities. Global data center energy consumption was approximately 286 TWh in 2016 (roughly 1.15% of global electricity use) and is projected to reach 321 TWh (nearly 1.9% of global electricity use) by 2030[1]. A large fraction of this energy is **non-IT load,** such as cooling and air handling. In fact, studies show that air conditioning and cooling systems account for approximately 38% of total data center energy consumption [2]. This makes the HVAC (Heating, Ventilation, and Air Conditioning) system a prime target for efficiency improvements to achieve greener and more cost-effective data center operations.

However, optimizing data center cooling is challenging because traditional control strategies are often static or overly simplistic. **Conventional HVAC controls** (e.g., fixed setpoints or PID loops) are typically tuned using heuristics and safety margins that do not capture the complex, nonlinear thermal dynamics of a data center [3][4]. These rule-based sequences fail to adapt to rapid changes such as spikes in server workload or shifts in outdoor temperature. Consequently, legacy controls tend to err on the side of caution – over-cooling to avoid hot spots – which leads to **suboptimal energy use** (i.e., wasted electricity on cooling). For example, a fixed supply-air temperature may be set lower than necessary for worst-case conditions, or all Computer Room Air Handler (CRAH) / Computer Room Air Conditioner (CRAC) units may run continuously at high fan speeds regardless of the actual IT load. This inflexibility results in poor efficiency, as cooling capacity is not matched in real-time to the data center's needs[5][4].

Data center operators commonly track **Power Usage Effectiveness (PUE)** as a key metric of facility efficiency. PUE is defined as the ratio of total facility energy consumption to IT equipment energy consumption; a PUE of 1.0 indicates that all power goes to IT equipment and none to overhead, such as cooling. In typical centers, PUE ranges from about 1.3 to 2.0+, indicating that for every 1 kW delivered to servers, an additional 0.3–1.0+ kW is consumed by cooling and other overhead[2]. Improving cooling efficiency directly reduces total energy use and moves PUE closer to 1.0 [6]. For instance, Google reported that applying AI to its cooling system yielded a ~15% reduction in PUE (from ~1.45 to ~1.25) at one data center, due to a 40% reduction in cooling energy use [7][6]. This highlights how impactful better cooling control can be on overall efficiency.

AI and IoT for adaptive control: In recent years, the convergence of Internet of Things (IoT) sensing and artificial intelligence has created new opportunities to overcome the limitations of static HVAC controls. Data centers are typically instrumented with thousands of sensors that monitor temperatures at server inlets/outlets, ambient conditions, humidity levels, equipment power draw, and other parameters. Leveraging this rich real-time data, machine learning algorithms can "learn" the complex relationships between cooling settings, IT load, and thermal response [3]. Unlike fixed logic, an AI agent can continually adapt and optimize decisions based on the current state of the system. Reinforcement learning (RL), in particular, has emerged as a promising approach for autonomous HVAC control because it can learn an optimal policy by interacting with the environment (in simulation or in real-world settings) and receiving feedback in the form of energy savings versus thermal performance [8]. Moreover, combining RL with predictive models (e.g., by forecasting future loads or temperatures) enables predictive control that proactively adjusts cooling before conditions drift out of range.

There have already been notable successes in applying AI to data center cooling. Google's DeepMind team famously deployed a machine learning system to Google's data centers that continuously analyzed sensor data and recommended optimal cooling adjustments. This system was able to **reduce cooling energy use by up to 40%** – an unprecedented improvement – while keeping server temperatures within allowed limits[7][10]. The ML system accomplished this by predicting how hot the data center would become over the next hour and preventing "over-cooling" beyond what was necessary [10]. Academic research also reinforces these gains: for example, a recent study found that deep RL methods can reduce cooling costs by 11–15% relative to traditional feedback controllers while satisfying thermal constraints [13]. These results demonstrate the potential of AI-driven control to significantly reduce energy consumption in cooling systems.

This research builds on those advances by proposing an AI-driven predictive control framework for data center HVAC. In the sections that follow, we describe the system architecture leveraging IoT sensors, the design of a hybrid RL + forecasting control algorithm, and a case study evaluation. We show that the AI-based approach dynamically adjusts cooling output to match demand, yielding 15–25% energy savings and a measurable improvement in PUE in our simulations, without compromising cooling reliability. We also discuss practical considerations – such as cybersecurity and integration with industry standards like ASHRAE Guideline 36 – for deploying such AI controllers in production data centers. The goal is to illustrate how combining real-time data, machine learning, and domain knowledge can enable more sustainable and efficient data center operations.

Methodology

IoT Sensor Architecture for Cooling Control: The proposed framework relies on a dense array of IoT sensors to monitor environmental and operational parameters throughout the data center. **Figure 1** illustrates a high-level architecture of the system. Temperature sensors are deployed at server inlets (to measure intake air temperature to the IT equipment) and at various points along the cold and hot aisles. Additional sensors measure return-air temperature, humidity, and the differential pressure between the aisles. Power meters provide IT load measurements (server power draw) and cooling system electrical demand (e.g., chiller and fan power). All these sensors feed data into a central Building Management/System (BMS) or Data Center Infrastructure Management (DCIM) platform via an IoT communication network (for example, using protocols like MQTT or BACnet over IP). The **controller** component (highlighted in the figure) ingests streaming sensor data and executes AI algorithms in real time. It then issues control commands back to the HVAC actuators – such as setpoint adjustments to Computer Room Air Handler (CRAH/CRAC) units, variable fan speeds, chilled water valve positions, or compressor settings – through the BMS interface. By continuously sensing and acting, this closed-loop system can respond quickly to changes in the data center's thermal state[14][15]. The IoT-based architecture thus provides the foundation for fine-grained monitoring and control, enabling predictive AI to observe current conditions and effect changes through the existing HVAC infrastructure.

