Power consumption and energy efficiency potentials in data centres: a case study from Switzerland

Jonas Müller
TEP Energy GmbH
Rotbuchstrasse 68
CH-8037 Zürich
Switzerland
jonas.mueller@tep-energy.ch

Martin Jakob
TEP Energy GmbH
Rotbuchstrasse 68
CH-8037 Zürich
Switzerland
martin.jakob@tep-energy.ch

Adrian Altenburger
Hochschule Luzern (HSLU)
Technikumstrasse 21
CH-6048 Horw
Switzerland
adrian.altenburger@hslu.ch

Keywords

data centres, information and communication technologies, power consumption, energy efficiency potential, market transformation

Abstract

Increasing digitalization and new trends such as Internet of Things (IoT), Big data, Industry 4.0, and cloud computing lead to a considerably higher demand for data storage and computing power in data centres (DC). Further trends like outsourcing and cloud-based applications are also boosting data centres dynamics. In 2018, power consumption in data centres accounted for approximately 2.7 % of the total power consumption in the EU, even more in Switzerland. At the same time, significant energy efficiency potential has been identified.

Energy efficiency measures in data centres exist for both, data centre infrastructure (cooling, ventilation, power consumption security) and IT infrastructure (storage technologies, backup systems, network, and other components). Despite the increasing importance of data centres, both in terms of energy and the economy, there is a lack of current empirical data on power consumption and existing energy efficiency potentials.

This article determines the power consumption and energy efficiency potential in Swiss data centres based on an extensive online survey. The most important parameters for estimating power consumption (installed capacity, power usage effectiveness, utilization) and energy efficiency measures (such as to increase the system temperature, adopting free cooling, using efficient storage and back-up systems), are collected from three DC segments: DC service providers, operators of in-house data centres and server rooms, and small and medium-sized enterprises. The results highlight that some efficiency potential has been tapped, partly because of promotion programs, especially on the DC infrastructure. Nevertheless, power consumption by DC steadily increased over the last years and accounts for about 3.6 % of the total power consumption in Switzerland. Moreover, there are still significant energy efficiency potentials of around 46 % of the total power consumption. Future demand is expected to increase even more, thus further monitoring and exploitation of remaining energy efficiency potential is crucial, particularly on the IT infrastructure.

Introduction

The worldwide demand for data storage and computing power is rapidly increasing due to new trends such as Internet of Things (IoT), Big data, Industry 4.0, and cloud computing. In 2018, power consumption in data centres accounted for approximately 2.7 % of the total power consumption in the EU (Montevecchi, F. et al., 2020), In Switzerland, data centres (DC) and server rooms (SR) are divided into the following two segments:

- DC service providers: offer services such as the rental of hardware (e.g. server space, servers) and/or IT services (e.g. software, storage, computing)
- In-house DC/SR: are operated for internal purposes of firms (private and public)

In the past, Switzerland has been shown to be a particular hotspot for data centres. According to ex-post analyses by the Swiss Federal Office of Energy (SFOE), the information and communication technology (ICT) sector accounts for more than 5 % of the power consumption in Switzerland (SFOE, 2020). More than half of this is due to server rooms and data centres, which already accounted for a share of 2.4 % to 3.3 % of the Swiss power consumption in 2014 (Altenburger et al., 2014). At the same time, significant energy efficiency (EE) potential has been identified in DC and SR (Puntsagdash et al., 2015). Due to an increasing use and shift to cloud-based services, the dynamics for data centres have also increased, which is expressed, e.g., by an increase in the number of data centres among data centre service providers from around 70 in 2017 to over 80 in 2020. Similarly, Swiss DC service space has also increased significantly in recent years (Netzmedien AG (2017, 2020)).

Despite the raising importance of DC, both in terms of energy and the economy, there is a lack of up-to-date empirical data on power consumption and energy efficiency. Thus, the objectives of the study are:

1. To collect statistical data from 2019 on the power consumption of data centres in Switzerland, divided into three segments: DC service providers, in-house data centres including server rooms, as well as server rooms of small and medium-sized enterprises (SME).
2. Identifying the remaining energy efficiency potential, also subdivided according to the various data centre segments.
3. Demonstrating the past development since the last study in 2014 (Altenburger et al., 2014) and the driving factors and, based on this, to present the development for the period 2020–2025.

Methodology

Power consumption and energy efficiency potentials in Swiss DC are determined empirically by adopting a sample and extrapolation approach. Therefore, a comprehensive online survey was developed for the subsequent analysis of power consumption and energy efficiency potentials.

SURVEY

The online survey was conducted in summer 2020. Three segments with differing questionnaires were defined for the online survey. The segmentation was carried out in collaboration with Profondia AG, an information service provider specialising in the ICT sector. The three segments are characterised as follows:

- Segment A (DC service provider): A full survey was aimed at because of the high energy significance of this segment and the lack of suitable estimators for extrapolation. For this segment, there was a close cooperation with Netzmedien AG, which published a market overview in 2020 (Netzmedien AG, 2020). This lists more than 80 data centres by name and size. The quantity structure of segment A essentially corresponds to this market overview, which is supplemented with providers from the Profondia database (Profondia, 2020).
- Segment B (In-house DC/SR): Are determined from the Profondia database, which results in a quantity structure of over 2,100 companies/organisations.
- Segment C (SME): Due to cost reasons, a geographical restriction for the sample was made. Lucerne as medium-sized Swiss canton with a heterogeneous economic structure is selected.

Within the online survey, information regarding power consumption, drivers of power consumption (e.g. installed capacity, area of DC space, number of servers), the quantitative indicator of energy efficiency (PUE) and further information on technical and operational characteristics of DCs and measures implemented to assess energy efficiency are collected.

ESTIMATION OF POWER CONSUMPTION

Approach

The approach for the estimation of the current power consumption of DC/SR in Switzerland (which includes the energy efficiency potential that has already been realised) is illustrated in Figure 1. The procedure for estimating power consumption is shown on the left in Figure 1. The power consumption of the sample is determined based on the data collected in the survey and by regression models with quantity-specific explanatory variables.

Using external data and further regression models with the explanatory variables number of employees, segment and economic sector, the power consumption is estimated for the remainder of the sample. Considering coverage levels and the assumption of an estimated number of unreported cases ("dark figure"), an extrapolation of the power consumption for the remaining power consumption of Swiss DC can be achieved. The individual shares for the estimation of power consumption are illustrated for each segment. The bar segments do not represent the true scale but are intended to illustrate the relative proportions.

