DOI: 10.1111/jiec.13528

JOURNAL OF INDUSTRIAL ECOLOGY WILEY

RESEARCH ARTICLE

Machine-learning-based demand forecasting against food waste

Life cycle environmental impacts and benefits of a bakery case study

Nicolas Hübner | Justus Caspers | Vlad Constantin Coroamă | Matthias Finkbeiner

Institute of Environmental Technology, Technische Universität Berlin, Berlin, Germany

Correspondence
Nicolas Hübner, Sekretariat Z1, Technische Universität Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany.
Email: nicolas.huebner@tu-berlin.de

Editor Managing Review: George Kamiya

Abstract

Rapid advancements in artificial intelligence (AI) are driving transformative changes in many areas, with significant environmental implications. Yet, environmental assessments for specific applications are scarce. This study presents an in-depth life cycle assessment of "Foodforecast," a machine learning (ML) cloud service designed to reduce food waste in bakeries by optimizing sales forecasting. It covers four impact categories: global warming, abiotic resource depletion, cumulative energy demand, and freshwater eutrophication. The assessment includes both the direct environmental impacts of the ML model and the underlying system hardware, as well as the indirect benefits of avoided bakery returns compared to traditional ordering methods, using real-world case study data. In 2022, "Foodforecast" led to an average 30% reduction in bakery returns, primarily bread and rolls, according to sales reports. The associated environmental benefits significantly outweighed the system's direct impacts by one to three orders of magnitude across impact categories and return utilization scenarios. The study identifies support activities such as service maintenance during deployment as major direct impact factors, surpassing those from cloud compute for ML operations. Data processing and inference dominate the latter, while the much-discussed ML training plays a minor role. The environmental consequences of AI are complex and dual sided. This case study demonstrates that AI might provide environmental benefits in certain contexts, yet results are constrained by methodological challenges and data uncertainties. There remains a need for further holistic LCAs across different ML applications to inform decision-making processes and ultimately guide the responsible use of AI.

KEYWORDS
artificial intelligence, bakery returns, demand forecasting, indirect effects, industrial ecology, life cycle assessment

1 | INTRODUCTION

Given the quick spread of artificial intelligence (AI) across society (Simon, 2019; Zhang et al., 2022), researchers and the public alike have expressed concerns about its environmental implications (Zhang & Dafoe, 2019). The direct—or first-order—impact of AI has been the subject of recent research and media attention, centering primarily on energy consumption and associated carbon emissions with the training of large-scale deep learning models (Strubell et al., 2019). The chase for AI accuracy comes at the cost of massive amounts of computing resources, data, and ever-larger models to diminishing accuracy gains. Consequently, the computing demand for the largest AI training runs has been doubling every 3.4 months since 2012 (Amodei & Hernandez, 2018). The widespread adoption of AI into business models and the mainstream use of AI chatbots, such as ChatGPT (OpenAI, 2023), can be expected to accelerate the growing energy demand of AI (IEA, 2022).

As a consequence, researchers started to include energy efficiency considerations in AI research, a concept known as "Green AI" (Schwartz et al., 2019). Calculator tools aim to support developers in estimating the "hidden" carbon emissions associated with machine learning (ML) systems (Lacoste et al., 2019). Due to their intricate nature, assessments require extensive data on factors such as computational demands, algorithm architecture, hardware utilization, frequency and duration of compute cycles, and the varying carbon intensity of the electrical grid (Koomey & Masanet, 2021; Lannelongue et al., 2021). Carbon footprint assessments often focus on the operational phase, disregarding the significant manufacturing footprint of computing hardware (Gupta et al., 2020). While some recent studies have taken a more holistic approach, considering both the life cycle of the model and its underlying equipment (Luccioni et al., 2022; Wu et al., 2021), they fall short in capturing impacts beyond carbon, such as on mineral and metal scarcity. Information and communication technologies (ICT) rely heavily on non-renewable materials, including critical raw materials such as gallium, silicon, and germanium, many of which face limited recycling (UBA, 2021; Wäger et al., 2015). Multi-criteria assessments can identify trade-offs and avoid misguided optimization strategies (Finkbeiner, 2009), with some recent work on the direct impacts of generative AI services (Berthelot et al., 2024).

ML's indirect environmental implications are diverse and significant, yet largely unexplored compared to economic or safety aspects, hindering large-scale applications of AI (Liao et al., 2022). As a transformative technology, ML can both mitigate and exacerbate environmental impacts, as recent studies highlight (Chen et al., 2023; Rolnick et al., 2022). For example, ML has the potential to reduce greenhouse gas (GHG) emissions through its immediate application (second-order effect) by optimizing energy production from renewable sources and enhancing smart grid management. Conversely, its application in fossil fuel exploration might increase emissions. Furthermore, higher-order effects, such as rebound effects, can arise when initial benefits lead to behavioral or systemic changes that counteract positive impacts, such as when improvements in fuel efficiency and traffic management by ML increase demand for individualized mobility (Fulton et al., 2017).

The food industry offers several opportunities for AI integration along the supply chain (Kumar et al., 2021), such as in enhancing waste reduction strategies in retail, where waste is concentrated in a few locations and accumulates impacts from upstream processes. Reducing food loss and waste, which account for one third of the food produced for human consumption and 8%–10% of anthropogenic GHG emissions (FAO, 2013; UNEP, 2021), aligns with the EU Waste Framework Directive's waste hierarchy priority on waste prevention (EC, 2024) and UN Sustainable Development Goal 12.3 (UN, 2015).

Fresh bakery products represent the largest waste fraction in European retail, mainly driven by uncertainties in customer demand (Stenmarck et al., 2011). In Germany, approximately 600,000 Mg (megagrams) of bakery products are wasted each year, with bakery return rates typically ranging from 1.5% to 19% (Jaeger, 2018). Reducing this waste could yield significant second-order benefits, and quantifying them is crucial to avoid neglecting AI's potential for the environment ("AI for Green").

Precise demand forecasting is essential but challenging due to daily fluctuations influenced by weather, seasonal changes, holidays, and local disruptions. The German startup "Foodforecast Technologies GmbH" offers a ML-based Software-as-a-Service (SaaS) "Foodforecast," aimed at German bakeries (Foodforecast Technologies GmbH, 2022). "Foodforecast" delivers demand-adjusted sales forecasts by analyzing historical sales data, thus replacing conventional order placement methods. The service automates the daily ordering process, aiming to reduce overproduction and early selloffs.

Claims highlighting AI's potential to enable environmental benefits through efficiency gains are rarely substantiated and may overestimate potential benefits (Rasoldier et al., 2022). Realistic insights into the environmental implications of AI in specific applications necessitate comparative, holistic life cycle assessment (LCA) approaches to balance AI's positive and negative effects (Pohl et al., 2019).

This paper aims to quantify and compare the environmental net impact of the implementation of "Foodforecast" in German bakeries, linking "Green AI" with "AI for Green." It thus represents one of the first case studies to quantify both first- and second-order effects of a specific AI application, covering the operation and manufacturing of system hardware in critical impact categories beyond climate change.

2 | METHODS

This section outlines the goal and scope of the study (Section 2.1), the collection of inventory data and their translation into a set of potential impacts (Section 2.2).

 JOURNAL OF INDUSTRIAL ECOLOGY WILEY

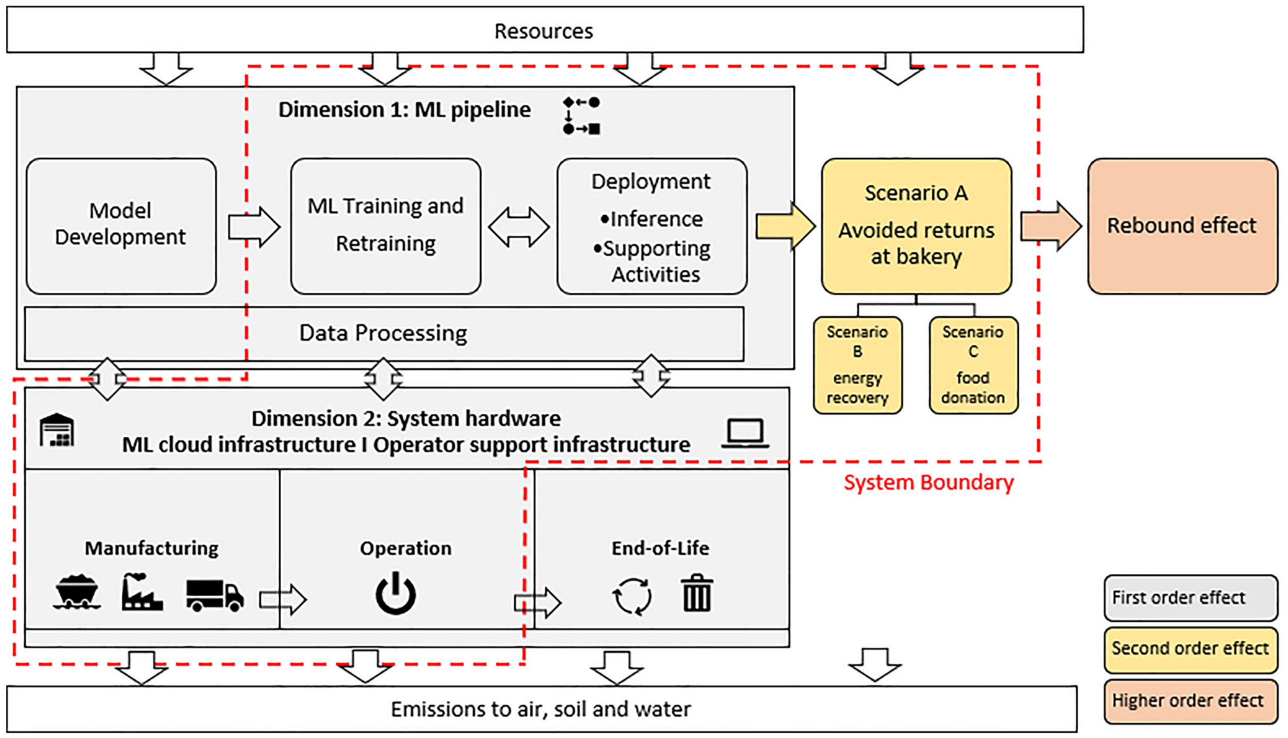

FIGURE 1 Scope of the multi-dimensional assessment.