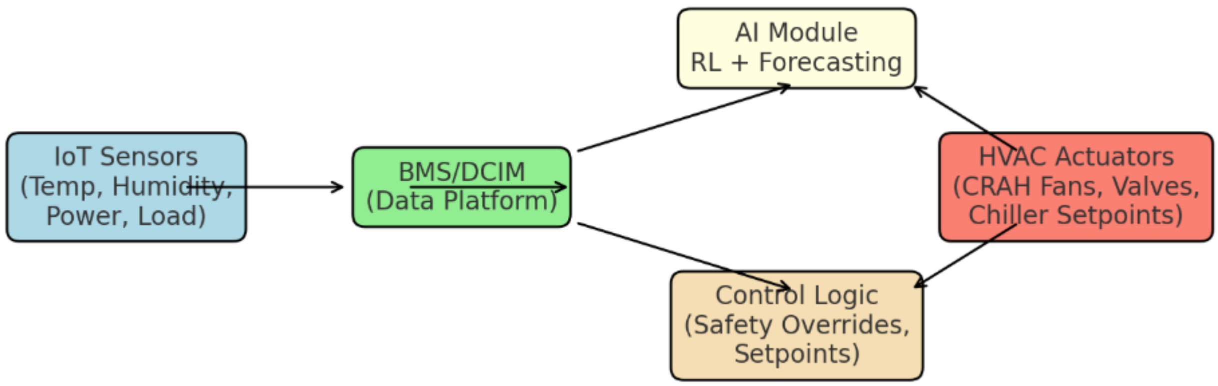

Figure 1: IoT-Enabled AI Predictive Control Architecture

Integration with the BAS Control Loop: Deploying the AI controller requires integration with the existing building automation system. Typically, the RL policy is implemented as a software service that runs either on a local industrial PC or an on-site edge computing device, or as part of the data center's central control software. Each control cycle (say every 5 minutes), the following loop is executed: (1) read current sensor data from the BMS (temperatures, humidity, IT load, etc.), (2) update the forecasting model to predict near-future values, (3) let the RL agent compute the optimal control action given the current state and forecast (this is a fast computation – a forward pass through the neural network policy), and (4) send the new setpoints/commands to the HVAC equipment via the BMS interface. The BMS then implements the changes (e.g., raises the CRAH supply-air temperature setpoint or reduces fan speed). This process repeats continuously. In essence, the **AI controller replaces or augments the traditional HVAC control logic** with an intelligent, adaptive policy. For safety, we include override mechanisms: if the AI output ever appears to drive conditions outside of allowable ranges (e.g., if a sensor detects temperature approaching an alarm threshold), the BMS can override to a default safe mode or fall back to a known-good control sequence (such as ASHRAE Guideline 36 sequence) as a fail-safe. In practice, the AI controller is often introduced gradually – for example, running in a **recommendation mode** in parallel with existing controls to build trust, before ultimately taking autonomous control once it has proven stable.

A crucial aspect of integration is managing **latency and response time**. The control loop interval should be tuned to the data center's thermal inertia. Large data centers have significant thermal mass and may not require ultra-fast control cycles; a 5–15-minute interval can be sufficient to adjust cooling gradually. This also gives time to collect data and make robust predictions. Our RL agent's decisions are explainable in that we log sensor inputs and selected actions for operators to review. Although the internal reasoning of a neural network is complex, the system can present its outputs as recommended setpoints that operators can compare with their own expectations. This helps with gaining operational acceptance of the AI. Over time, as the agent continues to learn (we can periodically retrain it on newer data or even allow limited online learning), it adapts to environmental changes, such as the deployment of new IT equipment or HVAC system upgrades. By tightly integrating sensing, prediction, and control, the AI-driven system continuously optimizes the cooling plant's efficiency while respecting all constraints.

Case Study and Simulation Setup

To evaluate the proposed AI-driven cooling control, we conducted a case study using a combination of simulation (digital twin) and a limited field test. The test environment represents a mid-size data center module, akin to a small section of a hyperscale facility or an enterprise data center. In the simulation, we modeled a data hall with 20 server racks (each drawing up to 5–10 kW under peak load) arranged in a hot-aisle/cold-aisle containment layout. Cooling was provided by two computer room air-handling (CRAH) units configured in an N+1 redundancy scheme (only one unit required under normal load, with the second for failover or high-load conditions). The CRAH units supply cold air beneath the raised floor to maintain the cold aisle and draw return air from the hot aisle. The simulation encompassed key thermal characteristics: server airflow and heat output, room air mixing, CRAH fan laws, and a chilled-water loop serving the CRAHs. This digital twin was calibrated against real data (temperatures and power readings) from a similar facility to ensure realism [18]. We also included models for **external influences** – for instance, the intake air temperature could be affected by outdoor conditions if the data center used some outside air economization. For simplicity, our base scenario assumed the data center is in a controlled environment with HVAC systems only (no direct outside air), though the methodology is equally applicable to facilities with air-side or water-side economizers.

We simulated various **server workload patterns** to test the controller under dynamic conditions. This included a steady baseline load (representing typical utilization) with periodic spikes in which additional compute jobs increased server utilization by 20–30% for a few hours (e.g., peak-traffic periods or batch-processing windows). We also simulated diurnal temperature changes: during "daytime" hours, the ambient temperature in the cooling intake rose (to mimic either warmer outdoor air intake or other heat contributions), whereas at “night” it was cooler. These variations required the cooling system to adjust capacity over time. The RL agent (with forecasting) was deployed in this simulation to control the CRAH supply air temperature setpoint and the CRAH fan speed. The agent’s goal was to minimize energy use of the CRAHs (fan power and chilled water cooling power) while keeping all server inlet temperatures below 27 °C (which is in line with ASHRAE’s recommended maximum for most data center equipment).

For **baseline comparison**, we implemented a conventional rule-based control strategy akin to those found in many operational data centers. The baseline used a fixed supply air temperature setpoint of 22 °C on the CRAHs, with a traditional PID loop modulating fan speed to maintain a set differential pressure between the cold and hot aisles (to ensure adequate airflow). Essentially, this mimics static setpoints that are configured to handle worst-case load (22 °C supply was chosen to meet cooling needs at peak load comfortably). No predictive adjustments were made in the baseline – it reacted to temperature deviations but had no knowledge of incoming load changes. This baseline also reflects the spirit of standard sequence controls (though not fully as advanced as ASHRAE Guideline 36, which would do more dynamic resets; we consider G36 in the Discussion section).

We ran the simulation for an extended period (several days' worth of virtual time) to cover different load scenarios. Key **performance metrics** were collected for both the baseline and the AI controller runs: total cooling energy consumption (kWh) over time; average and peak PUE; the time series of cold-aisle and server inlet temperatures; and any instances in which temperature thresholds were exceeded. We also measured fan utilization (i.e., how often fans operated at high speed) and control stability (e.g., whether the AI caused large oscillations or rapid changes). In addition, to validate real-world feasibility, we deployed the AI controller in a limited **field pilot** on an actual data center CRAH unit for a short duration (with operator oversight). This allowed us to verify that the RL agent’s actions in simulation translated well to the behavior of physical equipment and that no unexpected issues (such as sensor noise or actuator delays) would disrupt the control logic in practice.

