How Sustainable is Machine Learning in Energy Applications? – The Sustainable Machine Learning Balance Sheet

Simon Wenninger
Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Augsburg, Germany;FIM Research Center, University of Applied Sciences Augsburg, Augsburg, Germany, simon.wenninger@fim-rc.de

Can Kaymakci
Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart, Germany;Institute for Energy Efficiency in Production EEP, University of Stuttgart, Stuttgart, Germany, can.kaymakci@ipa.fraunhofer.de

Christian Wiethe
Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Augsburg, Germany, christian.wiethe@fit.fraunhofer.de

Jörg Römmelt
Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Augsburg, Germany, joerg.roemmelt@fit.fraunhofer.de

Lukas Baur
Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart, Germany, lukas.baur@ipa.fraunhofer.de

See next page for additional authors

Presenter Information
Simon Wenninger, Can Kaymakci, Christian Wiethe, Jörg Römmelt, Lukas Baur, Björn Häckel, and Alexander Sauer

How Sustainable is Machine Learning in Energy Applications? – The Sustainable Machine Learning Balance Sheet

Simon Wenninger[1,2], Can Kaymakci[3,4], Christian Wiethe[1], Jörg Römmelt[1], Lukas Baur[3], Björn Häckel[1,2] and Alexander Sauer[3,4]

[1] Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Augsburg, Germany
{joerg.roemmelt, christian.wiethe}@fit.fraunhofer.de

[2] FIM Research Center, University of Applied Sciences Augsburg, Augsburg, Germany
{simon.wenninger, bjoern.haeckel}@fim-rc.de

[3] Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart, Germany
{can.kaymakci, alexander.sauer}@ipa.fraunhofer.de

[4] Institute for Energy Efficiency in Production EEP, University of Stuttgart, Stuttgart, Germany

Abstract. Information Systems play a central role in the energy sector for achieving climate targets. With increasing digitization and data availability in the energy sector, data-driven machine learning (ML) approaches emerged, showing high potential. So far, research has focused on optimizing ML approaches' prediction performance. However, this is a one-sided perspective. ML approaches require large computation times and capacities leading to high energy consumption. With the goal of sustainable energy systems, research on ML approaches should be extended to include the application's energy consumption. ML solutions must be designed in such a way that the resulting savings in energy (and emissions) are greater than the energy consumption caused using the ML solution. To address this need, we develop the Sustainable Machine Learning Balance Sheet as a framework allowing to holistically evaluate and develop sustainable ML solutions which we validated in a case study and through expert interviews.

Keywords: Machine Learning, Sustainability, Green IS, Data-driven Approaches, Energy Informatics.

1 Introduction

The ambitious climate targets until 2030 require an expansion of renewable energies to reduce harmful emissions in the long term and ensure a sustainable energy supply [1]. In addition to the historically established efforts with a strong focus on efficiency improvements, the challenges of volatile energy supply from renewable sources must likewise be faced [2–4]. This results in a more complex energy supply and demand management, which is intensified by cross-sector and cross-energy source solutions in the context of integrated energy systems to ensure a low-emission energy supply [5].

17th International Conference on Wirtschaftsinformatik,
February 2022, Nürnberg, Germany

To enable the management of energy flows and infrastructure, the use of Information Systems (IS) becomes increasingly important [6, 7]. The concept of Green IS, understood as the planning, implementation, and management of IS to support sustainable action [8], and Green IT to ensure the sustainable and efficient use of information and communication technology [9] are central concepts that bundle corresponding work.

In particular, the use of data-driven approaches in a wide range of applications from anomaly detection analysis in energy consumption to the determination of annual energy consumption in buildings stands out and shows high potential compared to conventional, often engineering-based approaches [10, 11]. So far, work on data-driven solutions, mainly Machine Learning (ML), has been driven by the goal of achieving the highest possible quality of results, e.g., correctly identified consumers in energy consumption or prediction accuracy in energy consumption forecasts [12, 13]. However, this one-sided perspective ignores important aspects, as the focus is solely on the result and the preceding necessary steps are neglected. Data-driven approaches require significant computation times and capacities leading to high energy consumption [14]. The energy consumption of data-driven approaches has only been investigated in the research domain of computer science from a mostly non-application-oriented perspective [15]. With the goal of sustainable energy supply and at the same time increasing application of data-driven solutions due to the rising complexities of management between energy supply and energy demand, the research of data-driven solutions should be extended to include the component of energy consumption of the actual application. Data-driven solutions should be designed in such a way that the resulting savings in energy and emissions are greater than the energy consumption and emissions caused by the data-driven solution. This study therefore analyzes the following guiding research question:

How can a framework be designed that holistically quantifies the sustainability impact of data-driven solutions by considering both the energy consumption caused by data-driven solutions and the savings achieved?