2.1 | Goal and scope

LCA was applied to quantify the first-order, second-order, and resulting net second-order effect of "Foodforecast." LCA is a mature method to assess the environmental impacts of products and services, based on the ISO standards 14040 and 14044 (ISO, 2020a, 2020b) and more specific standards for ICT (ETSI, 2015; ITU-T, 2014).

The functional unit considered was the annual service use aggregated across "Foodforecast" customers in 2022. Processes that are estimated to account for less than 1% of the total environmental impact are subject to cut off.

The first-order assessment has a two-dimensional structure, including the higher-level service life cycle (dimension 1) and the lower-level system hardware life cycle (dimension 2). System boundaries are defined for both as summarized in Figure 1.

Dimension 1 covers the ML pipeline including data processing, model training, and deployment. The model development stage is cut off due to a negligible usage share of the widely available pre-trained recurrent neural network (RNN) used (Salinas et al., 2017). Model deployment involves inference, that is, the computation of sales forecasts based on data input, and supporting activities by the service operator that are integral to the effective deployment of the service, such as communications, accounting, maintenance, marketing, and others. Dimension 2 is divided into two categories: the ML cloud infrastructure dedicated to ML-related operations (data processing, model training, and inference), encompassing servers, storage disks, network devices, and data center facility systems, and the operator support infrastructure, comprising on-premises workstation ICT utilized by "Foodforecast" team members in home office settings. For both categories, this includes hardware operation (characterized by power consumption) and the upstream manufacturing processes of raw material extraction, production, and distribution. The end-of-life (EoL) stage is excluded due to uncertainties surrounding collection and recycling rates, which would require highly speculative estimates. Manufacturing of network equipment and data center facilities as well as direct emissions from data center cooling systems and backup diesel generators are excluded, based on marginal ratios of individual use, such as service traffic to total lifetime traffic.

The second-order assessment focuses on the change in bakery returns induced by the adoption of ML-powered demand forecasting compared to baseline returns from the previous conventional ordering process. Bakery returns in Germany are disposed or valorized through various means, including biogas plants and food banks (Jaeger, 2018). Three scenarios are evaluated, considering the life cycle of returns from cradle-to-gate (Scenario A), and for two EoL scenarios, including energy recovery through anaerobic digestion (Scenario B), and food donation to charitable organizations (Scenario C), as detailed in Supporting Information S1.

Due to their complex interactions and large uncertainties, higher-order effects at the systemic level are difficult to quantify within the scope of this single case study, so they are excluded with a focus on more measurable effects.

TABLE 1 Cloud usage categories and derived power coefficients.

Cloud usage		Power coefficient	Unit	Source
Storage	SSD	1.2	Wh/TBh	Adapted from extrapolated average industry disk wattage and capacity in 2020 from United States. Data Center Energy Usage Report (Shehabi et al., 2016)
	HDD	0.65	Wh/TBh	
Networking	Internal and external data transfer	1.64	Wh/GB	UTAMO model applied for VDSL data traffic at 1 GB/h, based on 2020 technology generation in Germany (IZM, 2021)
Compute	Server	Variable	W	Statistical approximation of power consumption based on CPU microarchitecture and utilization and comparable SpecPower_ssj2008 benchmarks (SPEC, 2023)
	Memory deviations	0.3915	W/GB	Memory size deviating from matched SpecPower_ssj2008 benchmarks is considered using average DDR4 SDRAM power draw provided by manufacturers (Crucial, 2023; Micron, 2017)

2.2 | Life cycle inventory and impact assessment

This section describes the procedure to compile a life cycle inventory (LCI) of the inputs and outputs related to the service use. Subsequently, the collected LCI data is converted into potential environmental impacts in the life cycle impact assessment (LCIA). Four impact categories were selected for their relevance to the studied system and data availability: Global warming (GW) over 100 years (kg CO_2 eq.), abiotic resource depletion, minerals and metals (AD) (kg Sb eq.), cumulative energy demand (CED) (MJ), and freshwater eutrophication (FE) (kg P eq.). These categories assess the contribution to climate change, resource scarcity, and the energy footprint, which are prevalent impacts in ICT. FE was added due to the application of the AI in the food sector, which predominantly contributes to global eutrophication through agricultural nutrient runoff (Poore & Nemecek, 2018). The Environmental Footprint 3.1 method was chosen for its coverage of these categories and alignment with the European Commission's recommendations (Pant & Zampori, 2019).

2.2.1 | ML pipeline

The SaaS uses dedicated services and platforms from two cloud providers, primarily Amazon Web Services (AWS) and to a lesser extent Microsoft Azure, to manage virtual machines (VMs), data transfer, and storage for ML operations. The respective data centers are located in Frankfurt, Germany and Amsterdam, the Netherlands (AWS, 2023; Azure, 2023). Connectivity between clients, operator workstations, and data centers relies on the already available Internet infrastructure. The cloud-based setup seamlessly integrates with existing enterprise resource planning (ERP) system interfaces of bakeries, eliminating the need for additional client-side infrastructure. The model is trained individually for each bakery, learning the complex relationships between historical sales and external factors such as holidays and weather conditions. Based on this, the ML algorithm computes demand-adjusted sales forecasts for each product and store. Data processing renders the data suitable for both training and inference, including tasks such as data preparation, transformation, and feature engineering.

The complexity of cloud computing architectures and AI usage patterns, coupled with limited information from cloud providers, required an extensive data collection process to approximate the environmental impact of workloads from multiple sources.

Operational power consumption

Operational power consumption was estimated using a procedure adapted from the open-source Cloud Carbon Footprint tool (CCF, 2023) and research published by Teads Engineering (Davy, 2021), which involves collecting cloud usage data and then multiplying it by energy coefficients. The data basis for the procedure is summarized in Supporting Information S1.

A cost and usage report (CUR) for "Foodforecast" was generated from AWS, detailing incurred usage for compute (runtime hours on specific VMs), networking (gigabytes [GB] of data transferred), and storage (GB-months of data stored). The model, along with raw and preprocessed data for training and inference, and outputs, are stored and managed across a mix of block-level, relational database, and object storage services using both HDDs and SSDs. Storage usage was adjusted for redundancy with replication factors of 2 or 3, as specified for each storage service (AWS, 2023). Energy coefficients are derived from industry reports and benchmarks, as summarized in Table 1.

JOURNAL OF INDUSTRIAL ECOLOGY WILEY

Given the wide diversity of server configurations, it was not feasible to rely on a single power coefficient for compute. Hence, an inventory of the various VMs with their distinct configurations running the SaaS was compiled. Cloud vendor documentation was consulted for performance-related specifications on attributes such as CPU type and cores, number of virtual CPUs (vCPUs), memory allocated, storage disc type, and capacity (AWS, 2023; Azure, 2023). Specifications of custom-built CPUs were derived from CPU-World (2023) and WikiChip (2023), as these details were not readily available in manufacturers' databases.

VMs represent a virtualized fraction of a server, and their size typically scales linearly based on vCPU ratios within a given family, with the largest size (maxVM) being equivalent to an entire physical server (AWS, 2023). Therefore, data was collected for maxVMs and then allocated to individual VMs based on vCPU ratios without overcommitting resources.

MaxVMs are configured without a Graphical Processing Unit due to the sequential nature of the RNN and have minimal internal storage, relying on external block-level storage. Therefore, their power consumption can be referenced from the SPECpower_ssj2008 (SPEC, 2023), hereafter SPECpower. This database provides power profiles of streamlined servers against a workload curve that is representative of real-world applications. The power consumption of servers is approximated from benchmarks with comparable configurations based on three factors: CPU type and number of cores (1), utilization (2), and memory capacity (3):

1. First, maxVMs are matched to benchmark systems with a similar CPU microarchitecture. The average matched benchmark power consumption per core is then scaled by the corresponding maxVM core count.
2. Second, the average utilization of the VMs in 2022 was determined based on the payment model and was aligned with SPECpower load levels applying linear interpolation. For VMs under on-demand pricing, utilization was set to 50%, reflecting projections for the average hyperscale utilization in 2020 (Shehabi et al., 2016). Utilization of burstable instance types was based on their baseline values from AWS documentation (AWS, 2023). For reserved instances, the average utilization was derived from early 2023 observations in the monitoring service AWS CloudWatch in the absence of logs from 2022 for "Foodforecast."
3. The memory per chip ratio of SPECpower systems is typically lower than that of the VMs, which is considered using the average power draw of memory (see Table 1) for additional memory.