Power consumption response sample

Depending on the data availability from the survey, one of the following methods is chosen to determine the power consumption of DCs from the response sample. For DC where the survey data does not allow the estimation via method 1 or 2, the power consumption is estimated using regression models (method 3).

Method 1

Calculated power consumption E_{DC} (MWh) based on installed IT capacity, *PUE* and utilization.

$$E_{DC} = P_{IT} \cdot PUE \cdot U \cdot 8{,}760\ h \qquad (1)$$

Where P_{IT} represents refers to the installed IT capacity (MW), U to the average utilization (%) of the IT components and *PUE* to the power usage effectiveness. The *PUE* expresses the ratio between the total energy consumed and the energy consumed specifically for IT. The closer the ratio approaches 1, the more energy-efficient the data centre is.

Method 2

Power consumption as indicated in the survey.

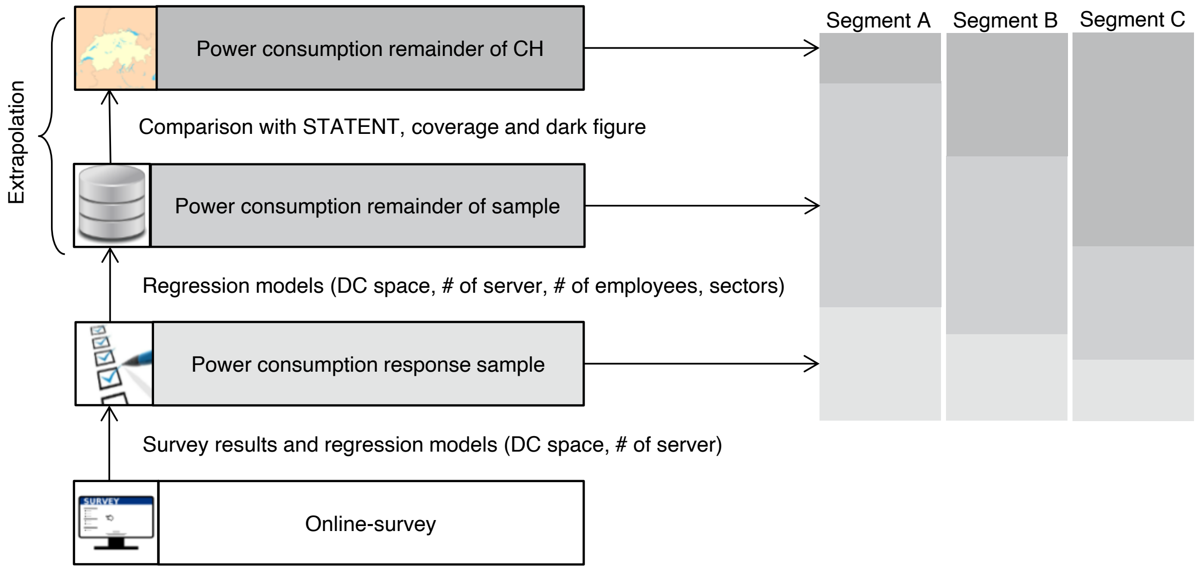

Figure 1. Approach to determine the power consumption of data centres and server rooms in Switzerland.

Method 3

Estimation of the power consumption by regression models with the number of physical servers PS_{DC}, the total area of DC space A_{DC} (m^2) and the segments A, B, and C as predictor variables.

$$\text{i. } \log(E_{DC}) = \beta_0 + \beta_1 \cdot \log(PS_{DC}) + \beta_2 \cdot \log(A_{DC}) + \beta_3 A + \beta_4 B + \beta_5 C \quad (2)$$

$\beta_0, \ldots, \beta_5$ represent the coefficients to be estimated and log indicates the natural logarithm.

Likewise, depending on data availability a simpler model is applied if only the total area of DC space or the number of physical servers was indicated in the survey (Jakob et al. 2021).

Power consumption remainder

The extrapolation approach of the power consumption is differentiated between DC service providers (A) and in-house DC (B/C).

For most of the DCs in segment A, data of the total DC space in sqm is available from the market survey (Netzmedien AG, 2020). For the mainly smaller data centre service providers, for which no DC space data is available from the market overview, power consumption is extrapolated using the attribute "number of employees" (EMP), which is available from the Profondia database. Since segment A is a special case regarding the number of employees, a separate model is specified for segment A only.

For segment A, a full survey is generally aimed for. This requires only an extrapolation to the Profondia database to consider the non-responding data centres. However, large public cloud providers such as Apple, Amazon, Google, Microsoft, etc. are increasingly entering the Swiss DC market. These providers not only want to cover the demand for server infrastructure, but also provide rental models for software use and data storage (Schröder M., 2019). Because some of the above-mentioned and other providers are neither included in the Profondia database nor in the market overview of network media, an estimated number of unreported cases (green bar in segment A) is assumed for segment A, based on expert opinion.

For segments B and C, no service area data is available from external data sources. For the extrapolation to all firms in the database, the attributes "number of employees" (*EMP*) and economic sector codes of the companies are used (see model ii.). The variables *BAN* (banks), *PHE* (public authorities/health/education), *IT* (information technology) and *R* (other sectors) represent dummy variables for the sector affiliation. Once again, the differences between segments B and C are reflected by the model.

$$\text{ii. } \log(E_{DC}) = \beta_0 + \beta_1 \cdot \log(EMP) + \beta_2 B + \beta_3 C + \beta_4 BAN + \beta_5 PHE + \beta_6 IT + \beta_7 R \quad (3)$$

A comparison of the employment figures of the selected segments from the Profondia database with the statistics on company structure (STATENT) of the Swiss federal statistical office reveals that total power consumption must be extrapolated for in-house DC/SR to determine the total power consumption (STATENT, 2018). Accordingly, a degree of coverage of the Profondia database is defined for the segments.

For segment C, the data from the canton of LU are further extrapolated to Switzerland using a regression model with the number of employees as predictor variable. Detailed information on the procedure and the assumed degree of coverage can be found in Jakob et al. (2021).