To this end, we introduce the Sustainable Machine Learning Balance Sheet (SMLBS) following a design science research approach and input from expert interviews in the AI research domain. The SMLBS allows deriving a holistic evaluation by considering all ecological effects in the form of a balance sheet, indicating the net sustainability impact of data-driven solutions. With the development of the SMLBS, we contribute to literature by being the first to demonstrate the consideration of energy consumption of data-driven solutions in energy applications. Furthermore, we provide a framework that enables future researchers and practitioners to perform a use case-specific and holistic analysis of their data-driven solutions. Using the SMLBS, data-driven solutions can be evaluated in terms of both energy consumption and emissions so that computing times can also be optimized for demand-side management and adaptation to the availability of renewable energy. We further highlight new perspectives for future research on data-driven approaches and ways to address both opportunities and challenges in equal measure so that energy informatics can contribute an important part to achieving climate goals.

2 Theoretical Background

2.1 The Potential of Green IT, Green IS, and Data-driven Approaches

IS contribute a large part to successful sustainable transformations in society and organizations [8, 16]. At the same time, the ever-increasing usage of IS is fueling research into energy-efficient IT to ensure sustainable and efficient use of information and communication technology. To this end, IS concerned with efficiency gains and energy-efficient IT coined the terms Green IS and Green IT. Energy Informatics as a subfield of Green IS research focuses on reducing energy consumption by using information and communication technology [17]. Both, Green IS and Energy Informatics accordingly come into focus of current research and applications in practice. A critical success factor for IS in energy applications is transparency about historical, current, and future energy flows and related measures. Consequently, data collected by sensors and IS are highly important to optimize decisions for managing and controlling energy supply and demand to minimize overall energy consumption and reduce emissions [4]. With the increasing amount of data, data-driven approaches have become more relevant in research for a wide range of applications in the energy sector. Examples are forecasting future energy consumption, detecting anomalies in energy consumption curves, and leveraging efficiency or flexibility potentials.

Most data-driven applications use methods from the area of Artificial Intelligence (AI) or, more precisely, ML. Such applications are bundled into AI systems [18] with their components and functionalities. In this context, the fundamental goal of an AI system is the value-adding transformation (data processing) of incoming data from data sources to outgoing data in the form of actions by an agent [18]. An ML algorithm that enables the actual application is the central component of an AI system in addition to data sources (e.g., sensors), a processing module to preprocess the incoming data, and an agent to execute actions.

ML algorithms can be classified as either supervised, unsupervised, or reinforcement learning techniques. Supervised ML algorithms learn a function that maps an input to an output, thus requiring labeled training data. Examples are applications of energy consumption prediction in energy consumption [19]. In unsupervised learning, there are no known target values in advance, and one tries to detect patterns in the available data [20]. Unsupervised learning can, for example, be used for anomaly detection in energy consumption if no labels are available (often expensive and time-consuming to generate) [11]. In reinforcement learning, an agent independently learns a strategy to maximize rewards and, therefore, often serves the purpose of optimization. Optimization tasks, e.g., in energy management or the optimization of bidding strategies of prosumers in local energy markets, can be addressed by reinforcement learning [21, 22].

ML shows high potential in the energy sector and surpasses approaches of the engineering discipline, which are often based on physical laws. For example, [10] achieved accuracy advantages of almost 50% in determining the annual energy consumption of residential buildings for energy performance certificates using data-driven methods compared to the engineering-based approach currently still required by law. Research on ML has also been successful in predicting future energy consumption, energy prices, or detecting anomalies and individual devices in energy consumption, as well as reinforcement approaches for energy management [11, 23–26].

Research on ML has been motivated by prediction performance. A single focus on performance evaluation measures, however, falls short when assessing the potential and value from a holistic perspective. Therefore, the achieved energy or emission savings and the energy necessary to run an often computation-intensive ML algorithm should be included when holistically assessing ML applications.

Approaches to quantifying energy and emission savings or efficiency increases are to the best of the authors' knowledge not yet the focus of IS research and mainly stem from engineering domains dealing with more technical problems. Interdisciplinary research and an extension of the scope of development goals are important to develop and operate sustainable ML applications.

2.2 Quantifying Machine Learning's Energy Consumption

Due to the large amount of data needed to train ML algorithms, powerful computer systems are required [14, 15]. With their powerful equipment and the additional components, such as Graphic Processing Units (GPU), these computer systems have an increased energy demand [27]. The high energy demand is already discussed and pointed out in some research groups [15, 28]. Measuring energy consumption is a growing part of computer science research, but as mentioned, still not yet established in IS research and not considered in real-world applications.

To quantify a computer system's energy consumption, one needs to define the scope of energy consumption measurement, such as at which point the energy consumption is measured and which components are included [29]. The most obvious is to measure the consumption of the system's individual components with additional hardware sensors [15]. This procedure will lead to the most accurate data but also requires a great deal of effort. Effort and expenses arise from additional sensors and extra wiring, which must be installed into the computer system, respectively, at the selected components [30]. An easier way is to use embedded sensors of the respective components, if available [15]. Alternatively, it is possible to measure the energy consumption of the whole computer system at the power supply, but with the drawback of more inaccurate quantification. In addition, the periphery in a data center or external components for computational performance or data storage / management can be included in the quantification to provide a holistic picture of the analyzed system's energy consumption.