As demonstrated in Equation (1), the resulting memory-adjusted maxVM power draw at the corresponding utilization ($P_{\mathrm{maxVM,U}}$) is multiplied by the corresponding VM's vCPU ratio and the runtime hours ($t_{\mathrm{VM,MLOp}}$) for a given ML operation specified in the CUR, either data processing, training, or inference.

$$E_{\mathrm{VM,MLOp}} = \left(P_{\mathrm{maxVM,U}} \times \frac{\mathrm{vCPU}_{\mathrm{VM}}}{\mathrm{vCPU}_{\mathrm{maxVM}}} \right) \times t_{\mathrm{VM,MLOp}} \tag{1}$$

To calculate the total operational power consumption for the ML cloud infrastructure (E_{MLCloud}), the individual power consumptions for computing ML operations ($E_{\mathrm{VM,MLOp}}$), storage (E_{S}) and networking (E_{N}) are aggregated as IT power consumption. This sum is then multiplied by the providers' power usage effectiveness (PUE)—1.15 for AWS (Sehgal & McDonnell, 2020) and 1.19 for Azure (Walsh, 2022)—to account for the overhead energy consumption of facilities, as summarized in Equation (2).

$$E_{\mathrm{MLCloud}} = \left(\sum E_{\mathrm{VM,MLOp}} + E_{\mathrm{S}} + E_{\mathrm{N}} \right) \times \mathrm{PUE}_{\mathrm{AWS}} + E_{\mathrm{IT,AWS}} \times \mathrm{CAF} \times \mathrm{PUE}_{\mathrm{Azure}} \tag{2}$$

Detailed usage data from Azure resources was not accessible for this research. Consequently, an alternative cost-based estimation of Azure power consumption was undertaken, assuming that electricity consumption is the main cost driver for hyperscale data center operators (Hamilton, 2008). It involved scaling IT power consumption at AWS ($E_{\mathrm{IT,AWS}}$) using a cost-based adjustment factor (CAF) derived from the distribution of the cloud service costs between AWS and Azure, as detailed in Supporting Information S1. Despite its uncertainties, this approach was deemed the most suitable option in lack of more specific data.

Support activity data was collected through a survey at "Foodforecast Technologies GmbH," recording the number of full-time equivalent (FTE) employees and the respective workstation devices, which included a laptop (Apple MacBook Pro 2021 or equivalent), router, and 24-in. liquid-crystal display (LCD) monitor. Specific employee usage profiles were not created for privacy concerns. Instead, representative usage patterns, data transfer, and resulting power consumption of workplace devices were taken from the "Green Cloud Computing" (GCC) project of the German Federal Environment Agency (UBA, 2021), considering 1760 FTE work hours and standby power consumption. However, routers are used by employees in home office privately outside of work hours and this consumption was not included.

Operational power consumption for ML operations and support activities was characterized using local electricity grid mix factors for the data centers in the Netherlands and Germany, sourced from Sphera's LCA database (content version 2023) (Sphera, 2023).

Manufacturing

The manufacturing impacts of cloud compute and storage usage are initially determined by applying LCIA indicators for server components and storage disks—for GW, AD, and CED from project GCC (UBA, 2021) and for FE from Peiró and Ardente (2015)—to the maxVM configurations and occupied external storage disks. These impacts are then allocated based on virtualized usage ratios.

Integrated circuits (ICs) significantly impact the environmental footprint of servers, directly related to their die area (Boyd, 2012; Teehan & Kandlikar, 2013). The die areas of RAM and SSD are calculated by dividing their capacities by the latest generation's average storage density. The CPU die area is calculated from the number of cores and the average die area per core specific to the CPU family, using data from the open-source project Boavizta (2021). CPU impact indicators are based on 300 mm wafers data from two major suppliers, Siltron and Sumco, and on Intel Xeon processors using 14 nm technology (UBA, 2021). These data are representative for the majority of utilized VM CPUs (see Supporting Information S1). The GCC indicators for the ICs were scaled based on die size, while also considering constant baseline contributions from the CPU heat sink and gold contacts. The EF indicators, however, do not specify a reference die size and constant factors were scaled based on the number of cores, RAM modules, and SSD disks of the maxVM configurations.

For HDD storage, motherboard, power supply units, assembly and remaining server components, including enclosure, cables, and fans, constant overhead impact factors based on the typical bill of materials for enterprise rack servers from the referenced reports were used, which may overestimate impacts for smaller maxVM configurations. Equation (3) summarizes the calculation of maxVM manufacturing impacts (I_{maxVM}) based on these LCIA factors (denoted as I), which are broken down in Supporting Information S1.

$$I_{\mathrm{maxVM}} = I_{\mathrm{CPU}} + I_{\mathrm{RAM}} + I_{\mathrm{SSD}} + I_{\mathrm{HDD}} + I_{\mathrm{Mainboard}} + I_{\mathrm{PSU}} + I_{\mathrm{RackRest}} + I_{\mathrm{Assembly}} \tag{3}$$

Specific EPDs, where available, were used to model the manufacturing of support activity hardware (Acer Inc., 2022; Apple Inc., 2021; EIZO Corporation, 2020), but data beyond GW was limited. Therefore, this dataset was complemented with generic data from ecoinvent 3.8 (Wernet et al., 2016) for the monitor and by treating the laptop as a blade server, calculating the indicators in a proxy approach as for the maxVMs.

The dynamic sharing of virtualized cloud resources under pay-as-you-go models requires a precise allocation of manufacturing impacts based on the individual usage ratios of "Foodforecast." Thus, the total embodied impacts of the server ($I_{\mathrm{maxVM,\,total}}$) are allocated to VMs and specific ML operations ($I_{\mathrm{VM,MLOp}}$) by multiplying the VM vCPU ratio, the ratio of the actual VM runtime for the ML operation ($t_{\mathrm{VM,MLOp}}$) to the total annual runtime, and the reciprocal of an expected 5-year lifetime (Amazon, 2022b).

$$I_{\mathrm{VM,MLOp}} = I_{\mathrm{maxVM,\,total}} \times \frac{\mathrm{vCPU_{VM}}}{\mathrm{vCPU_{maxVM}}} \times \frac{t_{\mathrm{VM,MLOp}}}{8760\mathrm{h}} \times \frac{1\ \mathrm{year}}{5\ \mathrm{year}} \tag{4}$$

The runtime term is omitted for resources reserved exclusively for "Foodforecast" throughout 2022. Similarly, allocation for external storage utilizes occupation ratio of storage disks, replacing runtime and vCPU ratio. The lifetime embodied impacts of support activity hardware from EPDs are allocated by spreading linearly over an expected 5-year lifetime.

2.2.2 | Bakery returns

The ML-based sales forecasting service changes the order and production quantities of client bakeries. Quantifying the induced change in bakery returns at product level requires extensive evaluation of ERP data and the accurate establishment of baseline returns under the replaced ordering process, which is no longer operational. The baseline definition faced with a hypothetical counterfactual is a well-known issue in the assessment of indirect impacts of ICT (Bieser & Hilty, 2018; Bremer et al., 2023; Coroamă et al., 2020). This baseline issue was addressed using two approaches:

Baseline Approach 1 (B1) sets historical pre-service returns as the baseline, using aggregated data collected by the service operator. However, B1 was not sufficient to determine the actual environmental effect of the changed returns for two reasons: it was not possible to disaggregate the results to the product level, which is crucial given the diverse range of bakery products, and it was not possible to account for confounding factors between various years, such as those caused by the Covid-19 pandemic or inflation.

Acknowledging these ontological uncertainties in B1, Baseline Approach 2 (B2) leaves the bakery ordering process unchanged by the SaaS to establish the baseline, while simulating the ML-based recommendations in parallel for the same stores to assess the alternative scenario. Here, actual returns represent the baseline, while virtual (simulated) returns represent the ML-based intervention. The crucial difference is that in this paradigm, virtual returns are perfectly deterministic. After an initial calibration phase, actual and simulated returns for a total of 52 products were compared in a medium-sized bakery chain in Germany, comprising 18 stores, a central administration, and a production facility. This analysis was conducted over a 4-week period from February 12 to March 12, 2023. Products were grouped into four representative clusters based on common attributes such as ingredients, baking process, and dough preparation:

TABLE 2 Life cycle impacts per kilogram return reference.

LCIA category	Unit	Scenario A	Scenario B	Scenario C
GW	[kg CO_2 eq.]	1.16	0.65	0.15
AD	[kg Sb eq.]	5.09E-06	5.01E-06	1.10E-06
CED	[MJ]	40.02	24.43	5.46
FE	[kg P eq.]	2.52E-04	2.50E-04	5.49E-05

1. Bread and rolls ($N = 31$)
2. Cakes and fine-yeast products ($N = 13$)
3. Croissants ($N = 5$)
4. Pretzels ($N = 3$)

"*N*" denotes the number of products considered within each cluster. For instance, the "Pretzels" cluster encompasses three products like Pretzel buns and sticks. The actual returns and virtual returns observed during B2 were aggregated across the clusters. The avoided returns were determined by subtracting actual returns from virtual returns.