Results

Energy Savings and Efficiency Improvements: The AI-driven predictive control demonstrated substantial energy savings in cooling operations compared with traditional control. In our simulations, the RL-based controller consistently reduced the cooling energy consumption by roughly 20% on average compared to the baseline rule-based strategy (with variations across different scenarios ranging from about 15% up to 25% savings). This outcome is in line with results reported by prior works – for example, [13] noted deep RL could save around 11–15% of data center cooling costs under stable conditions. In one scenario with high IT load variability, our AI controller used ~22% less cooling energy than the baseline over a 24-hour period. These savings directly translate to better PUE. For instance, during a high-load test day, the baseline PUE averaged 1.35, whereas with AI control, the PUE dropped to ~1.27, reflecting the lower overhead energy usage. Such improvement is very significant at scale – a 0.08 reduction in PUE means a large data center (tens of MW of IT load) can reclaim megawatts of power capacity or reduce electricity costs proportionally.

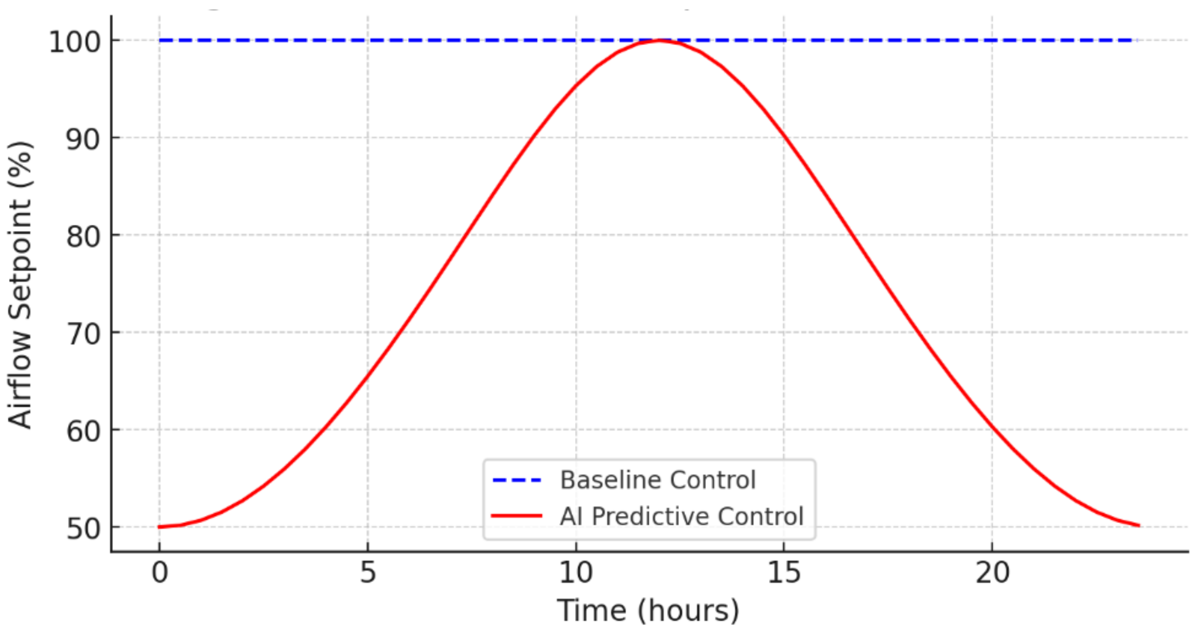

Figure 2: Comparison of cooling airflow setpoints over time under AI control (solid red line) versus the baseline BMS control (dotted blue line) for one week of operation.

The RL-based controller dynamically adjusts the supply air flow (and cooling output) in response to workload and temperature forecasts, often reducing airflow during low-demand periods and preemptively increasing it before peak demand. The conventional control maintains a relatively static or linear setpoint, resulting in higher airflow (and energy use) than necessary during off-peak periods. The AI controller's optimized profile reduces fan energy consumption while still meeting cooling requirements.

Figure 2 above illustrates how the RL controller's actions differed from the static baseline. The baseline (blue dashed line) maintained airflow at a high, relatively flat level to ensure cooling under peak-load conditions at all times. In contrast, the AI (red line) was able to **trim the airflow during lighter loads** (for example, each night when server usage dipped, the AI agent lowered the CRAH fan speeds significantly, saving energy). When a load increase was anticipated (e.g., just before midday peaks), the agent temporarily increased airflow setpoints in advance – a form of pre-cooling – so that temperatures stayed in range without needing an overshoot of cooling later. This predictive modulation is what enabled energy reduction without sacrificing performance. Notably, even during the peak periods, the AI's airflow was slightly lower than the baseline's, yet it satisfied cooling needs because it had optimized the balance of temperature and flow (running a bit warmer but still safe). These results confirm the effectiveness of the RL+forecasting approach in **right-sizing cooling output** to the actual demand.

Thermal Performance and Reliability: Crucially, the AI-driven control maintained thermal conditions within acceptable limits throughout our tests. In no instance did server inlet temperatures exceed the recommended maximum (we set 27 °C as the limit) under the RL controller's watch. The lowest observed cold aisle temperatures were around 18 °C (during over-provisioning by baseline controls), and the RL agent allowed cold aisle temps to rise closer to 25–26 °C during low-risk periods – which is still well within safe operating range for equipment, and in fact closer to ideal from an efficiency standpoint. At peak IT loads, the hottest server inlet temperature under AI control was approximately 27 °C, equal to our threshold but not exceeding it (whereas the baseline kept the hottest inlet around ~24 °C by brute-force cooling). This indicates that the AI was using the full allowable thermal headroom to save energy without compromising **reliability**. We also monitored relative humidity and found that it remained within the normal range (in this simulation, 45–50% RH) for both the baseline and AI cases; the AI did not adversely affect humidity control.

Another aspect of reliability is avoiding rapid oscillations or equipment wear. Our RL agent's policy resulted in **smooth control adjustments**: it typically made small setpoint changes and then observed, rather than constantly thrashing the equipment. This is partly due to the fact that we trained the agent with a penalty for instability. The frequency of setpoint changes under RL was approximately every 5–10 minutes, which is manageable for HVAC equipment. The baseline PID sometimes caused more frequent fan-speed hunting when conditions changed quickly, whereas the RL, with its predictive outlook, made more gradual adjustments. We observed no alarms or failures attributable to the AI control during the field test. For example, the CRAH fans under AI control ran at lower speeds most of the time (extending their lifespan) and only ramped up to high speeds when absolutely necessary (and even then, for shorter durations than in baseline control). In summary, the **thermal stability** under AI control was excellent: the system met all ASHRAE thermal guidelines (e.g., maintaining server inlet temperatures in the recommended 18–27 °C range [23]) and did so with fewer resources. The results provide confidence that an AI-driven approach can meet the stringent reliability requirements of data center environments while significantly improving efficiency.