To obtain measurement data in practice, two possibilities are presented in literature: first, the just mentioned use of sensors to measure the actual consumption [31], second the data from already existing measurements of a similar system with the same application purpose and transfer its results to their setup [32]. To date, the use of embedded sensors as measuring technology for energy consumption is widely used. Besides the limitation of available technologies, which speaks for a not yet intensely researched area [15], a recent literature review shows that energy consumption is not yet considered in the development of ML applications but should not be underestimated [33].

3 Methodology and Study Design

In this study, we structured our research following the design science in information systems research framework combining behavioral science and design science paradigms [34]. According to Gregor and Hevner [35] we contribute an *invention* to existing solutions in practice and research. We argue that the SMLBS forms a new solution for new problems becoming increasingly important. Despite growing steadily, the number of ML applications in energy informatics research is still relatively new and poorly established due to the often one-sided and not holistic research perspective. On the solution maturity side, there are no approaches that bridge the gap between the energy / emission savings and the energy demand / emissions of ML applications to the best of our knowledge. Therefore, we are the first to present this perspective on ML services in energy informatics research. In a first step, we derive potentials of ML approaches and existing solutions to quantify ML's energy consumption from the knowledge base conducting a semi-structured literature review and expert interviews (AI research experts – mostly in manufacturing-energy applications). We searched in the databases Google Scholar, Scopus, and AIS eLibrary with the keywords "AI", "ML", and "energy consumption". Expert interviews were conducted in a semi-structured manner with few predetermined questions and the possibility to spontaneously explore relevant parts in more depth. To ensure a high fit with the application domain, we consulted the experts before the initial introduction of the SMLBS to derive design and applicability criteria. Further interviews were conducted afterwards to review the results and correct design errors. The corresponding results were in part already presented in Section 2 and will be further elaborated in Section 4. In a second step, after several iterations of designing the SMLBS, we demonstrate, apply, and validate the SMLBS with a case study of AI-based anomaly detection in energy consumption (Section 5) and discuss the case study's results (Section 6). We communicate our results with this paper enabling practitioners and researchers to evaluate and design sustainable ML services.

4 The Sustainable Machine Learning Balance Sheet

To derive the necessary design and applicability requirements we conducted semi-structured expert interviews with practitioners from the AI research domain. The predetermined questions included the extent of AI use in the company, relevant KPIs in its design, implementation, and evaluation, as well as the interoperability of individual AI algorithms. Due to the semi-structured interview approach, we subsequently explored topics of particular interest to the individual experts regarding design and applicability requirements. To this end, the SMLBS should fulfill (1) applicability to different use cases, (2) applicability to different target variables, and (3) applicability to relevant subsystems. Additionally, acceptance criteria extracted from expert interviews are a simple understanding of the SMLBS and interdisciplinary use. We conceptualized the SMLBS based on these criteria before reconsulting the experts and iteratively shaping the SMLBS.

(1) With the SMLBS, we pursue the goal of use case-specific investigations of the net energy savings caused by data-driven solutions, considering both the achieved savings and corresponding energy consumption. This necessitates the balancing act between a generic framework, which serves its purpose for different applications, and a detailed procedure description, which provides further helpful use case-specific information, but at the same time restricts the application of deviating use cases. Different applications in the energy sector of supervised, unsupervised, reinforcement, or hybrid learning should be representable with the framework.

(2) Also, the SMLBS should allow for investigation under different target variables. For example, next to the mentioned energy consumption, efficiency improvements, energy flexibility, or emission reductions are useful target variables. Defining energy consumption as target variable for a load shifting use case in demand-side management would not achieve any benefit under these conditions. However, if emissions caused (e.g., CO_2 emissions) are a target variable, then load shifting can achieve significant savings in times with a higher share of renewable energy in the electricity mix. Moreover – if quantifiable –, social and governance impacts are also possible target variables to holistically depict sustainability [36]. Consequently, as in other optimization problems, the use case-specific definition of target variables is of central importance. Henceforth, we speak of energy savings for simplicity. All other target variables could be inserted as well.

(3) When determining the energy savings and the required energy consumption for the data-driven services, the SMLBS should also be restrictable to the relevant subsystems, e.g., which energy flows are considered, and which are not. Particularly in the case of complex applications that may be linked to other data-driven applications, it is important to define the system boundaries so that the savings can be correctly quantified and allocated. The same applies to the quantification of the ML service's energy consumption. For example, it must be weighed up whether the energy consumption of a higher-level data management system should also be used for the specific investigation. This allows correctly comparing different variants of (data-driven) services without distortions and to develop / select the optimal solution.

These requirements make up the strategic boundaries within which we developed the initial version of the SMLBS before reconsulting the experts. **Figure 1** depicts the final version of the SMLBS after the second interviews based on the balance sheet known from external accounting. It, however, contrasts savings and consumption of ML algorithms instead of assets and liabilities. The requirements are depicted at the top. The left-hand side describes the ML energy demand. It covers all steps within the relevant subsystem as defined in requirement (3), from the initial model setup and training to the energy consumption after deployment. Quantification methods for energy consumption during training and in operation have been discussed in Section 2.2. For instance, by implementing sensors or using measurements of similar systems. The right side covers all realized savings. We further divide the realized savings into direct savings (improvements in the target variable) and indirect savings (secondary objectives, e.g., explainability / trust, transparency). To this end, direct savings are easily quantified, as the value for the target variable is usually known before ML deployment. Under the assumption of fixed exogenous variables, the difference in the measured or calculated target variable gives the savings. Else, exogenous influences must first be extracted from the calculation.