ESTIMATION OF ENERGY EFFICIENCY POTENTIAL

Power consumption and efficiency potential depend on the energy efficiency measures that have already been implemented and the state of the art in the various system areas of infrastructure and IT. Cooling, ventilation, systems for power supply security, etc., belong to the former, storage technologies, backup systems, utilisation of IT components, etc., belong to the latter.

Figure 2 schematically illustrates the energy efficiency of a data centre as a function of the PUE and the IT efficiency

(ITE) indicator. For this purpose, a target value is defined, which represents the efficiency potential that can be realised in practice.

- The theoretical energy efficiency potential on the infrastructure side is represented by the black outlined area, the feasible energy efficiency potential on the infrastructure side is represented by the blue outlined area. The latter illustrates the reduction of the PUE from the actual value (PUE_{act}) to the possible target value (PUE_{targ}), which differs depending on the segment (A: 1.15, B: 1.2).
- The directly feasible IT energy efficiency potential, which can be exploited by reducing the IT efficiency indicator is characterised by the green outlined area. The ITE quantifies the energy efficiency of the IT and ranges between 0.5 (very efficient) and 1 (not efficient). The indirect effect of IT measures also includes the dark green and violet areas (assuming a constant PUE). However, to avoid double counting, the purple part is only assigned to the infrastructure side. The dark green area is neglected, taking into account a slight increase in PUE if extensive efficiency measures are implemented on the IT side while the infrastructure is not or cannot be adapted accordingly.

As such, the efficiency potential (EP) is calculated for the infrastructure and IT systems:

$$EP = \left(1 - \frac{PUE_{targ}}{PUE_{act}}\right) \cdot E_{total} + \left(1 - \frac{ITE_{targ}}{ITE_{act}}\right) \cdot E_{IT} \qquad (4)$$

Where PUE_{act} and PUE_{targ} represent todays actual PUE of the DC and the target PUE which is feasible with implemented EE measures. E_{IT} and E_{total} show the power consumption for the IT systems and the whole DC, respectively. ITE_{act} and ITE_{targ} illustrate the actual and target IT efficiency indicator. Both PUE_{act} and ITE_{act} are estimated by the implemented energy efficiency measures. For this purpose, the power demand for the infrastructure and IT systems are divided into different components. Each component is multiplied by defined factors according to the implemented measure (more details are found in Jakob et al. (2021)).

Status quo energy efficiency in Swiss data centres

Energy efficiency can be characterised based on a quantitative measure such as the PUE or by the information about already implemented energy efficiency measures. The PUE represents a recognised measure of the infrastructural energy efficiency of a DC. A PUE close to 1 reveals a lower amount of power for the infrastructure compared to the total power consumption, hence, a more efficient DC.

The question about the PUE was only asked in segments A and B (Figure 3, left). Most DC operated by service providers (segment A) indicate a PUE between 1.2 and 1.5. Compared to earlier studies, this already shows a strong improvement in energy efficiency: for 2013, a PUE of 1.4 to 1.8 (best guess 1.6) was assumed for this segment (Altenburger et al., 2014). This efficiency increase is in line with the implemented efficiency measures (cf. Figure 4). For in-house data centres, the distribution of responses is similar, although the proportion of very energy-efficient data centres (PUE<1.2) is smaller. For comparison: Altenburger et al. (2014) assumed PUEs between 1.4 and 2.1 (best guess 1.82) for internal data centres in Switzerland. The large proportion of in-house data centres or SRs for which the PUE is not known, or no information was provided is striking. The reasons for this can be manifold. On the one hand, there might be a perception that it is technically not easy to measure the PUE value correctly; on the other hand, the incentive for communicating the PUE may also be low. Thirdly, it might express a lack of awareness to monitor this important indicator as a first step to improve energy efficiency.

Besides the PUE, many key indicators for energy efficiency were asked in the survey. Raising system temperatures enables large energy savings, while system reliability is usually not or only little affected (El-Sayed et al, 2012). The system temperature is particularly important for free cooling, since it can significantly reduce chiller running times. High system temperatures are also a great advantage for waste heat recovery.

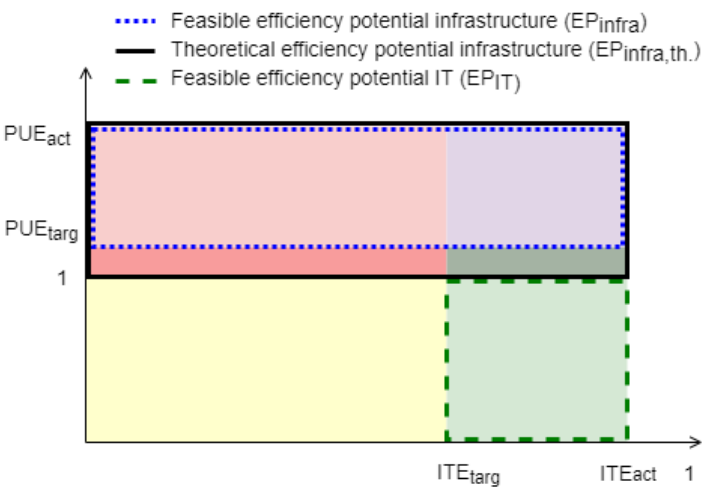

Figure 2. Schematic representation of the energy efficiency potential as a function of the PUE and the IT efficiency (ITE) indicator.

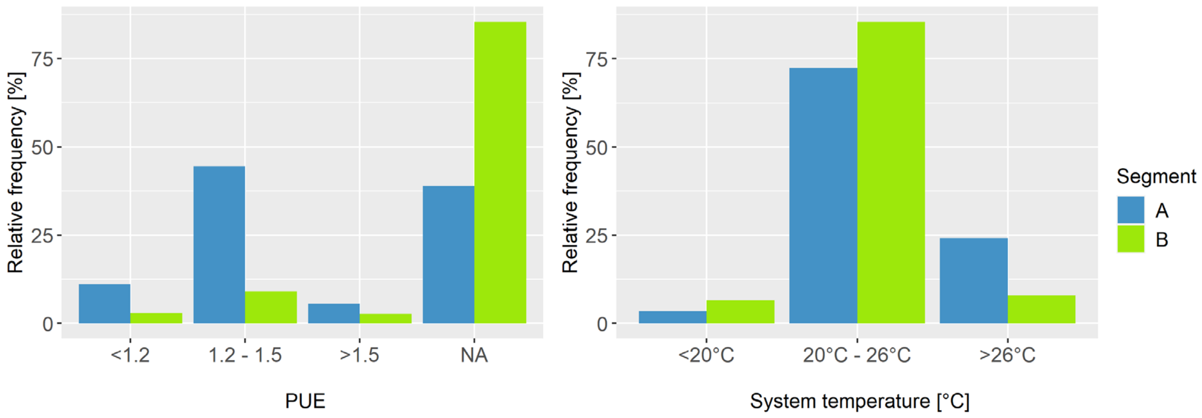

Figure 3. Left: PUE values (left, n = 22 (A) and 92 (B), without NAs). Right: System temperature (n = 29 (A), 517 (B)).