Discussion

Practical Challenges for Deployment: While the case study demonstrates clear benefits, implementing AI-driven HVAC control in real production data centers presents practical challenges. **Cybersecurity** is a top concern. Introducing IoT sensors and networked controllers opens potential attack surfaces in a mission-critical facility. If a malicious actor were to gain access to the cooling control system, they could theoretically manipulate it to disrupt operations (for instance, turning off cooling to cause overheating). In fact, cybersecurity analysts warn that building management systems and IoT devices (such as smart HVAC controllers) are increasingly targeted by hackers [24]. Scenarios have been discussed in which attackers exploit default passwords or vulnerabilities in connected thermostats/CRAH controllers to launch a "thermal attack," raising server temperatures and even forcing shutdowns [25]. To mitigate this, strong security measures must be in place: isolating the HVAC control network from external networks, using encryption and authentication for sensor data and control commands, and implementing strict access controls. Regular security audits and firmware updates for IoT devices are also necessary to patch any vulnerabilities. Industry bodies recognize the importance of cybersecurity; for example, U.S. federal guidelines have highlighted the need to secure data center infrastructure management and IoT systems to prevent such attacks[26][27]. When deploying an AI HVAC solution, operators must ensure that the system is as resilient and secure as other critical infrastructure in the data center.

Another challenge is the **lack of explainability and trust** in AI decisions. Data center operators and facility engineers are accustomed to control sequences that are easy to understand (e.g., *"if temperature > X, then increase cooling"*). In contrast, a reinforcement learning agent might make non-intuitive adjustments because it's optimizing long-term rewards. This "black box" nature can make operators uneasy – after all, the stakes are high if the cooling system misbehaves. Therefore, building trust in the AI system is crucial. One approach is to implement AI control gradually, starting with recommendations or advisory mode, in which the AI suggests setpoints and humans approve them, and then moving to closed-loop control once confidence is established. It also helps to add

interpretability features: for example, the system can display which sensor readings contributed most to the latest action, or what the predicted temperature vs. actual outcome was. Developing more transparent RL algorithms (or applying explainable AI techniques to RL) is an active research area that could benefit this application. In practice, we found that a well-defined reward function and constraints made the RL's behavior more predictable; for example, the agent was constrained to never exceed certain temperatures, which provided a layer of assurance. In summary, organizations may need to establish operational procedures and tools for AI controllers, including fallbacks, dashboards for monitoring AI actions, and, in critical facilities, an AI "safety officer" role to oversee these systems.

Comparison to ASHRAE Guideline 36 Sequences: ASHRAE Guideline 36 is a set of standardized, high-performance sequences of operation for HVAC systems, including building cooling systems. It represents the state of the art in rule-based control, including optimized reset schedules, trim-and-respond logic for supply-air temperature, static-pressure resets, and advanced economizer control. Guideline 36 (G36) has been shown to improve efficiency compared with older ad hoc control schemes significantly. For example, modeling studies suggest that G36 sequences can reduce HVAC energy use by more than 30% relative to conventional fixed controls, depending on the system and climate [28]. It achieves this through careful programming of how setpoints are adjusted in response to conditions (all firmly coded, not learned). Given that, a pertinent question is how an AI-driven approach stacks up against G36 – is it truly better, or just another way to reach similar outcomes?

Early evidence indicates that **AI can outperform G36** in terms of pure efficiency, though by a smaller margin, and with additional complexity. A recent field study by [29] compared a deep RL control with a G36-compliant control in a real building and found that the G36 logic saved approximately 42% of HVAC energy relative to a baseline, whereas the RL-based control achieved roughly a 54% reduction under the same conditions. In other words, the RL provided an additional ~12 percentage points of savings beyond the best conventional sequence. This is reasonable: the RL agent can fine-tune continuously and may identify micro-optimizations that a generic sequence cannot. It might, for example, learn an aggressive strategy tailored to the pattern of IT load in that data center, whereas G36 is a one-size-fits-all set of rules. That said, G36 is much easier to implement widely – it's deterministic, tested, and doesn't require site-specific training data or AI expertise to deploy. Many data centers today have not implemented G36 or similar advanced sequences; doing so would yield substantial improvements. We argue that AI-driven control is complementary to these standards. In fact, one could initialize an RL agent with a policy that mimics G36 (to provide a good starting point) and then let it learn further improvements. G36 can also serve as a **safety net** – if the AI ever behaves strangely, reverting to a G36 sequence would maintain safe operation with decent efficiency. Over time, as AI control becomes more proven, we might see standard sequences evolve to incorporate AI principles or guidelines on how to properly train and validate RL controllers for HVAC.

Furthermore, one must consider maintainability: G36's fixed sequences might be easier for an engineer to troubleshoot when something goes wrong (since they can step through logic charts), whereas an RL agent might require specialized skills to diagnose (one might have to check reward logs, or retrain if the environment changes). However, AI can adapt more effectively to nonlinear and changing conditions. For instance, if a data center adds a new cooling unit or rearranges the floor layout, a G36 sequence would require retuning by an engineer, whereas an RL agent could theoretically adapt through further learning. In summary, **AI versus G36 is not an either-or proposition: G36 provides a strong benchmark and starting point, and AI controllers have demonstrated the ability to further improve efficiency and address** complexities that pre-programmed logic might miss. The combination of both – using G36 as a baseline and allowing AI to optimize on top of it continuously – could be a powerful strategy for industry adoption.

Other Considerations: There are a few additional practical considerations. One is the **computational cost** of running AI models. Fortunately, modern data center control systems can leverage on-premises servers (or existing server infrastructure) to host the RL controller and forecasting models. In our implementation, inference (decision-making) takes only milliseconds on a CPU, which is negligible relative to a control interval of minutes. Training is more intensive but can be done offline on powerful machines (or cloud resources). Another consideration is **sensor reliability and calibration**: the AI's decisions are only as good as the data it gets. A mis-calibrated temperature sensor could mislead the agent. Thus, regular calibration and sensor health checks are vital. The AI system can also help: it can flag inconsistent readings (as anomalies) when one sensor deviates substantially from the others. Maintenance teams should include the sensor network in their routine checks when AI control is in place.

Lastly, deploying AI in a traditionally conservative domain requires addressing the **human factor**. Data center operators require training to understand how AI works and to interpret its actions. There might be organizational resistance initially – after all, it's a shift from manual or semi-manual control to trusting a machine. Pilots and gradual rollouts, as well as demonstrating success metrics (like "AI saved X MWh of energy this quarter"), will be key in gaining buy-in. Companies may designate champions or dedicated teams to oversee AI operations within facilities. With time, as success stories accumulate, AI-driven HVAC optimization could become a standard tool in the data center energy management toolbox.