Indirect savings are not as easily quantified because there is possibly no known benchmark. Nonetheless, there are several empirically observed effects. For instance, the application of explainable AI techniques could lead to further insights into energy-intensive processes which allow deriving implications for higher energy efficiency. These savings achieved through downstream activities must be budgeted accordingly - for example, with the application of conversion factors. Holistically evaluating the effect of ML applications thus requires the ex-ante determination of relevant key figures which can then be compared to later figures after the introduction of the ML application.

The difference between the two sides of the SMLBS then results in net energy savings or additional energy consumption, here schematically depicted on the left side as sustainable surplus. Note, that opposing effects may arise in similar applications due to the indirect rebound effect known from energy retrofitting, when an increase in energy efficiency may lead to fewer energy savings than expected [37]. For example, if a human user is involved in a process, their behavior may change over time and affect the amount of savings [38]. Consequently, iterative cycles and continuous monitoring with the SMLBS are necessary to ensure long-term sustainable outcomes.

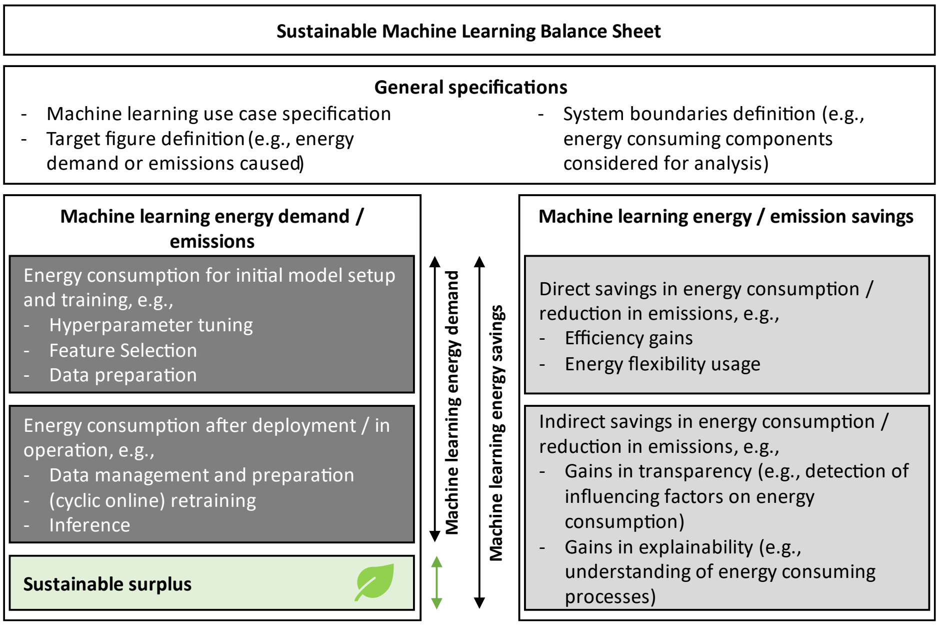

Figure 1. Sustainable Machine Learning Balance Sheet

We can now verify the consideration of the desired requirements stated above. (1) It becomes clear that the SMLBS does not restrict use cases since the design is use case-agnostic. (2) Despite depicting energy consumption in **Figure 1**, the target variable is not necessarily fixed. One might just as well list all efficiency gains, flexibility gains, or emission reductions, to name a few. (3) The balance sheet does not require the consideration of all impacts from either laterally linked ML applications or hierarchical data-management systems. This allows for a restriction to the relevant subsystem. Additionally, due to the similarity to the widely established balance sheet, the SMLBS is easy to understand and use, fulfilling the acceptance criteria.

After having defined the SMLBS, which consoles the foundations for the investigation of ML applications, we arrange the application of the SMLBS in two cases. For this purpose, we follow the four-phase model developed by [18] that supports developers and project managers in realizing AI systems. The authors define the phases of *planning*, *experimentation*, *implementation*, and *operation & optimization* over the life cycle of an AI system. The phases are run sequentially or iteratively (cf. **Figure 2** – grey arrows). From the four-phase model, we derive two application scenarios for the SMLBS, illustrated in **Figure 2**. On the one hand, the SMLBS can be applied in a pre-implementation phase comprising the steps of *planning* and *experimentation*. On the other hand, the SMLBS can be used in a post-implementation phase referring to the *operation & optimization* phase according to [18]. In the pre-implementation phase, the focus is on the design and conception of the AI system, which allows many options to be considered in terms of energy demand and energy savings. For example, different ML algorithms or training variants (e.g., online training) can be tested and compared to achieve an optimal energy demand and energy savings ratio. The evaluation of energy savings is often limited in this phase, as they must be collected conceptually since planning and experimentation usually do not occur in a productive environment. In the post-implementation phase, the energy demand and energy savings of existing AI systems can be recorded and empirically investigated. The results allow to modify and improve the AI system design. Of course, depending on the modularity of the AI system, the modification possibilities are smaller than in the pre-implementation phase, where greenfield development is possible. If the results are not satisfactory, iterative loops can be triggered on previous phases, analogous to the classical design science research process according to [39]. Here, the iteration loops can be triggered by both the net energy savings resulting from the SMLBS and prediction accuracy. Also, a multi-objective optimization problem can be defined for the design of an AI system that considers the sustainable surplus resulting from the SMLBS and the prediction accuracy.