Product clusters were modeled using a bottom-up approach by assigning and averaging cradle-to-gate indicators of matching LCA datasets for bakery products from the comprehensive Agribalyse 3.1 database (Colomb et al., 2015). For GW, the results range from 0.84 kg CO_2 eq. for bread and rolls to 3.00 kg CO_2 eq. for cakes and pastries, influenced by factors such as ingredients and processing steps. The weighted average of the cluster results based on the composition of the actual returns forms the cradle-to-gate impact of the reference return under Scenario A.

The EoL stages under Scenarios B (energy recovery) and C (food donation) were modeled using Sphera's Managed LCA Content (version 2023) (Sphera, 2023) and ecoinvent 3.8 (Wernet et al., 2016) database, based on parameters such as the average methane yield of bakery waste, transport distances, or the proportion of donated waste consumed (both from the B2 case study and literature, details in Supporting Information S1). Table 2 summarizes the derived impact factors per kilogram return reference by scenario.

3 | RESULTS

This section presents the results for the ML pipeline (see Section 3.1), the bakery returns (see Section 3.2), and the comparative assessment (see Section 3.3)

3.1 | ML pipeline

Figure 2 shows the direct impacts of the ML pipeline, which amount to a GW of 1656 kg CO_2 eq., AD of 2.05E-01 kg Sb eq., CED of 40,411 MJ, and FE of 3.63E-01 kg P eq. Operator support activities, primarily related to laptops and monitors, are the main contributors with shares ranging from 53% for CED and GW to 87% for AD and 97% for FE. Total networking accounts for 0%–8% of the direct impacts. Among the ML operations, cloud computing contributes 78%–93%, while data center facilities and data storage each contribute small shares ranging from 0% to 12%.

Figure 3 illustrates the contributions of life cycle phases to the direct impacts of operator support activities and ML operations. For the former, manufacturing plays a dominant role in all categories, accounting for 66% of GW, ~100% of AD, 61% of CED, and 99% of FE. Conversely, for ML operations, ICT operations are the main contributor to GW (79%) and CED (88%), while manufacturing remains the main contributor to AD (~100%) and FE (75%).

Of the cloud compute-related effects of ML operations, around two thirds are attributable to data processing and around 30% to ML inference, as detailed in Supporting Information S1. ML training accounts for the remaining 3% to 6%.

3.2 | Bakery returns

In 2022, an aggregate reduction of ~2000 tons in bakery returns across 175 bakeries was derived from B1, with an average 30% reduction in return weight. For B2, actual returns were 14,169 kg, with a 19.6% return rate. The simulation resulted in 10,251 kg virtual returns, reducing returns by 3917 kg or 27.7%, provisionally confirming B1's results. This implies a potential weekly reduction of 54 kg returns per bakery, lowering the return rate to 16.9%. Figure 4 illustrates the breakdown of actual, virtual, and avoided returns observed in B2.

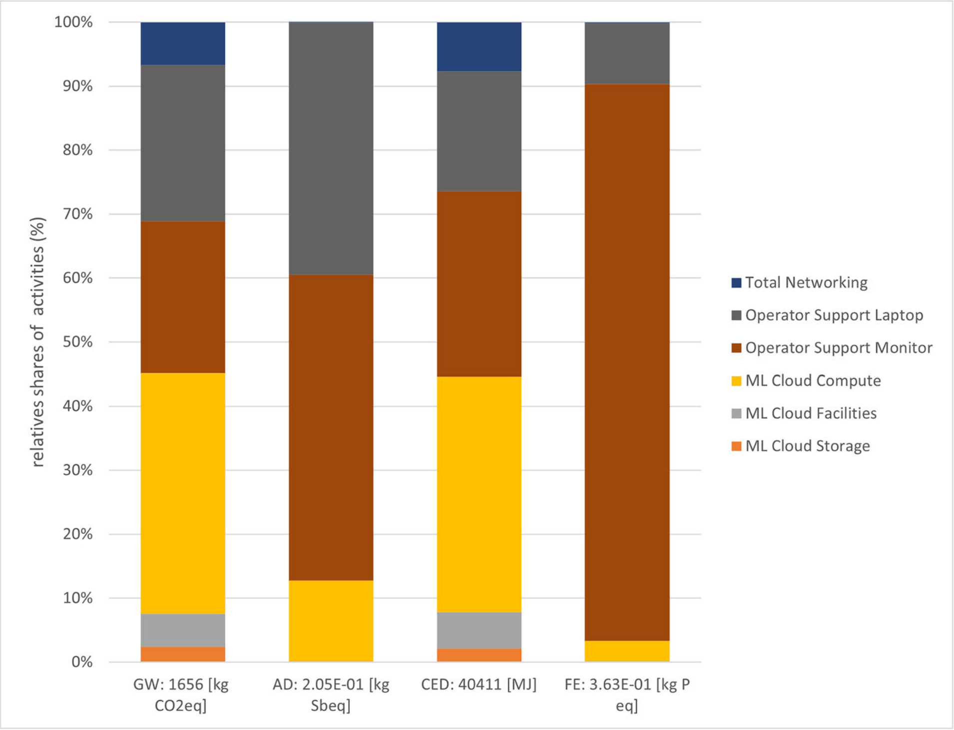

FIGURE 2 Direct machine learning pipeline results and relative contributions of activities, detailed data are provided in Supporting Information S1.

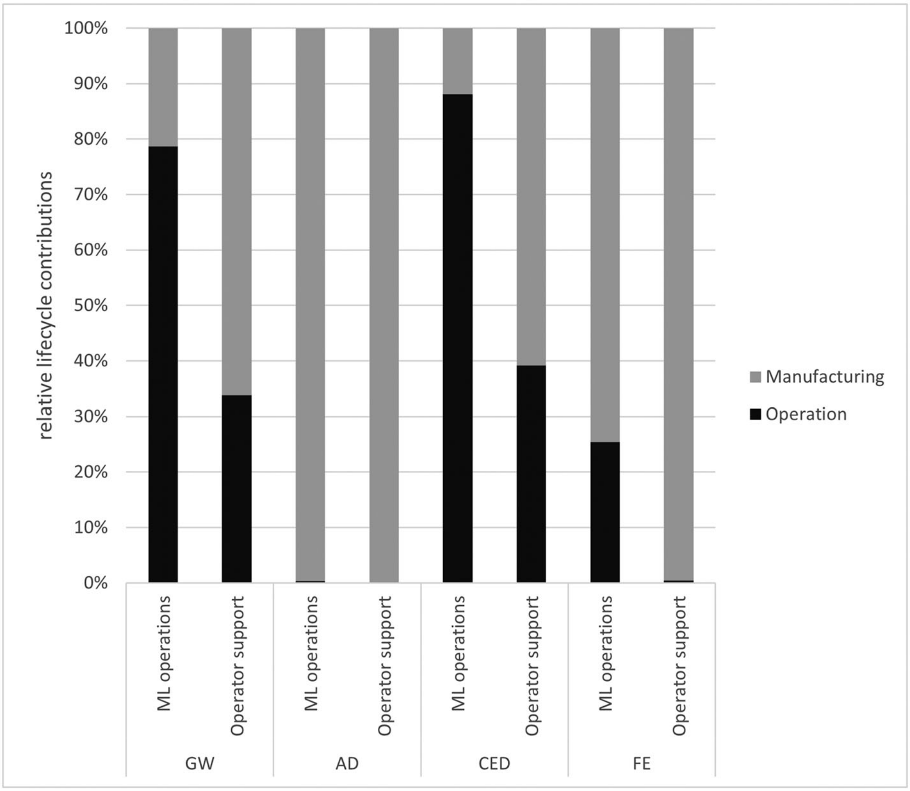

FIGURE 3 Life cycle contributions to machine learning operations and operator support activities, detailed data are provided in Supporting Information S1.

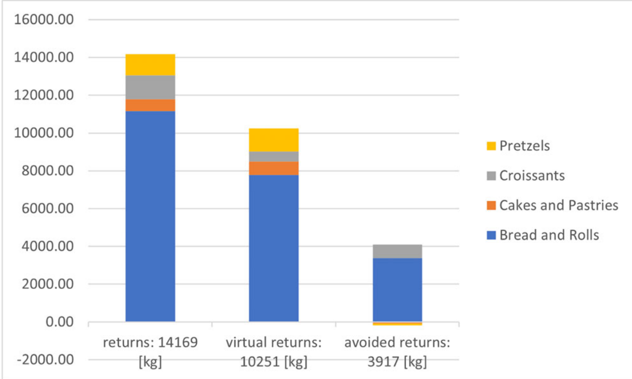

FIGURE 4 Observed return values of Baseline Approach 2 during the case study, detailed data are provided in Supporting Information S1.

TABLE 3 Aggregated second-order impacts of avoided bakery returns.

