International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 368

DATACENTERS AS PROSUMERS IN URBAN ENERGY SYSTEM: A REVIEW

Pei Huang[1], Xingxing Zhang[1*], Isabelle Löfgren[1], Mats Rönnelid[1], Jan Fahlen[2], Dan Andersson[2], Mikael Svanfeldt[2]
[1] Department of Energy and Built Environment, Dalarna University, Falun, 79188, Sweden
[2] EcoDataCenter, Falun, 79170, Sweden
Corresponding Author: xza@du.se; +46 (0) 23-77 87 89

ABSTRACT

As a large energy prosumer in the urban energy systems, on the one hand, datacenters consume a large amount of electricity to ensure the IT facilities and ancillary power supply and cooling systems work properly; on the other hand, datacenters produce a large amount of waste heat due to the high heat dissipation rates of the IT facilities. To date, a systematic review of datacenters from the perspective of energy prosumers, which considers both integration of the upstream green energy supply and downstream waste heat reuse, is still lacking. This study fills in this gap and provides such a review. By providing a full picture of datacenters in the urban energy systems, this study aims to search new opportunities for improving datacenter overall energy efficiency and reducing carbon emissions.

Keywords: Data center, review, urban energy system, renewable/green energy, waste heat, energy efficiency

1. INTRODUCTION

The rapid increase of needs for data processing, data storage and digital telecommunications has led to dramatic increase in the datacenter industry [1]. Datacenters are large energy end-users. The statistics indicate that datacenters now consume about 3 % of the global electricity supply and account for about 2 % of total greenhouse gas (GHG) emissions. A recent report of 2018 from Cushman & Wakefield predicted annual growth of data centers to be 12-14% over the next two to five years. This will directly result in 1/5 of Earth's electricity consumption by 2025.

On the other hand, datacenters are trying to curb their carbon footprint by increasing the share of renewable energy they use. For instance, Sheme et al. investigated the feasibility of using renewable energy to supply datacenters in 60° north latitude [2]. To maximize the utilization of renewable energy, advanced demand control methods and schedulers have been developed. For example, Aksanli et al. developed a datacenter demand response strategy, which cancels or reschedules jobs whenever the instant green energy availability is low [3].

Nearly half of the datacenters' consumed energy is used in the cooling systems to deliver the heat of the electronic facilities to the atmosphere, creating waste heat. The large amount of waste heat represents great opportunities for energy saving and environmental protection. Many studies have been conducted to explore new systems and ways to recover and reuse the waste heat from datacenters. For instance, Oró et al. analyzed the energy and economic feasibility of applying air-cooled datacenter waste heat in district heating [4].

However, most of the existing studies analyze the datacenter from the sole perspective of either energy consumers or energy producers. Due to a lack of global optimization, the potentials in datacenters' performance improvements are limited. Thus, this study conducts a systematic review of datacenters with both the upstream and downstream systems. By providing a full picture of datacenters in the urban energy systems, this study aims to seek new opportunities for improving datacenter overall energy efficiency and reducing carbon emissions.

2. DATACENTER OVERVIEWS

2.1 Major IT components

A datacenter is a repository for data and information storage, management, and dissemination organized around a particular body of knowledge or related to a particular business [1]. A datacenter typically includes the following IT components.

- Servers, which store, analyze and transmit enormous amounts of data, are the major components of a datacenter.

- Storage devices are the hardware capable of holding information either temporarily or permanently.
- Switches are connecting devices that connect multiple devices for data and information exchanges.
- Cables are used for connecting the servers, storages, and switches to form a network.
- Racks are the components that support various components such as servers, storage devices, and switches for easy management.

2.2 Environmental requirements

Datacenters contain large amounts of IT equipment. A poor control of the indoor environment can cause low computing efficiency or even severe faults/outage. Table 1 summarizes the four classes of thermal environment specified by the 2015 ASHRAE thermal guidelines.