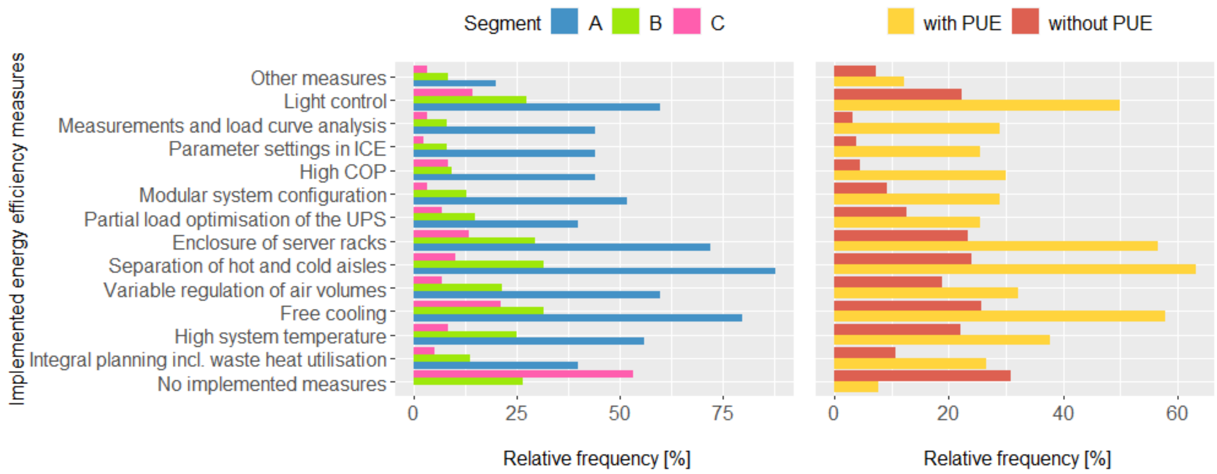

Figure 4. Left: Implemented energy efficiency measures in different segments (n = 29/574/136), Right: Implemented energy efficiency measures in segment B with indicated PUE and without indicated PUE.

The distribution of the system temperatures at the entrance of the servers in segments A and B are illustrated in the right panel of Figure 3. Significantly higher system temperatures are already accepted by DC service providers, with a total of almost 25 % above 26 °C. Nevertheless, even in segment A, most data centres are still operated with system temperatures between 20 °C and 26 °C.

A separate consideration of the segments is important when looking at implemented EE measures. In the case of data centre service providers, fully comprehensive energy efficiency measures are implemented significantly more often (Figure 4, left). Measures that are particularly relevant in terms of energy, such as separating the hot and cold aisles or enclosing server racks, raising system temperatures, variable regulation of the air volumes and use of free cooling, have already been implemented in many DCs of segment A. Contrarily, the number of data centres in which none of the measures have been implemented amounts to about 25 % in Segment B and 50 % in Segment C, while no feedback was received from the data centre service providers in this regard.

The right panel in Figure 4 illustrates relative frequencies of implemented measures for DCs where a PUE was given in the survey vs. DCs where no PUE was given. For all measures DCs with indicated PUE show a substantially higher frequency of implementation.

Further detailed information on many other energy-relevant data (cooling type, cooling water temperature etc.) was collected by the survey and enables a comprehensive overview of Swiss data centres. Detailed descriptive statistics can be found in Jakob et al. (2021).

Power consumption in Swiss data centres

STATISTICAL ANALYSIS

Based on the survey data, power consumption is determined for more than 360 DC. For further 270 DC power consumption is estimated by regression model i. Table 8 shows the model output of model i. As expected, both the area of DC space (A_{DC}) and the number of servers (PS_{DC}) have a strong, positive influence on the power consumption of the data centres. However, the relation substantially differs between the segments: With the same DC space and the same number of servers, the power consumption of DCs in segments B and C is significantly lower.

DC space and number of servers are familiar to most data centre service providers and operators. The model output shows that these are reliable parameters which are useful to also collect and monitor in the future.

Figure 5 illustrates the model from Table 1 (straight lines) compared to the underlying data (points). The power consumption data are plotted as a function of area. The colour of the points indicates the allocation to segments A, B and C and the size of the data points reflects the number of servers as another dimension. The model lines were plotted for the four figures a-d with a constant number of servers in each case, thus the cross-comparison between the four figures reveals the influence of the number of servers. As the number of servers increases (a to d), the straight lines shift vertically upwards, which allows the model to represent other size classes (in terms of number of server). For segment C, figure a) is relevant, as these usually have fewer than 10 servers. Regarding segment B, figures a) and b) are relevant. Figs. c) and d) show the significantly higher value for segment A.

POWER CONSUMPTION

For extrapolation reasons, three scenarios are distinguished to demonstrate the effect of uncertainty and in particular, the possible influence of a presumed number of unreported cases. The degree of coverage of the current results thus varies from low (scenario 1) to high (scenario 3), whereby a low degree of coverage corresponds to a higher number of unreported cases. Figure 6 shows the final extrapolated power consumption for server rooms and data centres in Switzerland for the scenarios 1–3 differentiated by segment.

Large consumers from segment B (allocated to segment B*) represent rare special individual cases, which nevertheless have

Table 1. Model output of model i. for the estimation of log(power consumption).