LCIA category	Scenario A (cradle-to-gate)	Scenario B (energy recovery)	Scenario C (food donation)
GW [kg CO_2 eq.]	−2.33E+06	−1.29E+06	−2.95E+05
AD [kg Sb eq.]	−1.02E+01	−1.00E+01	−2.20E+00
CED [MJ]	−8.00E+07	−4.89E+07	−1.09E+07
FE [kg P eq.]	−5.05E+02	−4.99E+02	−1.10E+02

The largest reductions were in bread and rolls (3381 kg) and croissants (718 kg), while increased overproduction is observed in the simulation for pretzels (110 kg) and cakes and pastries (72 kg). The integration of B1 and B2 establishes a robust foundation to determine the second-order benefits. Specifically, we use the total mass of aggregated avoided returns from B1, which is supported by a similar magnitude of relative reduction observed in B2. We then apply the impact indicators of the reference return to this data, based on the composition of returns analyzed in detail during the B2 case study. Table 3 reveals resulting significant avoided environmental (second-order) impacts in 2022 across all categories and scenarios, denoted by negative values.

Scenario A demonstrates the greatest absolute amount of avoided environmental impacts. In Scenario B, which includes energy recovery, avoided impacts are reduced by 45% in GW, 2% in AD, 39% in CED, and 1% in FE. Scenario C, which integrates the donation of returns, shows more pronounced reductions in avoided impacts: 87% in GW, 78% in AD, 86% in CED, and 78% in FE.

3.3 | Comparative assessment

Table 4 reports the aggregated net second-order results of the comparative assessment, which are consistently negative across all categories and scenarios, indicating that the environmental benefit resulting from avoided bakery returns outweighs the direct impacts associated with the ML pipeline. The magnitude of this second-order benefit (impact ratio) varies significantly across categories and scenarios, ranging from 11 times greater for AD in Scenario C to 1961 times greater for CED in Scenario A than the first-order impact. The gap for AD is consistently smaller than for the other indicators.

Break-even points were calculated to determine the amount of bakery returns that compensate the first-order impacts of the ML pipeline, ranging from 1 ton for CED in Scenario A to 186 tons for AD in Scenario C of avoided bakery returns. Again, the AD results stand out as they are consistently an order of magnitude higher than the other categories.

In comparison to Scenario A, Scenario B shows net impact reductions of 45% in GW, 39% in CED, 2% in AD and 1% in FE. Scenario C shows even greater reductions: 87% for GW, 86% for CED, 80% for AD, and 78% for FE.

4 | DISCUSSION

This section critically examines the results (see Section 4.1) by discussing various limitations and challenges (see Section 4.2).

TABLE 4 Net impact results of the comparative assessment.

LCIA category		Unit	Scenario A	Scenario B	Scenario C
GW	Net impact	kg CO_2 eq.	−2.33E+06	−1.29E+06	−2.93E+05
	Impact ratio		1381	767	175
	Break-even point	Tons of avoided returns	1.42	2.56	11.24
AD	Net impact	kg Sb eq.	−9.98E+00	−9.81E+00	−2.00E+00
	Impact ratio		49	48	11
	Break-even point	Tons of avoided returns	40.21	40.90	185.90
CED	Net impact	MJ	−8.00E+07	−4.88E+07	−1.09E+07
	Impact ratio		1961	1197	267
	Break-even point	Tons of avoided returns	1.01	1.60	7.41
FE	Net impact	kg P eq.	−5.04E+02	−4.99E+02	−1.09E+02
	Impact ratio		1384	1369	301
	Break-even point	Tons of avoided returns	1.44	1.45	6.61

4.1 | General discussion of the results

Previous research has assessed the environmental benefits of AI addressing food waste. Principato et al. (2023) used LCA to compare food waste generated in workplace canteens before and after the introduction of an AI-based digital tool but neglected the direct impact of AI itself. Their findings showed a significant reduction in both food waste and GHG emissions. Our case study advances this field by demonstrating that the implementation of ML-based demand forecasting can significantly reduce overproduction in bakeries. The resulting environmental benefits outweigh the direct impacts of the ML pipeline by one to three orders of magnitude.

The AD has the latest break-even points within impact categories, as the ML pipeline relies heavily on plastics, metals, and various critical raw materials for the production of system hardware (Wäger et al., 2015). Conversely, abiotic resource consumption is less pronounced in the life cycle of bakery returns, resulting rather indirectly from production of mineral fertilizers and energy. The major impacts of food waste are GHG emissions from soil processes and livestock farming, the energy use in agriculture and processing, and eutrophication from fertilizer application (Scherhaufer et al., 2018). Consequently, avoiding bakery returns results in a higher relative environmental benefit for the GW, CED, and FE.

The scenario analysis shows that the reduction of environmental impacts achieved by “Foodforecast” is closely linked to the scenario chosen for the handling of returns. Unsurprisingly, EoL valorization within the Scenarios B and C leaves less optimization potential and thus yields less environmental benefits when introducing the ML-based system. The two, however, are not mutually exclusive: Even with a return-minimizing ML system in place, EoL valorization should still be applied to unavoidable returns to minimize environmental impacts.

As shown in Figure 2, most direct environmental impacts originate from support activities rather than the primary ML-related operations. This perhaps counterintuitive finding is attributed to the efficient design of cloud infrastructure in hyperscale data centers. Scaling AI usage across bakery industry could potentially shift this balance by increasing the demand for cloud resources. The ML operations capitalize from maximized server utilization, specialized chips designed for ML tasks and advanced cooling systems, as evidenced by PUEs below 1.2. Furthermore, the share of embodied GHG emissions is often higher for decentralized computing compared to cloud computing (Malmodin & Lundén, 2018), as cloud platforms can dynamically allocate resources on demand, ensuring maximized utilization. By contrast, the operator support infrastructure remains idle during downtimes, resulting in an embodied emissions share of 66% for GW compared to 22% for ML operations, as shown in Figure 3. Prior work reports a comparable distribution, finding that manufacturing contributes about 30% of the carbon footprint of large-scale ML compute tasks in Facebook (Wu et al., 2021).

Current research primarily focuses on the operational carbon footprint of AI, often overlooking the considerable contribution of manufacturing, as found in this study. As data centers become more efficient and transition to renewable energy, the relative importance of manufacturing emissions is expected to increase (IEA, 2022). In addition, ML relies on finite and scarce resources, which are particularly depleted in hardware manufacturing as highlighted by the AD in Figure 3. This leads to conflicting environmental objectives concerning the optimal timing of hardware replacement: while newer, more efficient hardware can reduce operational emissions, extending lifespans conserves finite resources. This underscores the importance of including manufacturing and multiple impacts in environmental assessments of AI.

In the context of “Green AI” and “AI for Green,” prioritizing forecasting accuracy in “Foodforecast” is expected to yield greater environmental benefits than improving energy efficiency, due to significant second-order effects. However, potential trade-offs must be considered. Bakery returns are minimized when forecasts are set lower than actual sales. Therefore, although configuring “Foodforecast” to be more tolerant of shortages could reduce waste, it must be balanced against the risk of increased out-of-stocks and lost sales opportunities to ensure that the AI remains effective for its intended purpose.

However, measures to enhance model efficiency without compromising accuracy should not be neglected, as pursuing higher accuracy reaches a point of diminishing returns. Google's ML practices, which involve selecting efficient model architectures such as sparse models, have demonstrated the potential to reduce computation by 5–10 times (Patterson et al., 2022). The present study's detailed breakdown of results, down to the contributions of individual ML operations and bakery product clusters, enables the identification of areas for improvement in both dimensions, efficiency, and accuracy. For instance, our study indicates the significant impact of operator support activities and emphasizes the dominant share of bread and rolls in the returns of bakeries. In the ML pipeline, inference and data processing were identified as environmental hotspots. Both are executed at high frequency, unlike model training which contributes less than 10% across all impact categories. This distribution coincides with recent assessments of ML models in Facebook and Google (Park et al., 2018; Patterson et al., 2022; Wu et al., 2021).

Thus, reducing the computational requirements of each inference and data processing cycle may significantly reduce the direct environmental impact. This may be achieved by fine-tuning the pre-trained RNN used for "Foodforecast," which is not fully optimized for the specific use case and dataset, through measures such as reducing neurons and unused features, coupled with efficient data usage through techniques such as data sampling.

4.2 | Discussion of data-related issues, uncertainties, and limitations of the LCA

Major challenges und limitations of this study are methodological: assessing the enabling impact of ICT solutions requires comparison with a (speculative) hypothetical scenario without the solution. Different baseline approaches have been proposed (Coroamă et al., 2020) with this study combining an ex post (B1) and a simultaneous baseline (B2). B1 is limited in its temporal representativeness, as it assumes static returns relative to historical data, while B2's scope is restricted as it only considers a snapshot of AI's application. For instance, the amount and composition of (avoided) returns in a bakery might differ from the chosen reference scenarios influenced by demographics, and product assortment and there is currently no data available to validate them. Nonetheless, when applicable, such as during testing phases, B2 provides a valuable deterministic blueprint for environmental assessments of ICT services due to the unmodified, simultaneous nature of the reference activity, which removes some uncertainties inherent in purely hypothetical reference scenarios.