Table 1 2015 ASHRAE thermal guidelines for datacenters [5].

Class	Dry-bulb temperature	Humidity range	Maximum Dew Point
Recommended (suitable to all four classes)			
A1 to A4	18 to 27 °C	-9 °C DP to 15 °C DP (60% rh)	
Allowable			
A1	15 to 32 °C	-12 °C DP (8% rh) to 17 °C DP (80% rh)	17 °C
A2	10 to 35 °C	-12 °C DP (8% rh) to 21 °C DP (80% rh)	21 °C
A3	5 to 40 °C	-12 °C DP (8% rh) to 24 °C DP (85% rh)	24 °C
A4	5 to 45 °C	-12 °C DP (8% rh) to 24 °C DP (9s0% rh)	24 °C

2.3 Heat dissipation rates of components

There are significant differences in temperatures between different electronic components held in the IT server racks inside the datacenters. Consequently, the heat dissipation rates of different electronic components are different. A summary of the heat and temperature distribution of different components in the server is presented in Table 2. Different server types have different waste heat temperatures and heat densities.

Table 2 Summary of heat and temperatures in IT servers [6].

	Component	Proportion of total heat	Temperature
For standard server	Microprocessors	30%	85°C
	DC/DC conversion	10%	50°C
	I/O processor	3%	40°C
	AC/DC conversion	25%	55°C
	Memory chips	11%	70°C
	Fans	9%	30°C
	Disk drives	6%	45°C
	Motherboard	3%	40°C
For high performance cluster (HPC)	Microprocessors	63%	85°C
	DC/DC conversion	13%	115°C
	I/O processor	10%	100°C
	Memory chips	14%	40°C

3. COOLING SYSTEMS IN DATACENTERS

Due to the large variance in datacenter heat dissipation rates, different cooling techniques have been developed to satisfy the different cooling needs.

3.1 Air-cooled systems

Air-cooled systems make up most of the cooling systems in existing datacenters. There are four typical configurations in air-cooled datacenters [4].

- Computer room air conditioner units (CRAC): CRAC is usually applied in datacenters with low heat dissipation rates (<100kW). In CRAC systems, the warm air is cooled by a direct expansion unit. The cold air flows into the server racks from the cold aisles, after being heated up by the server racks, the warm air exits from the hot aisles [7].
- Computer room air handler units (CRAH): CRAH is commonly applied in datacenters of medium and bigger sizes (>100kW), due to its lower operational costs, compared with the CRAC [7]. The heated air from the IT equipment room is cooled by chilled water, and the return chilled water is cooled by a vapor compression chiller.
- In-Row cooling: In-row cooling is typically implemented for medium-high heat dissipation density (>10kW per rack) datacenters [4]. The liquid-to-air heat exchangers are placed between the server cabinets for more efficient cooling. The warm air exhausts directly from the hot aisles and is cooled by chilled water in heat exchangers. After that, the cool air is delivered to the cold aisles.
- Rear door cooling: The rear door cooling is mostly used in datacenters with high-density racks (>35kW) [4]. It is based on placing a liquid-to-air heat exchanger behind the server racks. The hot air is forced to flow through the heat exchanger by fans to be cooled. In this system, the room air has constant temperature and thus hot spots are avoided.

3.2 Water-cooled systems

To satisfy the increasing high heat removal needs, water-cooled systems have been developed. Water-cooled systems remove heat by direct contact with the server electronic components. Since water has a higher heat carrying capacity than air and higher convective heat transfer coefficients, direct contact with server components can produce higher heat transfer rates [8]. This allows low temperature differences between the cooling liquid and the server components, and thus liquid coolant with significantly high temperature can be used.

Substantial energy savings can be achieved in liquid cooling. Another benefit is the higher temperatures of the return water, which produces high quality waste heat to be easily recovered for reuse.