Conclusions and Future Work

This work presented an AI-driven predictive control approach for data center HVAC systems and demonstrated its potential to improve energy efficiency while maintaining thermal reliability significantly. By combining **reinforcement learning** with **time-series forecasting**, our framework can anticipate cooling demand and adjust HVAC operations more intelligently than static rule-based controllers. The case study results showed cooling energy reductions of 15–25%, which translate into lower operating costs and improved PUE for the facility. Importantly, these gains were achieved without violating temperature and humidity constraints; the AI controller maintained server inlet temperatures within safe limits at all times, thereby balancing performance and efficiency. This confirms that an appropriately trained AI policy can safely navigate the complex thermal dynamics of a data center. For industry professionals facing rising energy costs and sustainability goals, such AI-driven optimization offers a promising solution to reduce energy consumption in data centers (cooling systems) while also enhancing environmental control resilience.

For **future work**, several directions appear attractive to further develop and deploy this technology:

- **Scaling to Hyperscale Facilities:** Large cloud data centers comprise many rooms and a hierarchy of cooling systems (CRAH units, chillers, cooling towers, etc.). Scaling our single-module controller to coordinate at a whole-facility level is the next step. This may involve a multi-agent RL system or a hierarchical control approach (e.g., one agent optimizing chiller plant operation and others focusing on room-level airflow). Ensuring stability and interoperability in such a multi-zone environment will be important. This indicates that scaling is feasible, although engineering effort is required to integrate AI into each site's infrastructure.
- **Integration with Renewable Energy and Demand Response:** As data centers increasingly commit to using renewable energy, an AI cooling controller can help align

cooling power consumption with renewable energy availability. One idea is to incorporate dynamic electricity pricing or carbon-intensity signals into the control strategy. Our RL formulation can be extended by including the current power price or grid CO_2 factor in the state and reward – incentivizing the agent to, say, pre-cool (store "cooling" in thermal mass) when electricity is cheap/clean and allow temperatures to float a bit higher when electricity is expensive or in short supply. This effectively renders the cooling system a **flexible load** that can interact with the smart grid. Prior research on load shifting using RL for HVAC shows that this is a viable approach; for example, work by [31][32] used an RL agent with a recurrent network to respond to real-time electricity prices, yielding cost savings while keeping data center temperatures within the range. In practice, future data centers could integrate such algorithms to reduce cooling power draw during grid peak events automatically or to maximize the use of on-site solar/wind generation for cooling. Additionally, AI control could coordinate with **emerging cooling technologies** – e.g., adjusting liquid cooling pumps or phase-change cooling systems – and even facilitate **waste-heat reuse** by precisely controlling when and how excess heat is harvested, all in accordance with external energy conditions.

- **Continuous Learning and Adaptation:** Data center conditions are not static – IT loads evolve (especially with rapid growth in AI computing workloads), and facility upgrades or layout changes can alter cooling patterns. A deployed RL agent may need periodic retraining or adaptation to maintain optimal performance. Techniques such as **lifelong learning** or transfer learning in RL could be employed to enable the agent to continue learning from new data without forgetting prior knowledge. For example, physics-informed approaches such as the "Phyllis" framework have been proposed to enable safe, rapid retraining of data center RL controllers when the environment changes, achieving 8–10× faster adaptation to new conditions with minimal performance degradation [33]. Implementing an online learning loop – carefully, with validation steps to ensure safety – could allow the AI controller to improve itself over time and adjust to things like **capacity expansions, new cooling equipment, or climate change effects** on the facility. We also envision more advanced predictive models (e.g., ensemble forecasts or weather forecasts) to further enhance the controller's foresight.
- **Standardization and Compliance:** For widespread adoption, the industry may develop standards or guidelines for AI-driven control (just as ASHRAE Guideline 36 standardized best-practice sequences). Collaborations between AI researchers and organizations like ASHRAE could yield reference designs for "autonomous cooling optimization systems" that ensure interoperability with existing BMS products and compliance with safety standards. There is also room to incorporate explainability and fail-safe mechanisms into these standards, so that facility operators have confidence in delegating control to AI. Demonstrating reliability over long periods

(e.g., multi-year continuous operation without incidents) will be key to convincing stakeholders of the technology's maturity.

In conclusion, AI-driven predictive control for data center HVAC has demonstrated compelling benefits in energy efficiency and has a clear pathway to augmenting current best practices. As data centers continue to grow in scale and importance, such intelligent control systems will be instrumental in managing energy demand and reducing the environmental footprint. By integrating advanced sensors, machine learning algorithms, and robust control engineering, future data centers can be made smarter – **automatically optimizing cooling performance** in real-time, reacting to both internal IT needs and external grid conditions. This aligns with broader trends of automation and smart infrastructure in the industry. Continued research and pilot implementations will help address remaining challenges, and we anticipate that in the coming years, AI-based HVAC control could become a standard feature of green, efficient data centers, working hand-in-hand with human operators to achieve **sustainable computing at scale**.

Acknowledgments: *The authors thank the data center operations team for their support during the field tests and the anonymous reviewers for their insightful feedback.*

Author contact information

Name	Mamtakuamri A. Chauhan
Title	Controls System Engineer
Affiliation	Jones Lang LaSalle Inc.
Postal address	
E-mail address	Chauhanmamta55@gmail.com
Phone number	5512548389

References

[1] [2] [4] [8] [18] portal.research.lu.se

https://portal.research.lu.se/files/141036901/ACSOS_2022_RafsineRL_No_template_.pdf

[3] [5] [17] DeepMind AI Reduces Google Data Centre Cooling Bill by 40% - Google DeepMind

https://deepmind.google/discover/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40/

[6] [7] [10] Google harnesses the power of AI to cut energy use | World Economic Forum

https://www.weforum.org/stories/2016/07/google-harnesses-the-power-of-ai-to-cut-energy-use/

[13] [33] Phyllis: Physics-Informed Lifelong Reinforcement Learning for Data Center Cooling Control

https://tanrui.github.io/pub/Phyllis-eEnergy23.pdf

[14] [15] [31] [32] orbit.dtu.dk

https://orbit.dtu.dk/files/317883225/RLforHVACControl_IEEE_IoT_1_.pdf

[23] [PDF] ASHRAE TC9.9 Data Center Power Equipment Thermal Guidelines ...

https://www.ashrae.org/file%20library/technical%20resources/bookstore/ashrae_tc0909_power_white_paper_22_june_2016_revised.pdf

[24] [25] [26] [27] U.S. Data Centers Could Face Security Threats Amid AI Boom

https://broadbandbreakfast.com/u-s-data-centers-could-face-security-threats-amid-ai-boom/