Furthermore, LCA is not well suited to account for higher-order effects, and in particular rebound effects, which are crucial to comprehensively understand the net environmental impacts of ICT services (Coroamă & Mattern, 2019; Horner et al., 2016; Kaack et al., 2021; Lange et al., 2020). These effects are complex and generally underexplored (Bieser & Hilty, 2018), and are discussed here briefly. "Foodforecast" may increase the market competitiveness of bakeries by preventing overproduction and early sellouts. However, the overall consumption of bakery products in a society is likely to remain stable, as additional sales substitute products from other market participants—it represents an example for market saturation, where direct rebounds are unlikely (Coroamă & Mattern, 2019). Reduced staff time through automated ordering and the potential reinvestment of economic savings may also influence environmental impacts, triggering indirect rebound effects. However, not only are indirect rebounds notoriously difficult to grasp (Widdicks et al., 2023), they are likely limited due to the marginal income effects impact of bakery savings. This case study could thus be similar to the automated gas leakage discovery in von Fischer et al. (2017), which brings about important environmental benefits but only marginal financial ones, thus not triggering rebound mechanisms (Coroamă & Höjer, 2016).

Further limitations are related to epistemic and aleatoric uncertainties LCA input data, related to the complex, proprietary, and opaque nature of cloud-based ML services. While cloud usage data was obtained directly through CURs, generic data from a variety of industry reports and databases was used to model the impact of ML cloud workloads with little information from cloud providers. Mixing sources may have introduced inconsistencies (Pauer et al., 2020).

Server power consumption was estimated using the SPECpower database, which offers a more meaningful estimate for real-world workloads than thermal design power (TDP) of the CPU, commonly used as a proxy. Nevertheless, the representativeness of this approach is limited, especially for custom hardware. In some instances, the actual server power consumption can significantly exceed SPECpower benchmarks (Davy, 2021; Peiró & Ardente, 2015). The Server Efficiency Rating Tool may provide a suitable alternative for future assessments (SPEC, 2022).

The GW characterization of operational power consumption was based on the carbon intensity of the local electricity grid supplies (location-based accounting) (GHGP, 2015). Given that electricity is a major expense for data centers, investing in renewable energy sources (RES) can lead to cost savings and progress toward sustainability goals. Consequently, both AWS and Azure invest heavily in RES via power purchase agreements, reporting a share of 95% RES in their operations (Amazon, 2022a; Microsoft, 2018).

The operational GW of the ML cloud infrastructure was reevaluated using virtual electricity mixes reflecting a 95% RES share (market-based approach) with data from Sphera (Sphera, 2023). This adjustment led to a 62% decrease in GW for ML operations and a 29% reduction across the ML pipeline, as detailed in S1. Nevertheless, the impact on the net second-order GW was negligible (below 0.17%). Therefore, data-related uncertainties in the assessment of the ML pipeline operation may not significantly affect the overall findings of this study.

Furthermore, the present study has underscored the considerable contribution of manufacturing to the direct impacts of the ML service. Manufacturing data for (custom-made) electronic components and subassemblies is scarce, often poorly documented, limited to GW, and quickly outdated. EoL data for data center equipment is even scarcer and the selection of disposal companies is typically driven by compliance and distance over material recovery rates (Peiró & Ardente, 2015). Prior studies suggest that the EoL GW of servers is negligible (Boyd, 2012; Schödwell et al., 2018). However, this may not apply for the AD, where material recovery potentially offsets up to 60% of the AD from manufacturing (Peiró & Ardente, 2015). In 2022, of the 62 billion kg of e-waste generated globally, only 22.3% was properly collected and recycled with only a small fraction of valuable metals such as gold and copper being recovered (Baldé et al., 2024). Hazardous substances in this waste pose environmental and health risks when they leach from landfills or are released through informal recycling practices in lower-income countries. To conclude, further work is needed to assess the EoL stage, coupled with increased transparency by data center operators and ICT manufacturers to support holistic LCAs of ICT services.

Data availability and methodological challenges remain critical issues for the present study. Nevertheless, the relative magnitude of first- and second-order effects suggests that, despite uncertainty about the absolute extent, ML-based demand forecasting in bakery ordering processes is highly probable to yield substantial environmental benefits.

5 | CONCLUSION

We presented one of the first in-depth, multi-criteria LCAs of ML in a real-world application, comparing the first- and second-order effect, spanning the ML pipeline, system hardware, and the life cycle of avoided bakery returns.

This provides valuable insights into the practical application of LCA to ML-based services and reveals optimization potentials and potential trade-offs. The findings support the claim that ML can enable environmental benefits in specific contexts, as demonstrated by ML-based demand forecasting in bakeries. This application was shown to effectively reduce the amount of returns and their associated environmental impacts, with the benefits significantly outweighing the direct impacts of ML within the constraints of this study. This underscores both the need and the potential for sustainable solutions in the food industry, particularly for the pressing issue of food waste.

Data collection for cloud-based ML services is challenging due to the intricate and opaque nature of the cloud architectures and requires sensible approaches to bridge data gaps. Methodological challenges include the definition of an appropriate baseline, precise allocation of shared cloud resources based on the individual use ratio, and sophisticated setting of system boundaries. This study shows that holistic environmental assessments should go beyond GW and cover all dimensions of AI services, including data processing, ML training, inference, and operator support activities, as well as the entire life cycle of system hardware. Uncertainties remain and it is important to consider these limitations when interpreting the results of this study.

Despite AI's transformative potential, its environmental implications are not understood enough. The current focus on carbon emissions in company reporting, research, and media undermines the various other impacts involved with ML, such as the depletion of finite resources, as shown in this study. Holistic multi-criteria LCAs can uncover avenues for energy reduction, resource optimization, and "Green AI" system designs while quantifying the potentials of "AI for Green." As such, they are crucial to guiding AI development toward sustainability, avoiding exacerbation of global environmental issues, and harnessing AI's transformative power for positive impact. Increased transparency within the ICT sector is essential to assist environmental researchers and ML developers along the way.

ACKNOWLEDGMENTS

We would like to thank the team of "Foodforecast Technologies GmbH" and the bakery "Kettinger" for their assistance and willingness to provide data, which greatly facilitated our research.

Open access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

No funding.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supporting information of this article. Restrictions apply to the availability of data from Sphera and ecoinvent databases, which were used under license for this study and which are available at https://sphera.com/life-cycle-assessment-lca-database/ and https://ecoinvent.org/the-ecoinvent-database/. Other restrictions apply to specific (sales) data from bakeries' enterprise resource planning systems and specific cloud cost and usage report data from "Foodforecast," which are provided by the authors subject to the permission of the parties.

ORCID

Nicolas Hübner https://orcid.org/0009-0006-5455-5716
Justus Caspers https://orcid.org/0000-0001-7015-9779

REFERENCES

Acer Inc. (2022). *SGS life cycle assessment verification opinion: The inventory of life cycle assessment of LCD monitor (B247Y)*. https://www.acer-group.com/userfiles/LCA_B247Y.pdf

Amazon. (2022a). *Amazon's 2021 sustainability report*. https://sustainability.aboutamazon.com/2021-sustainability-report.pdf

Amazon. (2022b). *Q4 2021 Amazon.com, Inc. Earnings conference call*. https://ir.aboutamazon.com/events/event-details/2022/Q4-2021-Amazoncom-Inc-Earnings-Conference-Call/default.aspx

Amodei, D., & Hernandez, D. (2018). *AI and compute. OpenAI*. https://openai.com/blog/ai-and-compute/

Apple Inc. (2021). *14-inch MacBook Pro product environmental report*. https://www.apple.com/environment/pdf/products/notebooks/14-inch_MacBook_Pro_PER_Oct2021.pdf

AWS. (2023). *AWS documentation*. https://docs.aws.amazon.com/index.html

Azure. (2023). *Cloud computing services | Microsoft Azure*. https://azure.microsoft.com/de-de/

Baldé, C. P., Kuehr, R., Yamamoto, T., McDonald, R., D'Angelo, E., Althaf, S., Bel, G., Deubzer, O., Fernandez-Cubillo, E., Forti, V., Gray, V., Herat, S., Honda, S., Iattoni, G., Khetriwal, D. S., Luda di Cortemiglia, V., Lobuntsova, Y., Nnorom, I., Pralat, N., & Wagner, M. (2024). *Global E-waste monitor 2024*. International Telecommunication Union (ITU) & United Nations Institute for Training and Research (UNITAR).

Berthelot, A., Caron, E., Jay, M., & Lefèvre, L. (2024). Estimating the environmental impact of Generative-AI services using an LCA-based methodology. In L. Settineri & P. C. Priarone (Eds.), *Proceedings of CIRP LCE 2024, the 31st Conference on Life Cycle Engineering* (pp. 1–10), June 2024, Turin, Italy. HAL: hal-04346102v2f.