3.3 Two-phase cooled systems

The scales and power densities have exceeded 1kW/cm^2 in some datacenters. Two-phase cooling has been developed to deal with such high density heat dissipation. It cools the racks by taking advantage of the high convection heat transfer efficiency associated with the nucleate boiling. The two-phase cooled systems have shown capability to remove heat fluxes between 0.79~27kW/cm^2 [9]. There are two ways to drive the refrigerant circulation, liquid-pump driven and vapor-compressor driven.

Table 3 summarizes the typical datacenter heat sources and streams in different cooling systems. The waste heat quality of air-cooled systems is low, and thus heat pumps are usually used for upgrading waste heat. The waste heat quality in two-phase cooled systems is the highest.

Table 3 Typical datacenter heat sources and streams [8].

	Parameter	Value
Air-cooling	Cold aisle (CRAC supply) temp.	10–32 °C
	Hot aisle (CRAC return) temp.	50–60 °C
	Temp. rise over servers	10–20 °C
	Airflow per rack	200–2500 CFM
	Chiller water supply to CRAC	7–10 °C
	Chilled water return from CRAC	35 °C
Water-cooling	Water supply to server	20–60 °C (std) 70–75 °C (max)
	Water exit from server	2–5 °C temp. rise over servers
	Water flow rate per rack	5-10 GPM
	∆T from water to lid	5–18 °C
	Buffer heat exchanger flow rate	5-10 GPM
	Buffer heat exchanger supply temp.	3–5 °C above ambient
Two-phase cooling with liquid pump	Coolant supply to evaporator	60 °C saturated liquid (std.) 70–75 °C (max)
	Coolant exit from evaporator	62 °C at 30% quality (std.) 75–80 °C (max)
	Condenser cooling fluid inlet	30 °C
	Condenser cooling fluid outlet	45-90 °C
Two-phase cooling with vapor compressor	Coolant supply to evaporator	60 °C saturated liquid (std.) 70–75 °C (max)
	Coolant exit from evaporator	60 °C saturated liquid (std.) 70–75 °C (max)
	Coolant temperature at the exit of vapor compressor	~90 °C
	Condenser cooling fluid inlet	30 °C
	Condenser cooling fluid outlet	~90 °C

4. DATACENTERS AS CONSUMERS - INTEGRATION WITH RENEWABLE ENERGY GENERATIONS

This section introduces the different ways of integrating renewable energy with datacenters.

4.1 Datacenter with generation of renewable energy

In this way of renewable integration, the datacenters generate their own energy and have a direct control or influence of the energy resources. The renewable energy can be generated either 'on-site' or 'off-site'. There are three kinds of renewable generations [10].

- On-site generation from on-site renewables: Datacenters install renewable energy generation systems within their own facilities. The generation of usable form of energy takes place within the infrastructure footprint or site. The renewable energy sources are directly available on the site, such as the solar energy and wind energy.
- On-site generation with off-site renewables: Datacenters install their own renewable energy generation systems, but they have to rely on renewable energy sources outside the buildings. The generation of usable form of energy takes place on the project site. Example of this type of integration is the transportation of biomass or biogas from outside the datacenter to produce the needed electricity.
- Off-site generation: Due to factors such as insufficient renewable energy potentials and limited spaces available, in some places generating on-site renewables is not suitable. The datacenter owners invest in a renewable energy plant in places with sufficient renewable sources or in a community system. The power grid or heat/cooling networks are used as the carrier of produced energy.

By generating their own usable forms of energy, the datacenters can significantly reduce the dependence on the grid. The on-site generation can significantly lower the energy losses as the generated power undergoes less conversions and is not transmitted over long distances. However, the generation capacity may be limited for the on-site generation. The off-site generation is much more flexible in the site planning of renewable energy systems, and thus the capacity can be larger. However, the losses due to energy transmission are inevitable, and the penetration of renewable energy may reduce the reliability and efficiency of the existing energy networks.