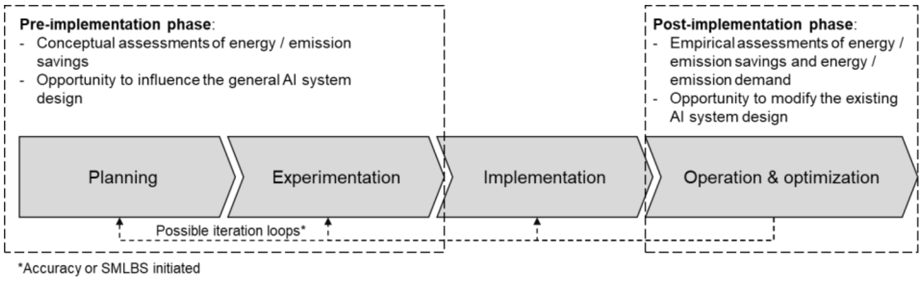

Figure 2. Application of the SMLBS in a pre- and post-implementation phase in the four-phase model of Kaymakci et al. [18] for AI system design

5 Validation – An Anomaly Detection Case Study in Energy Consumption

For demonstration purposes, the SMLBS is applied in an anomaly detection case study in the pre-implementation phase, allowing us to quantify energy savings and energy demand conceptually. For this purpose, a conceptually developed AI system according to [18] detects anomalies in the energy consumption of an open access dataset from Qatar University (QUD) [40]. Energy consumption is the target variable for our case study. The system boundary is set for the achieved direct savings by the power meter alone. For the energy demand, energy consumption for the initial training, retraining, and prediction is considered by quantifying the Central Processing Units'(CPU)'s and GPU's consumption.

In addition to the second-by-second energy consumption, the QUD contains a label called "micro-moments" that classifies the energy consumption into five classes, which allows identifying states such as a significantly too high energy consumption. Thus, we can formulate a supervised ML problem and identify conditions with too high energy consumption with the AI system, with the underlying assumption that the labels of the different classes in energy consumption can only be determined retrospectively. The approach early detects anomalous behavior and allows countermeasures to come back to normal energy consumption, resulting in energy savings. To test the SMLBS, we divide the data collected over a period of three months into a training and test part and create a balanced training dataset to avoid training bias. A linear support vector classifier (LinearSVC) model then learns to predict the energy consumption class which we determined by using cross-validation of different models such as a random forest classifier or a multinomial naïve Bayes classifier. After the initial training of the LinearSVC model, we predict the class of high energy consumption and retrain the model once an hour.

To calculate the net energy savings according to the SMLBS, the resulting energy savings are calculated by taking the difference between the energy consumption classified as too high and the energy consumption in a normal state. The assumption is that if we detect too high energy consumption at an early stage, we can initiate countermeasures and reduce energy consumption to a normal state. The difference between the actual too high energy consumption and the reduced state consequently results in energy savings. The energy consumption for initial training of the LinearSVC model, retraining, and anomaly detection is obtained by applying the python package “pyJoules” [41]. The package measures the energy consumption of the CPU (Intel Core i5-6300U) cores, cache, integrated GPU, and the RAM (16 GB) of the computer used for deploying our service. The deployed service gets the energy consumption in a specific time and generates a prediction of the energy consumption class.

In the case study, there were savings in energy consumption of 0.36 kWh over the test period of three days and eight hours (cf. Figure 3). This corresponds to a share of 13% in energy savings of the application represented by the QUD. For our service we measured an energy consumption of 0.09 kWh for the AI system which is far less than the savings resulting in net energy savings of 0.26 kWh. To better understand the results, we have approximated the energy consumption of a conventional notebook (15 W) and a RaspberryPi 3 (2 W) for the respective period, which is presented in Figure 3 [42, 43]. The conventional notebook's consumption of 1.19 kWh is clearly above the measured values and above the achieved savings. Consequently, the application of a notebook would not be sustainable for our case study. The consumption of the RaspberryPi 3 with 0.16 kWh is just above the measured consumption and would achieve net savings. For our case study, we conclude that under the specifications and system boundaries defined at the beginning, the use of ML makes sense and the savings achieved exceed the consumption of the ML application. Using an ML algorithm on the hardware of a notebook, in contrast to a RaspberryPi 3, is not recommended because the consumption clearly exceeds the savings.