[28] Energy performance evaluation of the ASHRAE Guideline 36 control and reinforcement learning–based control using field measurements | Request PDF

https://www.researchgate.net/publication/385643243_Energy_performance_evaluation_of_the_ASHRAE_Guideline_36_control_and_reinforcement_learning-based_control_using_field_measurements

[29] Simulation-based assessment of ASHRAE Guideline 36 ...

https://www.sciencedirect.com/science/article/abs/pii/S0360132323003980

Technology Collaboration Programme
by iea

Heat Pumping Technologies
MAGAZINE
Heat Pumps in Industrial Applications: Pioneering High-Temperature Solutions
Vol.43 No3/2025
A HEAT PUMP CENTER PRODUCT

National Market

Sweden: Heat Pump Market Report

DOI: 10.23697/p2sk-7x35

Metkel Yebiyo, Research Institute of Sweden, Sweden

Sweden remains one of the world's most mature and saturated heat pump markets. With a high penetration rate in single-family homes, the market dynamic has shifted significantly from "first-time adoption" to "replacement and modernization." While the industry comes off a record-breaking period, 2024 has introduced a notable cooling in immediate demand due to economic headwinds, even as the long-term outlook remains overwhelmingly positive. Swedish heat pump statistics are collected by the Swedish Heat Pump Association (Svenska Kyl & Värmepumpföreningen, SKVP), which has tracked sales since 1982 and publishes quarterly data (see Figure 1). The data used in this report has been extracted from the SKVP website.

Market Size & Sales Development

The heat pump industry has continuously collected statistics on heat pumps sold in Sweden since 1982. Each quarter, the most recent sales figures are reported, and the dataset provides a comprehensive overview of market development from 1982 to the present.

As illustrated in Figure 1, heat pump sales in Sweden have undergone several distinct phases over the past four decades. Following relatively modest volumes during the 1980s and early 1990s, the market expanded rapidly from the late 1990s, driven by rising energy prices, policy support, and growing awareness of energy efficiency. This growth accelerated in the early to mid-2000s, when annual sales exceeded 100,000 units, marking the first major expansion of the Swedish heat pump market.

The graph also highlights significant shifts in the composition of technologies over time. While ground-source (brine-to-water) and exhaust air heat pumps played an important role in the early growth phase, air-source heat pumps, particularly air-to-air and air-to-water systems, have become increasingly dominant since the mid-2000s. In recent years, air-to-water heat pumps have experienced strong growth, reflecting their suitability for hydronic heating systems and their potential to replace fossil-fuel-based boilers in single-family homes.

Market fluctuations are clearly visible, with periods of decline corresponding to economic downturns, changes in subsidy schemes, and external shocks. Notably, sales peaked again around 2021–2022, reaching historically high levels, before declining in 2023–2024. This recent downturn likely reflects a combination of high interest rates, reduced household purchasing power, market saturation in some segments, and the normalization of conditions following exceptionally strong demand during the energy price crisis.

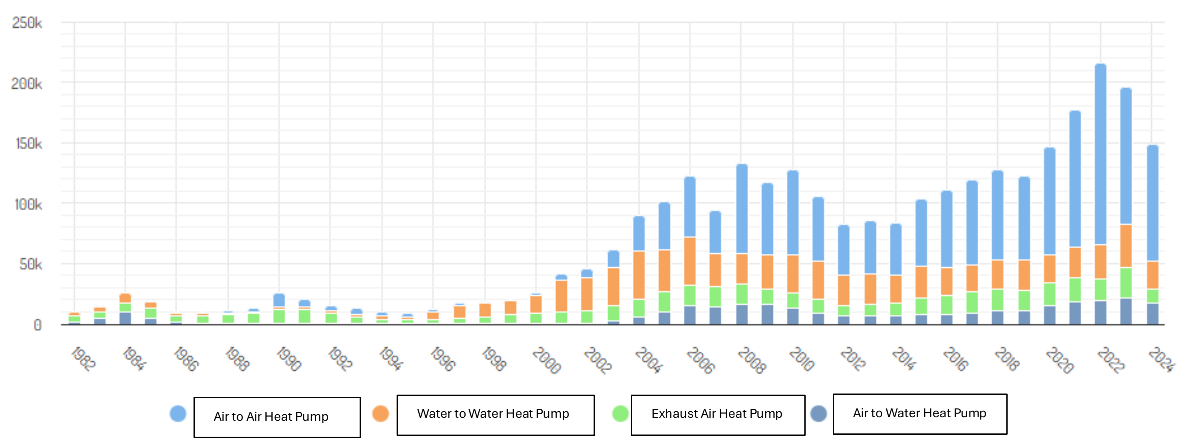

Figure 1: Heat Pump Sales in Sweden 1982 – 2024 (Source: https://skvp.se/nyheter-o-statistik/statistik/varmepumpforsaljningen)

Installer Sentiment: (The Pulsen Survey)

After a historic peak in sales and interest, the Swedish market is currently experiencing a correction. High inflation and interest rates have impacted consumer purchasing power, leading to a visible shift in installer sentiment compared to the previous year.

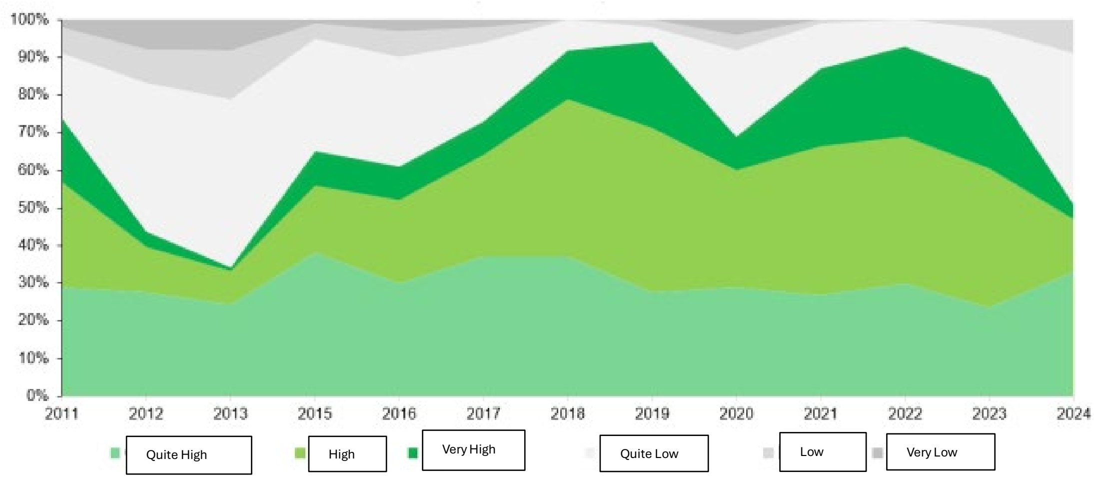

Figure 2: Market Sentiment / Demand Trend 2011–2024 (Source: https://skvp.se/nyheter-o-statistik/statistik/varmepumpforsaljningen)

As shown in **Figure 2**, the perceived demand has dampened significantly. In the 2023 survey, **85%** of installers described demand as positive (Very High, High, or Quite High). In 2024, that figure has dropped to **51%**. However, this dip is widely viewed as a temporary stabilization rather than a structural decline. When asked about the next three years, confidence remains robust, with **94%** of respondents predicting positive demand ahead, actually an increase from 92% the previous year.