Variable	Coefficient	Std. error	t-value	p-value
Intercept	0.91	0.56	1.64	1.01e-01
$\log(A_{DC})$	0.41	0.11	3.88	1.28e-04***
$\log(PS_{DC})$	0.61	0.08	7.40	1.27e-12***
Segment B	-0.92	0.44	-2.12	3.46e-02*
Segment C	-1.19	0.50	-2.37	1.87e-02*
RSE: 1.53, R2: 0.61, Adj. R2: 0.61			*** 0.001, ** 0.01, *0.05	

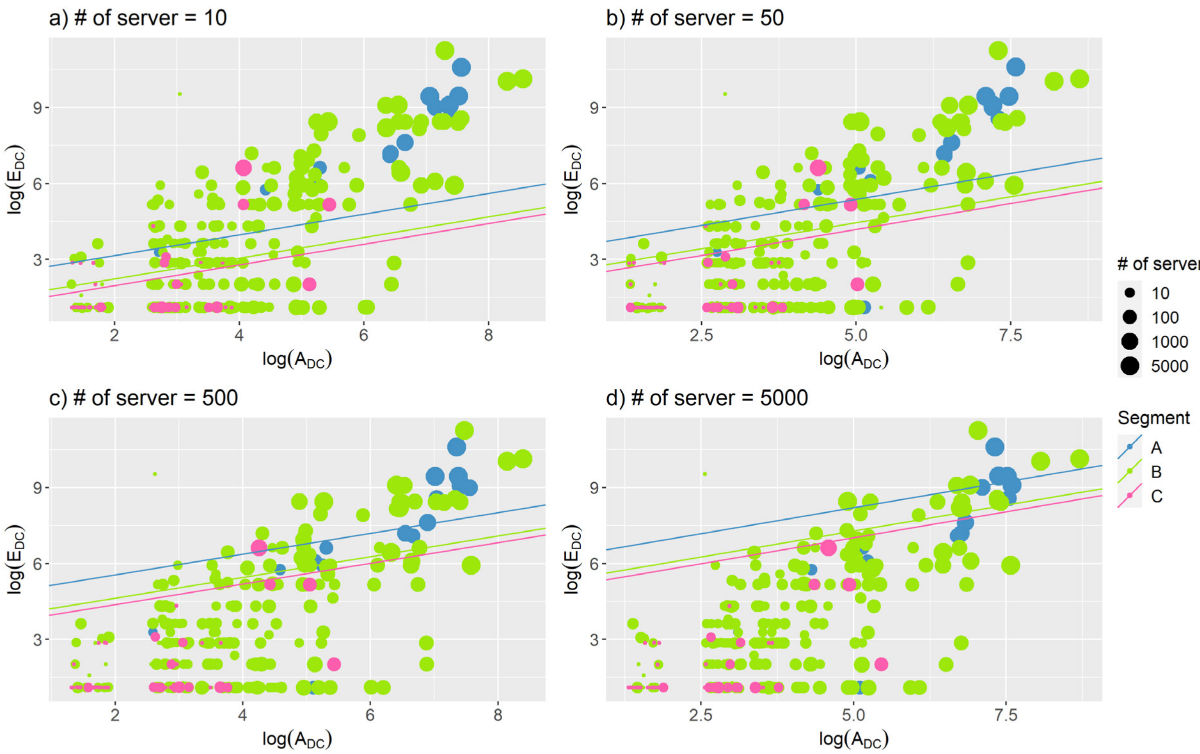

Figure 5. Log(power consumption in MWh) as a function of DC space (A_{DC}), number of servers and segment using a log-log approach according to model i. The size of the data points reflects the number of servers, whereby the number of servers was kept constant for each figures a)–d) to illustrate the model slope (according to the labelling of the graph).

a high energy relevance. The 5 largest consumers each have a power consumption of more than 20 GWh p.a. and account for a share of approximately 60 % in the entire segment B (incl. B*). However, the regression approach cannot consider such extreme values with simultaneously small number of cases. Accordingly, an individual case analysis is necessary for these large consumers. For this purpose, estimated values are imputed for the largest companies from the financial and insurance sectors, for pharmaceuticals and hospitals, TV and telecommunication as well as for a large research institution, based on expert estimates and findings from the survey.

The power consumption for segment A varies between almost 700 GWh (scenario 3) and around 900 GWh (scenario 1), depending on the dark figure. The two segments B and B* together contribute to about 1,075 GWh (scenario 2) of power consumption, with the value varying between 960 GWh and 1,220 GWh depending on the scenario. The extrapolated demand in segment C and scenario 2 amounts to about 230 GWh.

Accordingly, power consumption across all segments varies between 1.85 TWh and 2.37 TWh, with the medium scenario showing power consumption of around 2.09 TWh, which corresponds to around 3.6 % of Swiss electricity consumption in 2019 (57.2 TWh).

Energy efficiency potential in Swiss data centres

Similar to power consumption, scenarios are defined to illustrate the impact of uncertainty on the estimated EE potential, since the estimation of the PUE and ITE are strongly dependent on the survey data regarding energy efficiency measures. However, to also consider participants who did not answer all the questions, different values are imputed for the missing answers depending on the scenario:

1. Scenario ”Median”: The median from the responses received is used in the case of missing data. Due to the very high number of DCs that did not provide any information on PUE and with reference to the PUE figure as a function of the measures implemented (see Figure 4, right), it can be assumed that the median value is distorted in comparison to the population.
2. Scenario “Normal case”: Missing answers are replaced with normal case according to previous observations and expert knowledge.
3. Scenario “Worst case”: Missing answers are replaced with worst case. This corresponds to the hypothetical case in which these DCs have not yet implemented any measures.
4. Scenario “Best case”: Missing answers are replaced with best case. This corresponds to a data centre that has already implemented many measures and can therefore already show a lower PUE.

For each measure, a best, normal and worst case is defined based on previous studies (Puntsagdash et al., 2014, Altenburger et al., 2015). For more information see Jakob et al. (2021).