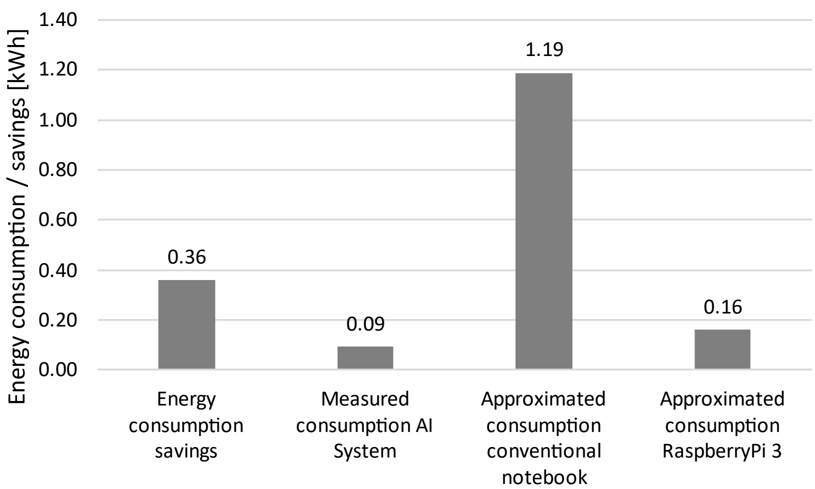

Figure 3. Representation of energy consumption savings and demand over the period of three days and eight hours

6 Discussion

Our presented case study allows us to validate and discuss the SMLBS with findings from the conducted expert interviews. In doing so, we first derive and discuss practical implications before we list limitations and prospects for further research. With our case study we show that the overall outcome in terms of sustainability is highly use case-specific. Depending on data availability, model complexity, retraining necessities, etc. the degree of sustainability and net energy savings may vary. In fact, when considering non-energy-intensive processes, continuous retraining may outweigh the benefits gained from the AI application. This clearly shows the need for the SMLBS, as trivial guidelines may easily fail, and sustainability projects may otherwise actually be counterproductive. The SMLBS then provides the quantitative decision support to settle on alternative approaches, e.g., less calculation-intensive AI applications, rule-based systems, or longer retraining cycles. Furthermore, the comparison with conventional notebooks and RaspberryPis as running hardware showed that small computers have an advantage. In practice, this means that when developing AI systems, computing capacity should be shared across multiple applications and should not be oversized.

In research and practice, there is often a prioritization of a single evaluation criterion when implementing an ML application, e.g., accuracy, explainability, or economics. However, multi-objective optimization provides solutions which are closer to the desired business strategy as pointed out by the interviewed experts. The SMLBS allows for the additional consideration of overall sustainability, which previously could only be partially considered. For instance, the SMLBS can also be used to evaluate whether a transparent and by design explainable ML algorithm (e.g., QLattice [12]) offers advantages in terms of both accuracy and energy consumption compared to sequentially applied post-hoc techniques (e.g., SHAP values [44]) in XAI applications. It is also possible to consider the higher energy consumption of energy-intensive applications (e.g., large server farms) and the resulting costs due to CO_2 prices, which means that energy consumption can be quantified directly in economic considerations as an important aspect in practice as mentioned by several experts interviewed.

The SMLBS provides extension and integration potential due to its (technology) agnostic and general approach. It can be considered as a meta model which may integrate already existing frameworks and tools. For instance, data-driven approaches may be incorporated by building upon the CRISP DM [45] and [18].

Our study disposes of some limitations. First, we validated the SMLBS with a highly specific case study and semi-structured expert interviews which mostly stem from a manufacturing background. Also, the expert interviews led to some corrections of early design flaws, indicating the relevance to consult practitioners. Further structured interviews with practitioners from other domains appear fruitful. Second, we designed the SMLBS highly model-agnostic. Despite the aforementioned advantages, this provides little guidance during application / implementation. Third, the energy savings might be hard to quantify, or the ML application only represents a support process of a Green IS application. Fourth, the SMLBS falls short of providing benchmarks or alternatives with better or worse ratios of energy demand and savings of ML applications.

These limitations give rise to further research potential. Further validation of the applicability of the SMLBS through alternative case studies and application in other fields such as energy-intensive industries is advisable. Thereby, further guidelines and quantification recommendations for energy savings and consumption could be obtained allowing for a wide-spread application. Additionally, future research might focus on deriving more specific details for highly used adaptations overcoming the generic nature of the SMLBS. For practical application developing an “SMLBS-as-a-Service” which can be deployed with ML models/data-driven solutions (for instance in an open access code repository) is advisable in future research projects. Last, upcoming studies could attempt to find application-agnostic benchmarks or values for good to very good ratios of energy demand and savings.

7 Conclusion

In our study, we pursued the goal of developing a framework that allows a holistic use case-specific investigation of data-driven solutions regarding a target variable (e.g., energy consumption) and quantifies the achieved savings. We, therefore, introduced the Sustainable Machine Learning Balance Sheet (SMLBS) that allows considering all relevant effects of data-driven solutions in the form of a balance sheet, indicating the net impact in the target variable (e.g., sustainability) of data-driven solutions. In addition, we provide ways to apply the SMLBS in the development of data-driven solutions, which we rank based on the four-phase model for the development of AI systems according to [18]. With our developed artefacts, we contribute an important framework to the under-researched consideration of energy consumption aspects of data-driven solutions.