4.2 Datacenters with renewable energy provided by a third party

The datacenter operators can also purchase renewable energy from other entities to reduce the carbon footprint. There are two important mechanisms for this type of renewable energy integration.

- Energy certificate systems: Electricity certificates provide an efficient and reliable tracking mechanism for the energy origin of the electricity system. An Energy certificate includes the information of the generation attributes (e.g. renewable fuel type, capacity and age of the plant) of the related electricity production. Examples of the certificate system include the Guarantees of Origin in the EU and Renewable Energy Credits in the US [1].
- Power purchase agreements (PPAs): A PPA is a contract between a supplier and a consumer which specifies how much electricity the supplier has promised to provide to the power gird, and how much the consumer will take off [11]. A PPA also specifies an electricity price.

5. DATACENTERS AS PRODUCERS - WASTE HEAT REUTILIZATION

5.1 Locations for waste heat recovery

In the existing studies, waste heat from datacenters has been applied in multiple ways, including district heating (most common), domestic water heating, Organic Ranking cycles for power generation, absorption chillers, desalination, biomass processing, piezoelectrics and thermoelectrics. The quality and quantity of captured waste heat in datacenter are strongly affected by the type of cooling systems and the locations of waste heat recovery [8].

- Air-cooled systems: The optimal location to capture waste heat in air-cooled datacenters is the rack exhaust prior to room air mixing. Alternatively, waste heat can be collected at the air return to CRAC or at the chilled water returns.
- Water-cooled systems: In water-cooled systems, the temperature of water exit from servers/racks are the highest. The optimal location to recover waste heat is in the water exit from servers/racks.
- Two-phase cooled systems: Due to the system complexity, the optimal location for collecting waste heat in two phase cooled system is at the condenser.

5.2 Waste heat reuse for district heating networks

Waste heat reuse in district heating network is an effective and efficient way to connect the datacenters with the urban energy systems. This section reviews the different prototypes for integrating datacenter waste heat with district heating systems.

5.2.1 Connection at the datacenter side

Fig. 1 presents the connection at the datacenter side for reuse of CRAH cooling system waste heat. There are two locations that can capture the waste heat: return hot aisle and the chiller condenser. For heat recovery from the return hot aisle of CRAH systems, a water-to-air heat exchanger is installed in the return hot stream air from the white space, as depicted by Fig. 1(1). For heat recovery from the chiller condenser of CRAH systems, a water-to-refrigerant heat exchange is installed in parallel with the condenser (or dry cooler) of the chiller, as shown in Fig. 1(2). For in-row cooling and rear-door cooling systems, since the operational air temperature is very low, it is inefficient to recover heat at the air side. A feasible solution is to capture heat from the chiller condenser. A water-to-refrigerant heat exchanger is installed in parallel with the condenser (or dry cooler) of the chiller, similar to CRAH (see Fig. 1(2)). The heat recovery in CRAC cooling systems is not considered, as it is normally equipped in small datacenters with less waste heat.

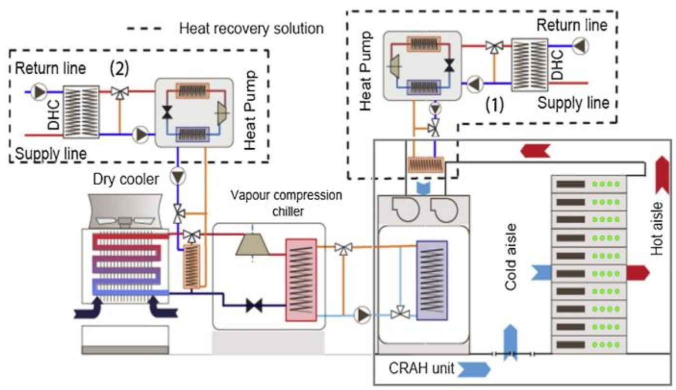

Fig. 1 Schematics of waste heat reuse for air-cooled system.