This results in different efficiency potentials for power consumption depending on the scenario, which are shown in Figure 7. The estimation reveals a slightly lower energy efficiency potential of the infrastructure systems compared to the IT systems. Depending on the scenario, the infrastructural energy

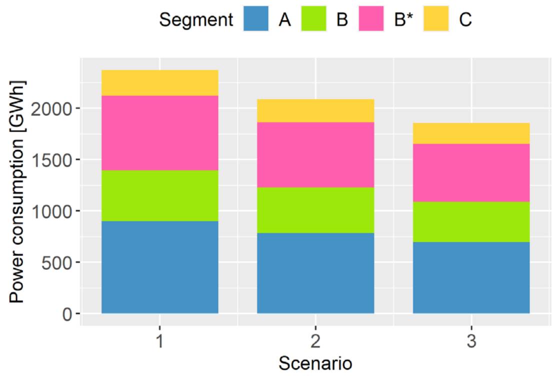

Figure 6. Estimated power consumption for the different segments and scenarios 1–3.

efficiency potentials vary between 350 GWh (Best case: most efficiency measures already implemented) and around 500 GW (Worst case: few efficiency measures implemented). The “Best guess” scenario results in a total energy efficiency potential of about 410 GWh, which corresponds to around 20 % of the total power consumption (scenario 2 in Figure 6).

For the first three scenarios, the absolute energy efficiency potential of the IT is in a similar range as for the DC infrastructure. However, the differences between the scenarios are greater than on the infrastructure side. This is, among other things, due to the fact, that not all variables were completely surveyed. The “Best case” scenario is rather unlikely. For the “Best guess” (normal case scenario), the total power-related efficiency potential is about 550 GWh, thus, around 26 % of the total power consumption of the DCs and SR. Based on the “Best guess” scenario the overall the energy efficiency potential is estimated to about 46 % of the total power consumption.

Discussion

PAST DEVELOPMENT IN THE DATA CENTRE MARKET

The focus of the online survey conducted in this work was on the status quo in 2019. Thus, no questions concerning past developments of power consumption were asked. Nevertheless, a comparison with a previous study (Puntsagdash et al., 2014), which estimated the power consumption in 2013, allows to speculate concerning the developments in the Swiss DC market within the last 6 years.

The results from the online survey and the estimation of the power consumption reveal a moderate increase in power consumption of around 25 % since the previous study for the year 2013 (Puntsagdash et al., 2014). However, due to the different methodology and the different data, these values cannot be directly compared, and the influencing factors illustrated in Figure 8 must be considered. The bars refer to the influencing factor and the development over time since 2013, i.e., the estimated power consumption for 2019 is e.g., 227 GWh higher than in 2013 due to the extension of the system boundary (consideration of segment C). By contrast, power consumption has decreased by 161 GWh due to an infrastructural efficiency increase. Uncertainties are expected in both studies, which are at least partially represented by the orange bar. The factors can be

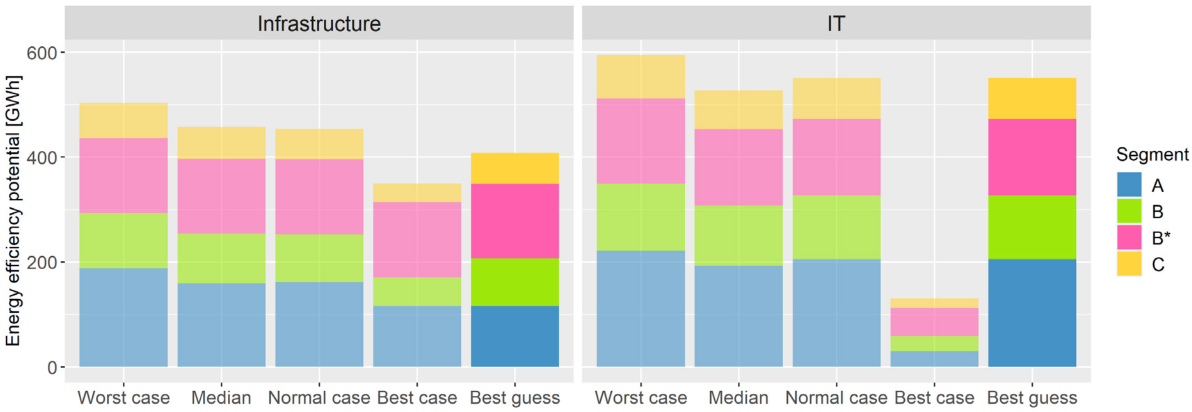

Figure 7. Estimated energy efficiency potential of infrastructure systems (left) and IT systems (right).

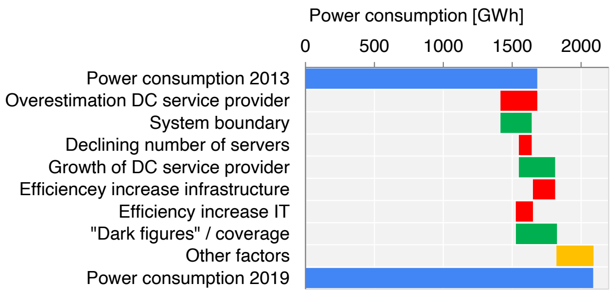

Figure 8. Increase of power consumption between 2013 and 2019 and the explanatory factors. Green and orange bars indicate an increase, red bars a decrease in power consumption.

divided into methodological (overestimation DC service provider, system boundary, dark figures/coverage) and explanatory (remaining factors) factors. Neglecting the methodological differences, the adjusted real development growth accounts to 8 %.

FUTURE DEVELOPMENT IN THE DATA CENTRE MARKET

Energy efficiency in data centres will remain an important topic in the future. On the one hand, the share of total power consumption is substantial (3.6 % of Swiss power consumption); on the other hand, there is further potential for energy efficiency. Simple technical and operational measures can often reduce power consumption significantly and thereby also operating costs.

Drivers of power consumption

Future analyses are important in view of the further expected dynamics in the data centre market. Large public cloud providers are pushing into the Swiss market. Microsoft, for example, is present in Switzerland since 2019 (Schröder, 2019). Amazon is not only a front-runner in global online commerce, but also dominates the DC market, and aims to open several data centres for its cloud subsidiary Amazon Web Services from 2022. Customers will benefit from data storage on Swiss servers and shorter latency (Ott, 2020). Oracle also opened its first data centre in Switzerland in 2019 (Schröder, 2019). According to a recently published article, Google has also been operating its own data centres in Switzerland since 2019 (Huwiler, 2021). Alibaba Cloud already operates its own data centres in Europe, focusing primarily on Chinese customers who are active in Europe (Schröder, 2019).