National Sales Statistics (Hard Data)

While the Pulsen survey tracks sentiment, the national sales statistics from the Swedish Refrigeration & Heat Pump Association (SKVP) reveal the actual volume trends. Following the exceptional "boom" year of 2023, the market experienced a significant correction in 2024, though specific segments showed resilience by year-end.

- **Full Year 2024 Performance:** The total market contracted in 2024 as distributor stocks remained high and households paused investments due to economic uncertainty.
    - **Exhaust Air Heat Pumps (Frånluft):** This segment was hit hardest, with sales declining by **54%** for the full year.
    - **Ground Source Heat Pumps (Bergvärme):** Sales decreased by **34%** compared to the 2023 peak.

  - **Air-to-Water Heat Pumps:** Showed more resilience but still finished the year down by **20%**.
- **The Turning Point (Q4 2024):** Despite weak annual figures, the market began to turn a corner in the fourth quarter. As seen in Figure 3, Air-to-Water heat pumps broke the negative trend, recording a 16% increase in sales during Q4. This signals that inventory issues have cleared and demand is returning as interest rates stabilize.

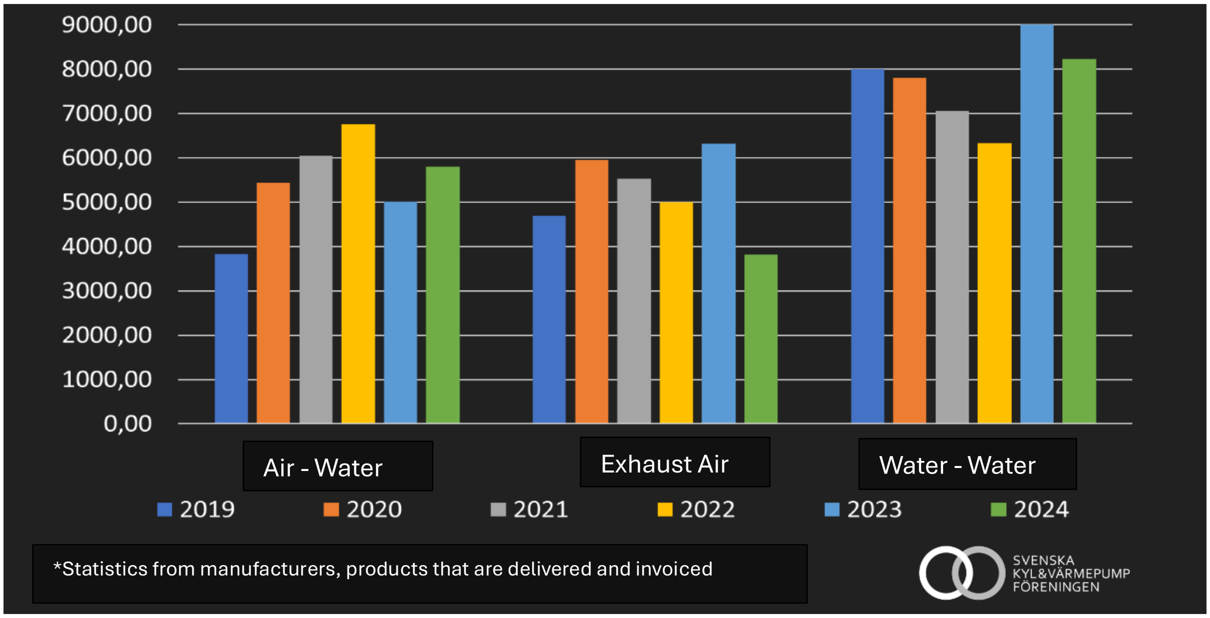

Figure 3: Sales volumes (units) for Q4 2024 compared to previous years. Note the recovery in Air-to-Water (Luft-vatten), represented by the green bar surpassing the 2023 blue bar, contrasting with the continued decline in Exhaust Air (Frånluft) (Source: https://skvp.se/nyheter-o-statistik/statistik/varmepumpforsaljningen)

1. **The Replacement Era**

Sweden has effectively entered the "Replacement Era." The massive wave of installations from the 1990s and 2000s has now reached the end of life, creating a reliable baseload of sales volume that is less dependent on new-construction trends.

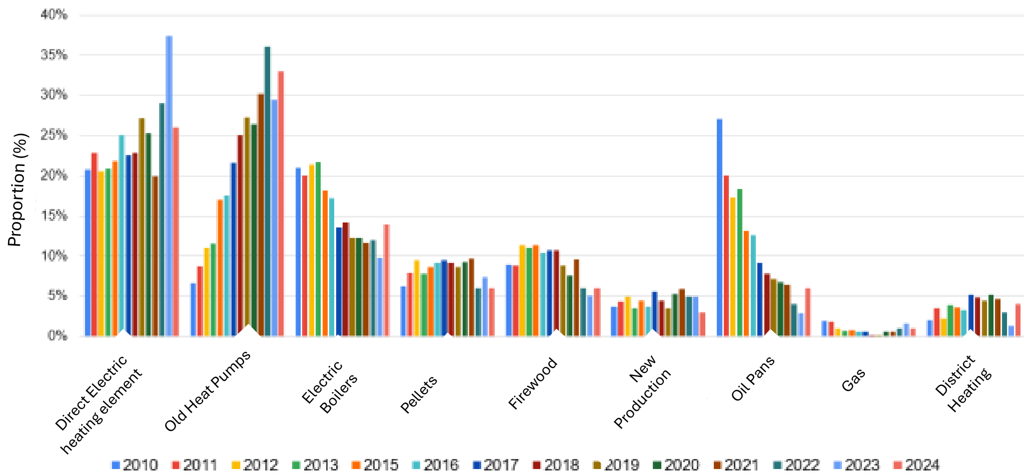

Figure4: Distribution of heating systems replaced or supplemented by new heat pump installations (2010–2024). The data highlight Sweden's shift to a "replacement market," in which replacing an older heat pump (Äldre värmepump) is now the leading installation scenario, surpassing direct electric heating (Direktverkande el). (Source: https://skvp.se/nyheter-o-statistik/statistik/varmepumpforsaljningen)

Figure 4 illustrates this pivotal shift. For the first time, replacing an **older heat pump** is the single most common reason for a new installation, accounting for **33%** of all projects.