By applying the SMLBS in a case study of anomaly detection regarding energy consumption, we could show that for the analyzed case, the energy consumption of an ML application does not exceed the achieved energy consumption savings and is therefore suitable from a sustainability perspective. Based on the case study and conducted expert interviews, we could also derive first practical implications. For example, the hardware of an AI system has a decisive influence on energy consumption. Thus, it is recommended not to over-dimension the computing capacity when designing AI systems to avoid unnecessarily high energy consumption. Although our work has some limitations and prospects for further research, we made an important contribution to the design of AI systems in energy-related applications to realize the full potential of data-driven applications to achieve climate goals.

References

1. Sachs, J.D., Schmidt-Traub, G., Mazzucato, M., Messner, D., Nakicenovic, N., Rockström, J.: Six Transformations to achieve the Sustainable Development Goals. Nat Sustain 2, 805–814 (2019)

2. Langevin, J., Harris, C.B., Satre-Meloy, A., Chandra-Putra, H., Speake, A., Present, E., Adhikari, R., Wilson, E.J., Satchwell, A.J.: US building energy efficiency and flexibility as an electric grid resource. Joule (2021)
3. Bank, L., Wenninger, S., Köberlein, J., Lindner, M., Kaymakci, C., Weigold, M., Sauer, A., Schilp, J.: Integrating Energy Flexibility in Production Planning and Control : An Energy Flexibility Data Model-Based Approach. 2nd Conference on Production Systems and Logistics (CPSL) (2021)
4. Lindner, M., Wenninger, S., Fridgen, G., Weigold, M.: Aggregating Energy Flexibility for Demand-Side Management in Manufacturing Companies – A Two-Step Method. Wissenschaftliche Gesellschaft für Produktionstechnik - Jahreskongress 2021 - "Produktion im Grenzbereich – Durch Innovation den Wandel gestalten" (2021)
5. Ramsebner, J., Haas, R., Ajanovic, A., Wietschel, M.: The sector coupling concept: A critical review. WIREs Energy Environ. 10 (2021)
6. Singh, M., Sahu, G.P.: Towards adoption of Green IS: A literature review using classification methodology. International Journal of Information Management 54, 102147 (2020)
7. Bauer, D., Hieronymus, A., Kaymakci, C., Köberlein, J., Schimmelpfennig, J., Wenninger, S., Zeiser, R.: Wie IT die Energieflexibilitätsvermarktung von Industrieunternehmen ermöglicht und die Energiewende unterstützt. HMD (2020)
8. Melville, N.P.: Information systems innovation for environmental sustainability. Management information systems : mis quarterly 34, 1–21 (2010)
9. Murugesan, S.: Harnessing Green IT: Principles and Practices. IT Prof. 10, 24–33 (2008)
10. Wenninger, S., Wiethe, C.: Benchmarking Energy Quantification Methods to Predict Heating Energy Performance of Residential Buildings in Germany. Bus Inf Syst Eng 63, 223–242 (2021)
11. Kaymakci, C., Wenninger, S., Sauer, A.: Energy Anomaly Detection in Industrial Applications with Long Short-term Memory-based Autoencoders (Under Revision). 54th CIRP Conference on Manufacturing Systems (2021)
12. Wenninger, S., Kaymakci, C., Wiethe, C.: Explainable long-term building energy consumption prediction using QLattice (in Press). Applied Energy
13. Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L.: ImageNet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248–255. IEEE (2009 - 2009)
14. Lane, N.D., Bhattacharya, S., Georgiev, P., Forlivesi, C., Kawsar, F.: An Early Resource Characterization of Deep Learning on Wearables, Smartphones and Internet-of-Things Devices. In: Xu, C., Zhang, P., Sigg, S. (eds.) Proceedings of the 2015 International Workshop on Internet of Things towards Applications, pp. 7–12. ACM, New York, NY, USA (112015)
15. García-Martín, E., Rodrigues, C.F., Riley, G., Grahn, H.: Estimation of energy consumption in machine learning. Journal of Parallel and Distributed Computing 134, 75–88 (2019)
16. Donnelly, J., John, A., Mirlach, J., Osberghaus, K., Rother, S., Schmidt, C., Voucko-Glockner, H., Wenninger, S.: Enabling the Smart Factory : A Digital Platform Concept for Standardized Data Integration. 2nd Conference on Production Systems and Logistics (CPSL) (2021)