The impact of these increasingly large data centre service providers on overall power consumption in Switzerland is uncertain. To date, a significant amount of private data (photos, videos, social media, etc) is stored abroad. Large providers such as Apple, Microsoft or Google might enter the Swiss market to supply data and IT service needs and possibly even to serve end customers abroad. If such an export market emerges, a significantly faster increase in power consumption could occur in the coming years. The online survey also shows that more than half of the data centre service providers already serve customers from abroad.

In addition to the penetration of large public cloud providers and the resulting development of additional customer groups (abroad, private customers), the expected increase in power consumption in the coming years is also likely to be driven by an increasing demand. Big data, Internet of Things (IoT), Industry 4.0, etc. are likely to be the main drivers for the increased demand for data storage and computing power. A quantification of these drivers was not the subject of this work but could be conducted in a subsequent study.

Dynamics in Swiss DC market

The possible emergence of new providers and the expected growth in new usable space and IT services were also surveyed as part of another SFOE project (7pro solution, 2020). Data centre operators, data centre planners, engineering firms, power supply companies, etc. were contacted by telephone and asked about data centre construction and expansion projects. The estimates show that a significant increase in IT performance from data centres can be expected in the next few years. A significant number of large projects are already planned, particularly around Zurich, and many feasibility studies are also being carried out in the region of Lake Geneva. Some of these projects are already publicly known (Netzmedien AG, 2020).

If all these projects were to be realised, without this being due to a further significant shift from in-house data centres to service providers, and if efficiency considerations were not given due attention, the power consumption of data centres in Switzerland could increase from around 2 TWh today to around 3 to 4 TWh in the coming years. However, it remains to be seen which projects will be realised. Recent years have also shown that certain data centres have been closed again (cf. Netzmedien AG, 2017 and Netzmedien AG, 2020). In addition, there is considerable efficiency potential in both existing and new data centres. Based on these and other considerations (see next section) we expect the medium-term power consumption to be between 2.7 and 3.5 TWh.

Switzerland remains an attractive location for large data centre service providers, particularly due to political stability, qualified staff, secure electricity supply and the central location in Europe. However, Switzerland is in international competition. Other regions such as Scandinavia or Ireland are favoured by the climatic conditions, the connection to transatlantic fibre-optic links, low electricity prices or tax advantages.

Despite other popular locations in Europe, a further comparison with an EU-wide study shows that power consumption of DC in Switzerland is still above average. This is, additionally the above-mentioned reasons, related to the rapidly advancing digitalization in Switzerland (Dib, 2019). In the period from 2015 to 2018, power consumption in DCs in the EU has increased from 2.3 % to 2.7 % compared to total power consumption. The study examined DC/SR with at least 3 physical servers, which is comparable to the system boundary in this study. By 2030, an EU-wide share of 3.2 % is expected (Montevecchi, F. et al., 2020), a value already exceeded in Switzerland. The specific power consumption in DC/SR per capita in Switzerland (0.25 MWh/capita in 2019) is almost 40 % higher than in the EU (0.18 MWh/capita in 2018).

Component level development

Besides emerging DC service providers and increasing computation and storage demand, further reasons for an increasing power consumption may be due to technical limits.

Diving into the chip level reveals that the point at which technical advances (more efficient chips) can compensate for rising power consumption will soon be reached. Moore's law (doubling the number of transistors on integrated circuits every 18 to 24 months) and Dennard's scaling (exponential improvements in energy efficiency at circuit level) have slowed down (Moore) or lost their validity (Dennard). While it has been possible to reduce the size of chip components over decades, the limits are now being reached for silicon-based elements. According to expert estimates, energy efficiency can no longer compensate for the increase in power consumption due to the faster rise in data and computing requirements.

Nevertheless, progress can be expected in certain areas by producing chips designed for specific applications. Graphics processing units (GPUs) allow significant speed improvements through parallel computing, which help emerging disciplines such as machine learning or artificial intelligence. In recent years, classical CPUs have not been able to keep up with these speed improvements (Lee, 2020). Accelerators for artificial intelligence (AI) offer a further improvement over GPUs and allow speed improvements for processes in neural networks. Google and Microsoft, for example, are developing their own AI chips for their specific needs. However, for everyday software, including virtual machines, operating systems, data management or cloud applications from third-party providers in data centres, no specific chip solutions are expected soon. Insights from basic research are missing to improve general-purpose chips (Rotman, 2020). Therefore, none of the above-mentioned technologies will help to fully compensate for the exponentially increasing data and computational demand in the coming decades. Accordingly, this can only be covered by an expansion of the installed electrical power, i.e., by new and larger DCs.

Consequently, not only in Switzerland power consumption in data centres is expected to increase more strongly in the medium term than in recent years. The implementation of energy efficiency measures in existing data centres is therefore of enormous importance to at least cushion the expected increase in computing power in the coming years. When designing, planning, and realising new data centres, it is also crucial to configure them adequately, to select the right components and to optimize the processes. With these approaches, an energy-efficient construction and operation of data centres is feasible.

Conclusion and recommendations

This study provides a comprehensive empirical basis for many fundamental energy related parameters in modern data centres. The results of this study show that power consumption in DC/SR accounts for between 3.3 % and 4.1 % (Best guess: 3.6 %) of the total power consumption in Switzerland. This implies that power consumption has again increased moderately in the last 5 years, whereby similar trends are also evident in the EU.

At the same time, efforts for more energy efficiency in data centres have already had an effect and the PUE values could be reduced compared to earlier studies. The survey results also manifest this development through the measures already implemented, especially by DC service providers. Nevertheless, a considerable energy efficiency potential of around 46 % of the total power consumption remains, which is to be further exploited by measures considered in this study.

Future studies should consider other factors influencing energy efficiency in the area of servers and networks (e.g. quality of power supply, intelligent switches), which have not been part of this work. Future studies should also provide more information about the DC service providers customers. If many large public cloud providers enter the Swiss market and possibly even serve customers from abroad with storage services and computing power, power consumption may increase rapidly.

We recommend that based on the findings from this study and from earlier work (Jakob et al., 2015, Altenburger et al., 2015) lists of measures (e.g. EnergieSchweiz, 2020) should be further developed and disseminated. This may especially help, to foster energy efficiency for in-house DC. Such findings can be used at various levels, whereby action must be taken quickly due to the larger number of new DC projects in the pipeline and the expected increase in computing demand. In the Swiss context, the following findings are helpful:

- to inform, educate and train planners, investors, and operators of new and existing data centres,
- for the federal government to support cantons and municipalities within the framework of existing legislation, e.g. in the issuing of building permits or in the implementation of the large-scale consumer article,

- for the PUEDA+ funding programme within the ProKilowatt framework,
- to support voluntary approaches such as the efficiency labels by the SDEA.