- **Direct Electric Heating:** Still a major driver, accounting for **26%** of replacements.
- **Electric Boilers:** Account for **14%** of replacements.
- **District Heating:** There is also a noted increase in heat pumps replacing district heating connections, driven by rising prices in that sector.

Technology Trends: The Shift to Propane (R290)

European F-gas regulations are driving a rapid technological transformation in the Swedish hardware stock. Manufacturers are quickly pivoting toward natural refrigerants to future-proof their product lines.

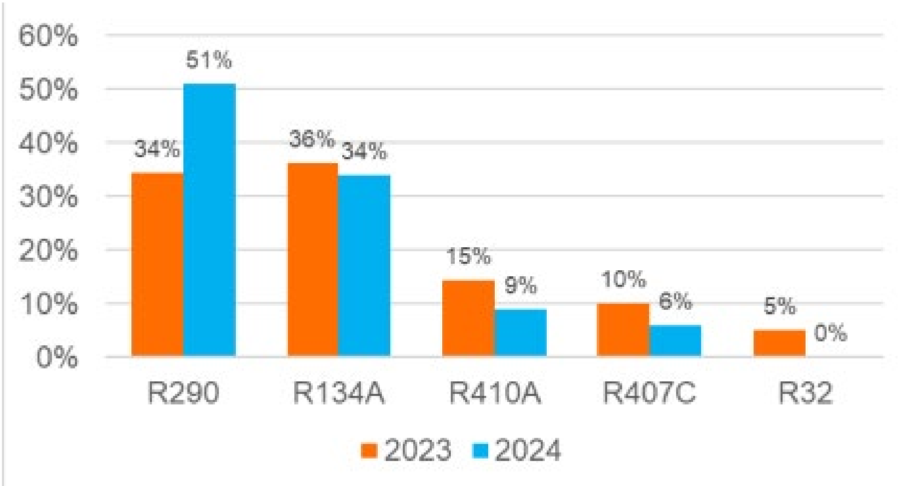

Figure 5: Refrigerant Trends (Exhaust Air) (Source: https://skvp.se/nyheter-o-statistik/statistik/varmepumpforsaljningen)

As detailed in **Figure 5**, the adoption of R290 (Propane) has accelerated sharply between 2023 and 2024:

- **Exhaust Air Heat Pumps:** The shift is most dramatic here. The share of units using R290 jumped from **34%** in 2023 to **51%** in 2024, making it the dominant refrigerant in this segment.
- **Air-to-Water Heat Pumps:** R290 share grew from **2%** to **12%** in just one year.
- **Air-to-Air Heat Pumps:** R32 remains the standard, commanding **95%** of the market.
- **Ground Source Heat Pumps:** This segment is moving more slowly, with R410A still holding **53%** of the market, though R32 has grown to **21%**.

Price Stability

Following a period of volatility and price hikes in 2023, the cost of installing a heat pump in Sweden has stabilized.

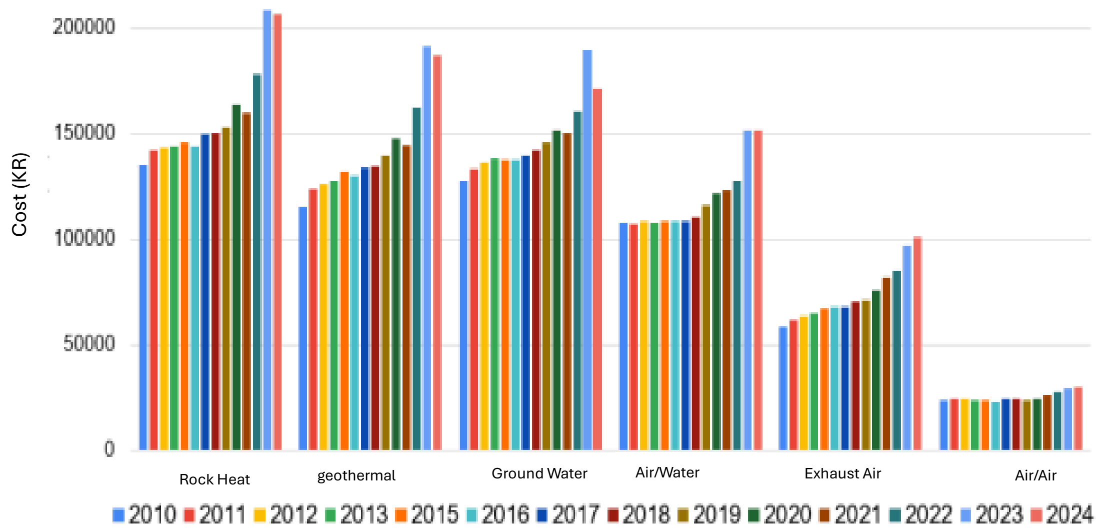

Figure 6: Total installation cost (including the heat pump) development (2010–2024) for a standard single-family home with a heating demand of 20,000 kWh/year. (Source: https://skvp.se/nyheter-o-statistik/statistik/varmepumpforsaljningen)

Figure 6 tracks the total cost (hardware + installation) for a standard residential project. After sharp price increases in previous years, 2024 shows a stabilization trend, with costs for Geothermal (Bergvärme) and Air-to-Water (Luft/Vatten) systems remaining largely in line with 2023 levels.

- **Geothermal (Bergvärme) and Air-to-Water:** Prices have flattened after previous increases.
- **Exhaust Air (Frånluft):** Saw a modest increase of approx. **4%**.
- **Air-to-Air (Luft/Luft):** Continued a trend of slight increases, rising by approximately **2%**.

Conclusion

While 2024 presents a challenging economic environment, the fundamentals of the Swedish market remain strong. The high replacement volume, combined with a 94% positive 3-year outlook from industry professionals, suggests that Sweden will continue to serve as a leading model for heat pump maturity and technological innovation in Europe.

Author contact information

Name	Metkel Yebiyo
Title	PhD, Senior Researcher
Affiliation	Research Institute of Sweden

E-mail address	Metkel.yebiyo@ri.se

Reference

[1]. Swedish Heat Pump Association (Svenska Kyl & Värmepumpföreningen, SKVP)