Bieser, J. C. T., & Hilty, L. M. (2018). Assessing indirect environmental effects of Information and Communication Technology (ICT): A systematic literature review. *Sustainability*, *10*(8), 2662. https://doi.org/10.3390/su10082662

Boavizta.org. (2021). *Digital & environment : How to evaluate server manufacturing footprint, beyond greenhouse gas emissions?*. https://www.boavizta.org/en/blog/empreinte-de-la-fabrication-d-un-serveur

Boyd, S. B. (2012). *Life-cycle assessment of semiconductors*. Springer. https://doi.org/10.1007/978-1-4419-9988-7

Bremer, C., Kamiya, G., Bergmark, P., Coroama, V. C., Masanet, E., & Lifset, R. (2023). Assessing energy and climate effects of digitalization: Methodological challenges and key recommendations. *nDEE Framing Paper Series. SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4459526

Chen, L., Chen, Z., Zhang, Y., Liu, Y., Osman, A. I., Farghali, M., Hua, J., Al-Fatesh, A., Ihara, I., Rooney, D. W., & Yap, P.-S. (2023). Artificial intelligence-based solutions for climate change: A review. *Environmental Chemistry Letters*, *21*(5), 2525–2557. https://doi.org/10.1007/s10311-023-01617-y

Cloud Carbon Footprint (CCF). (2023). *Cloud Carbon Footprint—An open source tool to measure and analyze cloud carbon emissions*. https://www.cloudcarbonfootprint.org

Colomb, V., Colsaet, A., Ait-Amar, S., Basset-Mens, C., Mevel, G., To, V., & van der Werf, H. (2015). AGRIBALYSE: The French public LCI database for agricultural products: high quality data for producers and environmental labelling. *OCL Oilseeds and fats crops and lipids*, *22*(1), D104. https://doi.org/10.1051/ocl/20140047

Coroamă, V. C. & Mattern, F. (Eds.) (2019). Digital rebound—Why digitalization will not redeem us our environmental sins. *Proceedings of the 6th international conference on ICT for Sustainability (ICT4S 2019)*. https://ceur-ws.org/Vol-2382/ICT4S2019_paper_31.pdf

Coroamă, V. C., Bergmark, P., Höjer, M., & Malmodin, J. (2020). A methodology for assessing the environmental effects induced by ICT services. In R. Chitchyan, D. Schien, A. Moreira, & B. Combemale (Eds.), *Proceedings of the 7th International Conference on ICT for Sustainability* (pp. 36–45). ACM. https://doi.org/10.1145/3401335.3401716

Coroamă, V. C., & Höjer, M. (2016). Assessing GHG benefits induced by ICT services in practice: A case study and resulting challenges. In P. Grosso, P. Lago, & A. Osseyran (Eds.), *Proceedings of ICT for Sustainability 2016* (pp. 29–35). Atlantis Press. https://doi.org/10.2991/ict4s-16.2016.4

CPU-World. (2023). CPU-World: BSP/CPU/Microprocessor/Co-processor/FPU/Microcontroller families. https://www.cpu-world.com/CPUs/

Crucial. (2023). *How much power does memory use?* https://www.crucial.com/support/articles-faq-memory/how-much-power-does-memory-use

Davy, B. (2021). *Estimating AWS EC2 instances power consumption—Teads engineering*. https://medium.com/teads-engineering/estimating-aws-ec2-instances-power-consumption-c9745e347959

EIZO Corporation. (2020). *Environmental product declaration: FlexScan EV2457*. https://www.eizoglobal.com/company/csr/promise2/epeat/LCA_EV2457.pdf

European Commission (EC). (2024). *Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008 on waste and repealing certain directives text with EEA relevance*. http://data.europa.eu/eli/dir/2008/98/2024-02-18/eng

European Telecommunications Standards Institute (ETSI). (2015). *ES 203 199 - V1.3.1: Environmental engineering (EE); Methodology for environmental life cycle assessment (LCA) of information and communication technology (ICT) goods, networks, and services*. https://www.etsi.org/deliver/etsi_es/203100_203199/203199/01.03.01_60/es_203199v010301p.pdf

Finkbeiner, M. (2009). Carbon footprinting—Opportunities and threats. *The International Journal of Life Cycle Assessment*, *14*(2), 91–94. https://doi.org/10.1007/s11367-009-0064-x

Food and Agriculture Organization of the United Nations (FAO). (2013). *Food wastage footprint: Impact on natural resources: Summary report*. https://www.fao.org/3/i3347e/i3347e.pdf

Foodforecast Technologies GmbH. (2022). *Künstliche Intelligenz (KI) für die optimale Bestellmenge*. https://foodforecast.com/

Fraunhofer Institute for Reliability and Microintegration (IZM). (2021). *Umweltbezogene Technikfolgenabschätzung Mobilfunk in Deutschland: Projekt UTAMO [Environmental impact assessment of mobile communications in Germany: Project UTAMO]*. https://www.umweltbundesamt.de/sites/default/files/medien/479/publikationen/texte_26-2023_umweltbezogene_technikfolgenabschaetzung_mobilfunk_in_deutschland.pdf

Fulton, L., Mason, J., & Meroux, D. (2017). *Three revolutions in urban transportation: How to achieve the full potential of vehicle electrification, automation, and shared mobility in urban transportation systems around the world by 2050*. https://trid.trb.org/View/1466512

Greenhouse Gas Protocol. (2015). *The greenhouse gas protocol: A corporate accounting and reporting standard* (Rev. ed.). World Business Council for Sustainable Development; World Resources Institute.

Gupta, U., Kim, Y. G., Lee, S., Tse, J., Lee, H.-H. S., Wei, G.-Y., Brooks, D., & Wu, C.-J. (2020). *Chasing carbon: The elusive environmental footprint of computing*. https://arxiv.org/pdf/2011.02839.pdf

Hamilton, J. (2008). *Why blade servers aren't the answer to all questions—Perspectives*. https://perspectives.mvdirona.com/2008/09/why-blade-servers-arent-the-answer-to-all-questions/

Horner, N. C., Shehabi, A., & Azevedo, I. L. (2016). Known unknowns: Indirect energy effects of information and communication technology. *Environmental Research Letters*, *11*(10), 103001. https://doi.org/10.1088/1748-9326/11/10/103001

IEA. (2022). *Data centres and data transmission networks—Analysis—IEA*. https://www.iea.org/reports/data-centres-and-data-transmission-networks

International Organization for Standardization (ISO). (2020a). *ISO 14040:2006, Environmental management: Life cycle assessment: Principles and framework*. https://www.iso.org/standard/37456.html

International Organization for Standardization (ISO). (2020b). *ISO 14044:2006, Environmental management: Life cycle assessment: Requirements and guidelines*. https://www.iso.org/standard/38498.html

International Telecommunication Union Telecommunication Standardization Sector (ITU-T). (2014). Recommendation ITU-T L.1410, Methodology for environmental life cycle assessments of information and communication technology goods, networks and services. https://www.itu.int/rec/T-REC-L.1410-201412-I/en

Jaeger, S. (2018). *Unser täglich Brot: Von überschüssigen Brotkanten und wachsenden Brotbergen [Our daily bread: On surplus bread edges and growing bread mountains]*. https://www.wwf.de/fileadmin/fm-wwf/Publikationen-PDF/WWF-Studie-Unser-taeglich-Brot_Von-ueberschuessigen-Brotkanten-und-wachsenden-Brotbergen_102018.pdf

Kaack, L., Priya, D., Emma, S., George, K., Felix, C., & David, R. (2021). *Aligning artificial intelligence with climate change mitigation*. https://hal.archives-ouvertes.fr/hal-03368037

Koomey, J., & Masanet, E. (2021). Does not compute: Avoiding pitfalls assessing the Internet's energy and carbon impacts. *Joule*, *5*(7), 1625–1628. https://doi.org/10.1016/j.joule.2021.05.007

Kumar, I., Rawat, J., Mohd, N., & Husain, S. (2021). Opportunities of artificial intelligence and machine learning in the food industry. *Journal of Food Quality*, *2021*, 1–10. https://doi.org/10.1155/2021/4535567

Lacoste, A., Luccioni, A., Schmidt, V., & Dandres, T. (2019). *Quantifying the carbon emissions of machine learning*. https://arxiv.org/pdf/1910.09700.pdf

Lange, S., Pohl, J., & Santarius, T. (2020). Digitalization and energy consumption. Does ICT reduce energy demand? *Ecological Economics*, *176*, 106760. https://doi.org/10.1016/j.ecolecon.2020.106760

Lannelongue, L., Grealey, J., & Inouye, M. (2021). Green algorithms: Quantifying the carbon footprint of computation. *Advanced Science*, *8*(12), 2100707. https://doi.org/10.1002/advs.202100707

Liao, M., Lan, K., & Yao, Y. (2022). Sustainability implications of artificial intelligence in the chemical industry: A conceptual framework. *Journal of Industrial Ecology*, *26*(1), 164–182. https://doi.org/10.1111/jiec.13214

Luccioni, A. S., Viguier, S., & Ligozat, A.-L. (2022). *Estimating the Carbon footprint of BLOOM, a 176B parameter language model*. arXiv. https://doi.org/10.48550/arXiv.2211.02001

Malmodin, J., & Lundén, D. (2018). The energy and carbon footprint of the global ICT and E&M sectors 2010 - 2015. In B. Penzenstadler, S. Easterbrook, C. Venters, & S. I. Ahmed (Eds.), *ICT4S2018 EPiC series in computing*, *52* (pp. 187–164). EasyChair. https://doi.org/10.29007/nvd8

Micron. (2017). *TN-40-07: Calculating memory power for DDR4 SDRAM*. https://www.micron.com/-/media/client/global/documents/products/technical-note/dram/tn4007_ddr4_power_calculation.pdf

Microsoft. (2018). *The carbon benefits of cloud computing A study on the Microsoft Cloud in partnership with WSP*. https://www.semanticscholar.org/paper/The-carbon-benefits-of-cloud-computing-A-study-on/3c5294d32900fda8b212200244499a7eb9790643

OpenAI. (2023). *GPT-4 technical report*. https://arxiv.org/pdf/2303.08774

Pant, R., & Zampori, L. (2019). *Suggestions for updating the Product Environmental Footprint (PEF) method* (EUR, Vol. 29682). Luxembourg: Publications Office of the European Union.