To be able to use residual heat from data centres for the decarbonization of the building sector, cooperation with the municipalities is particularly important in the planning of thermal networks. Here, the concept of PUE^{DA} of the PUEDA programme can be used (Jakob et al. 2015).

In the future, further voluntary measures as well as normative bases should be considered, whereby cooperation with the industry should take place on the infrastructure and IT side. Further, such basic principles can also serve to shape new building regulations.

To monitor future developments in the DC market, external data sources used for this study (Profondia, 2020, Netzmedien, 2020) should also be available in the future. Additionally, the authors suggest that the empirical data basis created here should be updated periodically in the coming years. Monitoring of the power consumption as well as the implemented energy efficiency measures will be beneficial for further observations of the market. Similar approaches could be used in other European countries or regions to evaluate power consumption and EE potentials in the highly dynamic market of data centres.

References

Altenburger, A., Hauri, D., Puntsagdash, G., & Deininger, S. (2014). Rechenzentren in der Schweiz – Energieeffizienz: Stromverbrauch und Effizienzpotenzial. (*Data centres in Switzerland – Energy efficiency: Power consumption and energy efficiency potential.*) IWSB und Amstein+Walthert i.A. ASUT und Swiss Federal Office of Energy. Bern.

SFOE (2020). Analyse des schweizerischen Energieverbrauchs 2000–2019 nach Verwendungszwecken. (*Analysis of Swiss energy consumption 2000–2019 by specific use.*) Infras, Prognos, TEP i.A. Swiss Federal Office of Energy. Berne.

Dib, G. (2019). 2019 Enabling Digitalization Index. Euler Hermes Economic Research.

El-Sayed, N., Stefanovici, I. A., Amvrosiadis, G., Hwang, A. A., & Schroeder, B. (2012, June). Temperature management in data centers: Why some (might) like it hot. In Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems (pp. 163–174).

EnergieSchweiz. (2020). Massnahmenkatalog für Rechenzentren. (*List of measures for data centres.)* Retrieved February 15, 2021, from https://dcenergy.ch/actions/.

Huwiler, S. (2021). Gross und verschwiegen: Die "Hyperscaler" kommen in die Schweiz. (*Big and secretive: The hyperscalers are entering Switzerland.*) Tages Anzeiger.

Jakob, M., Forster, R., & Altenburger, A. (2015). Wettbewerbliche Ausschreibungen für Effizienzmassnahmen im Elektrizitätsbereich. Abschlussbericht Programm 2-Pg125 LowEx-Rechenzentrum. Amstein+Walthert, TEP Energy, Jobst Willers Engineering.

Jakob, M., Müller, J., Altenburger, A. (2021). Rechenzentren in der Schweiz – Stromverbrauch und Effizienzpotenzial. (*Data centres in Switzerland – Power consumption and efficiency potential.*) TEP Energy, HSLU i.A. SFOE.

Lee, K. C. (2020). Parallel Computing – Upgrade your data Science with a GPU. Retrieved January 12, 2020, from https://towardsdatascience.com/parallel-computing-upgrade-your-data-science-with-a-gpu-bba1cc007c24.

Montevecchi, F., Stickler, T., Hintemann, R., Hinterholzer, S. (2020). Energy-efficient Cloud Computing Technologies and Policies for an Eco-friendly Cloud Market. Final Study Report. Vienna. Retrieved from https://ec.europa.eu/digital-single-market/en/news/energy-efficient-cloud-computing-technologies-and-policies-eco-friendly-cloud-market.

Netzmedien AG. (2017). Marktübersicht Schweizer Rechenzentren 2017. (*Market overview Swiss data centres 2017*.)

Netzmedien AG. (2020). Marktübersicht Schweizer Rechenzentren 2020. (*Market overview Swiss data centres 2020*.)

Ott, A. (2020). Amazon plant Rechenzentren in der Schweiz. (*Amazon plans data centres in Switzerland*.) Retrieved November 5, 2020, from https://www.srf.ch/news/schweiz/geschaeft-mit-daten-amazon-plant-rechenzentren-in-der-schweiz.

Profondia AG. (2020). ICT Database.

Puntsagdash, G., Hass, O., Bersier, R. L., Zwygart, M., & Blättler, P. (2015). Studie zur Stromeffizienz bei Rechenzentren in der Schweiz – Potentialanalyse und Massnahmenkatalog für Effizienzsteigerung IT-Hardware und Infrastruktur. (*Studie zur Stromeffizienz bei Rechenzentren in der Schweiz – Potentialanalyse und Massnahmenkatalog für Effizienzsteigerung IT-Hardware und Infrastruktur*.) Amstein+Walthert (A+W) i.A. Swiss Federal Office of Energy. Berne.

Rotman, D. (2020). We're not prepared for the end of Moore's Law. Retrieved January 12, 2020, from https://www.technologyreview.com/2020/02/24/905789/were-not-prepared-for-the-end-of-moores-law/.

Schröder, M. (2019). Die Hyperscaler in der Schweiz. (*The hyperscalers in Switzerland.*) Retrieved November 4, 2020, from https://www.computerworld.ch/business/cloud/hyperscaler-in-schweiz-1715605.html.

STATENT. (2018). Arbeitsstätten und Beschäftigte nach Kanton und Grössenklasse. (*Workplaces and employees by canton and size class*.) Retrieved October 27, 2020, from https://www.pxweb.bfs.admin.ch/pxweb/de/px-x-0602010000_109/px-x-0602010000_109/px-x-0602010000_109.px.

7pro solution. (2020). Bericht RZ Trends CH – Umfrage und Erhebung der Trends bei den Rechenzentren in der Schweiz (Bau- und Ausbaupläne/Leistungen) in den nächsten 5 Jahren. (*Report DC trends CH – Survey of data centre trends in Switzerland (construction and expansion plans/services) in the next 5 years:* unpublished).

Acknowledgement

This project was commissioned by SwissEnergy, the SFOE program that supports voluntary measures for the implementation of the Swiss energy policy.