17. Watson, Boudreau, Chen: Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community. MIS Quarterly 34, 23 (2010)
18. Kaymakci, C., Wenninger, S., Sauer, A.: A Holistic Framework for AI Systems in Industrial Applications. 16. Internationale Tagung Wirtschaftsinformatik (2021)
19. Ahmad, T., Chen, H., Huang, R., Yabin, G., Wang, J., Shair, J., Azeem Akram, H.M., Hassnain Mohsan, S.A., Kazim, M.: Supervised based machine learning models for short, medium and long-term energy prediction in distinct building environment. Energy 158, 17–32 (2018)
20. Ahlrichs, J., Wenninger, S., Wiethe, C., Häckel, B.: Impact of socio-economic factors on local energetic retrofitting needs - A data analytics approach. Energy Policy 160, 112646 (2022)
21. Bose, S., Kremers, E., Mengelkamp, E.M., Eberbach, J., Weinhardt, C.: Reinforcement learning in local energy markets. Energy Inform 4 (2021)
22. Xu, X., Jia, Y., Xu, Y., Xu, Z., Chai, S., Lai, C.S.: A Multi-Agent Reinforcement Learning-Based Data-Driven Method for Home Energy Management. IEEE Trans. Smart Grid 11, 3201–3211 (2020)
23. Bourdeau, M., Zhai, X.q., Nefzaoui, E., Guo, X., Chatellier, P.: Modeling and forecasting building energy consumption: A review of data-driven techniques. Sustainable Cities and Society 48, 101533 (2019)
24. Spiliotis, E., Doukas, H., Assimakopoulos, V., Petropoulos, F.: Forecasting week-ahead hourly electricity prices in Belgium with statistical and machine learning methods. In: Mathematical Modelling of Contemporary Electricity Markets, pp. 59–74. Elsevier (2021)
25. Salem, H., Sayed-Mouchaweh, M., Tagina, M.: A Review on Non-intrusive Load Monitoring Approaches Based on Machine Learning. In: Sayed-Mouchaweh, M. (ed.) Artificial Intelligence Techniques for a Scalable Energy Transition, pp. 109–131. Springer International Publishing, Cham (2020)
26. Kim, S., Lim, H.: Reinforcement Learning Based Energy Management Algorithm for Smart Energy Buildings. Energies 11, 2010 (2018)
27. Henderson, P., Hu, J., Romoff, J.: Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning. MICROTOME PUBL, BROOKLINE (2020)
28. Chen, Y.-H., Krishna, T., Emer, J.S., Sze, V.: Eyeriss: An Energy-Efficient Reconfigurable Accelerator for Deep Convolutional Neural Networks. IEEE J. Solid-State Circuits 52, 127–138 (2017)
29. García-Martín, E., Lavesson, N., Grahn, H., Casalicchio, E., Boeva, V.: How to Measure Energy Consumption in Machine Learning Algorithms. In: Alzate, C., Monreale, A., Assem, H., Bifet, A., Buda, T.S., Caglayan, B., Drury, B., García-Martín, E., Gavaldà, R., Koprinska, I. et al. (eds.) ECML PKDD 2018 Workshops, 11329, pp. 243–255. Springer International Publishing, Cham (2019)
30. Goel, B., McKee, S.A., Själander, M.: Techniques to Measure, Model, and Manage Power. In: Advances in Computers Volume 87, 87, pp. 7–54. Elsevier (2012)
31. D. Economou, S. Rivoire, C. Kozyrakis, P. Ranganathan: Full-System Power Analysis and Modeling for Server Environments. In: (2006)
32. Ermao Cai, Da-Cheng Juan, Dimitrios Stamoulis, Diana Marculescu: NeuralPower: Predict and Deploy Energy-Efficient Convolutional Neural Networks. In: Min-Ling

Zhang, Yung-Kyun Noh (eds.) Proceedings of the Ninth Asian Conference on Machine Learning, 77, pp. 622–637. PMLR (2017)
33. Kramer, K.J., Rokoss, A., Schmidt, M.: Do we really know the benefit of machine learning in production planning and control? A systematic review of industry case studies. Conference on production systems and logistics (2021)
34. Hevner, March, Park, Ram: Design Science in Information Systems Research. MIS Quarterly 28, 75 (2004)
35. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37, 337–355 (2013)
36. Rusche, S., Rockstuhl, S., Wenninger, S.: Quantifizierung unternehmerischer Nachhaltigkeit in der Fertigungsindustrie: Entwicklung eines zielorientierten Nachhaltigkeitsindex. Z Energiewirtsch (2021)
37. A. Greening, L., Greene, D.L., Difiglio, C.: Energy efficiency and consumption — the rebound effect — a survey. Energy Policy 28, 389–401 (2000)
38. Rockstuhl, S., Wenninger, S., Wiethe, C., Häckel, B.: Understanding the risk perception of energy efficiency investments: Investment perspective vs. energy bill perspective. Energy Policy 159, 112616 (2021)
39. Peffers, K., Tuunanen, T., Gengler, C.E., Rossi, M., Hui, W., Virtanen, V., Bragge, J.: Design Science Research Process: A Model for Producing and Presenting Information Systems Research. Proceedings of the First International Conference on Design Science Research in Information Systems and Technology (2006)
40. Himeur, Y., Alsalemi, A., Bensaali, F., Amira, A.: Building power consumption datasets: Survey, taxonomy and future directions. Energy and Buildings 227, 110404 (2020)
41. University of Lille: PyJoules, https://pypi.org/project/pyJoules/
42. Federal Ministry for Economic Affairs and Energy: Laptop oder PC – beim Stromverbrauch haben Sie die Wahl, https://www.deutschland-machts-effizient.de/KAENEF/Redaktion/DE/Standardartikel/Dossier/A-computer-energieeffizient.html
43. Geerling, J.: Power Consumption Benchmarks, https://www.pidramble.com/wiki/benchmarks/power-consumption
44. Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.-I.: Explainable AI for Trees: From Local Explanations to Global Understanding (2019)
45. Wirth, R., Hipp, J.: CRISP-DM: Towards a Standard Process Model for Data Mining (2000)