Park, J., Naumov, M., Basu, P., Deng, S., Kalaiah, A., Khudia, D., & Smelyanskiy, M. (2018). *Deep learning inference in Facebook data centers: Characterization, performance optimizations, and hardware implications*. https://arxiv.org/pdf/1811.09886

Patterson, D., Gonzalez, J., Holzle, U., Le, Q., Liang, C., Munguia, L.-M., Rothchild, D., So, D. R., Texier, M., & Dean, J. (2022). The carbon footprint of machine learning training will plateau, then shrink. *Computer*, *55*(7), 18–28. https://doi.org/10.1109/MC.2022.3148714

Pauer, E., Wohner, B., & Tacker, M. (2020). The influence of database selection on environmental impact results. Life cycle assessment of packaging using GaBi, Ecoinvent 3.6, and the environmental footprint database. *Sustainability*, *12*(23), 9948. https://doi.org/10.3390/su12239948

Peiró, L. T., & Ardente, F. (2015). *Environmental footprint and material efficiency support for product policy: Analysis of material efficiency requirements of enterprise servers* (EUR, Scientific and Technical Research Series Vol. 27467). EU Publications Office.

Pohl, J., Hilty, L. M., & Finkbeiner, M. (2019). How LCA contributes to the environmental assessment of higher order effects of ICT application: A review of different approaches. *Journal of Cleaner Production*, *219*, 698–712. https://doi.org/10.1016/j.jclepro.2019.02.018

Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science*, *360*(6392), 987–992.

Principato, L., Marchetti, S., Barbanera, M., Ruini, L., Capoccia, L., Comis, C., & Secondi, L. (2023). Introducing digital tools for sustainable food supply management: Tackling food loss and waste in industrial canteens. *Journal of Industrial Ecology*, *27*(4), 1060–1075. https://doi.org/10.1111/jiec.13391

Rasoldier, A., Combaz, J., Girault, A., Marquet, K., & Quinton, S. (2022). How realistic are claims about the benefits of using digital technologies for GHG emissions mitigation? In *Proceedings of the Eighth Workshop on Computing within Limits (LIMITS 2022)* (Article No. 39). Virtual, France. https://inria.hal.science/hal-03949261

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., … Bengio, Y. (2022). Tackling climate change with machine learning. *ACM Computing Surveys*, *55*(2), 1–96. https://doi.org/10.1145/3485128

Salinas, D., Flunkert, V., & Gasthaus, J. (2017). *DeepAR: Probabilistic forecasting with autoregressive recurrent networks*. https://arxiv.org/pdf/1704.04110

Scherhaufer, S., Moates, G., Hartikainen, H., Waldron, K., & Obersteiner, G. (2018). Environmental impacts of food waste in Europe. *Waste Management*, *77*, 98–113. https://doi.org/10.1016/j.wasman.2018.04.038

Schödwell, B., Zarnekow, R., Liu, R., Gröger, J., & Wilkens, M. (2018). *KPI4DCE Kennzahlen und Indikatoren für die Beurteilung der Ressourceneffizienz von Rechenzentren und Prüfung der praktischen Anwendbarkeit*. Umweltbundesamt. https://www.umweltbundesamt.de/publikationen/kennzahlen-indikatoren-fuer-die-beurteilung-der

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2019). *Green AI*. https://arxiv.org/pdf/1907.10597

Sehgal, A., & McDonnell, D. (2020). *Four trends driving global utility digitization*. https://aws.amazon.com/de/blogs/industries/four-trends-driving-global-utility-digitization/

Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J., & Lintner, W. (2016). *United States data center energy usage report*. https://www.scholars.northwestern.edu/en/publications/united-states-data-center-energy-usage-report

Simon, J. P. (2019). Artificial intelligence: Scope, players, markets and geography. *Digital Policy, Regulation and Governance*, *21*(3), 208–237. https://doi.org/10.1108/DPRG-08-2018-0039

Sphera. (2023). *Managed LCA content (GaBi databases): Content version 2023*. https://sphera.com/life-cycle-assessment-lca-database/

Standard Performance Evaluation Corporation (SPEC). (2022). *SERT suite*. https://www.spec.org/sert/

Standard Performance Evaluation Corporation (SPEC). (2023). *Benchmarks and tools*. https://www.spec.org/benchmarks.html#sfs

Stenmarck, Å., Hanssen, O. J., Silvennoinen, K., Katajajuuri, J.-M., & Werge, M. (2011). *Initiatives on prevention of food waste in the retail and wholesale trades*. Nordic Council of Ministers. https://doi.org/10.6027/TN2011-548

Strubell, E., Ganesh, A., & McCallum, A. (2019). *Energy and policy considerations for deep learning in NLP*. https://arxiv.org/pdf/1906.02243

Teehan, P., & Kandlikar, M. (2013). Comparing embodied greenhouse gas emissions of modern computing and electronics products. *Environmental Science & Technology*, *47*(9), 3997–4003. https://doi.org/10.1021/es303012r

Umweltbundesamt (UBA). (2021). *Green cloud computing: Lebenszyklusbasierte Datenerhebung zu Umweltwirkungen des Cloud Computing [Report]*. https://www.umweltbundesamt.de/en/publikationen/green-cloud-computing

UNEP. (2021). *UNEP Food Waste Index Report 2021. Nairobi*. https://www.unep.org/resources/report/unep-food-waste-index-report-2021

United Nations (UN). (2015). Transforming our world: The 2030 Agenda for Sustainable Development. https://sdgs.un.org/2030agenda

von Fischer, J. C., Cooley, D., Chamberlain, S., Gaylord, A., Griebenow, C. J., Hamburg, S. P., Salo, J., Schumacher, R., Theobald, D., & Ham, J. (2017). Rapid, vehicle-based identification of location and magnitude of urban natural gas pipeline leaks. *Environmental Science & Technology*, *51*(7), 4091–4099. https://doi.org/10.1021/acs.est.6b06095

Wäger, P. A., Hischier, R., & Widmer, R. (2015). The material basis of ICT. In L. M. Hilty & B. Aebischer (Eds.), *ICT innovations for sustainability* (Vol. 310, pp. 209–221). Springer International Publishing. https://doi.org/10.1007/978-3-319-09228-7_12

Walsh, N. (2022). *Wie Microsoft den Wasser- und Energieverbrauch von Rechenzentren misst, um die Nachhaltigkeit von Azure zu verbessern*. https://azure.microsoft.com/de-de/blog/how-microsoft-measures-datacenter-water-and-energy-use-to-improve-azure-cloud-sustainability/

Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-Ruiz, E., & Weidema, B. (2016). The ecoinvent database version 3 (part I): Overview and methodology. *The International Journal of Life Cycle Assessment*, *21*(9), 1218–1230. https://doi.org/10.1007/s11367-016-1087-8

Widdicks, K., Lucivero, F., Samuel, G., Croxatto, L. S., Smith, M. T., Holter, C. T., Berners-Lee, M., Blair, G. S., Jirotka, M., Knowles, B., Sorrell, S., Rivera, M. B., Cook, C., Coroamă, V. C., Foxon, T. J., Hardy, J., Hilty, L. M., Hinterholzer, S., & Penzenstadler, B. (2023). Systems thinking and efficiency under emissions constraints: Addressing rebound effects in digital innovation and policy. *Patterns*, *4*(2), 100679. https://doi.org/10.1016/j.patter.2023.100679

WikiChip. (2023). *Chips & Semi*. https://en.wikichip.org/wiki/WikiChip

Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., & Hazelwood, K. (2021). *Sustainable AI: Environmental implications, challenges and opportunities*. https://arxiv.org/pdf/2111.00364

Zhang, B., & Dafoe, A. (2019). Artificial intelligence: American attitudes and trends. *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.3312874

Zhang, D., Maslej, N., Brynjolfsson, E., Etchemendy, J., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., Sellitto, M., Sakhaee, E., Shoham, Y., Clark, J., & Perrault, R. (2022). *The AI Index 2022 annual report*. arXiv. https://doi.org/10.48550/arXiv.2205.03468

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hübner, N., Caspers, J., Coroamă, V. C., & Finkbeiner, M. (2024). Machine-learning-based demand forecasting against food waste: Life cycle environmental impacts and benefits of a bakery case study. *Journal of Industrial Ecology*, *28*, 1117–1131. https://doi.org/10.1111/jiec.13528