5.2.2 Connection at the district heating network side

The simplest connection of datacenter with district heating load is to add an in-between heat pump, as shown in Fig. 2. The low-grade recovered waste heat in the datacenter is fed into the heat pump. After being upgraded to the required temperature, the heat is delivered to the district heating end-users directly. This connection is easy to implement, and the initial investments are relatively low. But, since there is no thermal storage, the excessive waste heat will be released to the atmosphere. The COP of the heat pump can reach 4.3.

Fig. 2 Schematics of the energy sharging system.

To increase the utilization of waste heat, a ground source heat pump integrated with a borefield system can be added, as displayed in Fig. 3. The energy sharing heat pump is used only when the simultaneous cooling and heating loads are equal. When there is mismatch between the heat and cooling loads, the ground source heat pump will operate in either cooling or heating mode. The borefield is used as a large thermal energy storage to alleviate the mismatch. As the accumulated waste heat from the datacenter may not be equal to the accumulated heating demand of end-users, a cooling tower is installed to offset the heat injected to the borefield. The COP of this connection can reach as high as 8.2 for the cooling mode and 3.5 for the heating mode.

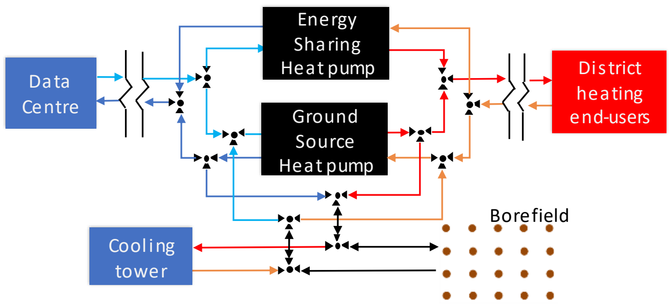

Fig. 3 Schematics of the one-borefield system [12].

6. METRICS AND PERFORMANCE INDICATORS

This section reviews the metrics related to the datacenter integration in the urban energy systems from aspects of energy, economy, and environment.

6.1 Energy-related metrics

Power usage effectiveness (PUE) compares the total energy consumption of a datacenter with the energy consumption of the IT facilities (i.e. for computing, storage and network equipment) [4]. It is defined as the ratio of the datacenter power usage ($Power_{datacenter}$ kW·h) to the IT facilities power usage ($Power_{IT}$ kW·h), see Eqn. (1).

$$\mathrm{PUE} = \frac{Power_{datacenter}}{Power_{IT}} \tag{1}$$

Datacenter energy productivity is a metric that quantifies the useful work that a datacenter produces based on the amount of energy it consumes.

$$\mathrm{DCeP} = \frac{Useful\ work\ produced\ in\ datacentre}{Total\ datacentre\ energy\ consumed} \tag{2}$$

To take into account the waste heat reuse, the energy reuse effectiveness (ERE) has been developed by the Green Grid [13]. The ERE is calculated by Eqn. (3), where Power_{reuse} (kWh) denotes the power supplied to a secondary waste heat reuse application [13].

$$\mathrm{ERE} = \frac{Power_{datacenter} - Power_{reuse}}{Power_{IT}} \tag{3}$$

Energy reuse factor (ERF) quantifies the percentage of reused waste heat energy to the datacenter total energy usage, as described by Eqn. (4) [4].

$$\mathrm{ERF} = \frac{Power_{reuse}}{Power_{datacenter}} \tag{4}$$

He et al. proposed a composite coefficient of performance (CCOP) for district heating system driven by datacenter waste heat, as shown by Eqn. (5) [14].

$$\mathrm{CCOP} = \frac{chille\ cooling\ capacity + heat\ pump\ heating\ load}{district\ heating\ system\ power\ consumption} \tag{5}$$

Green energy coefficient (GEC) quantifies the proportion of a datacenter's energy that comes from renewable energy sources.

$$\mathrm{GEC} = \frac{green\ energy\ used\ by\ datacenter}{total\ datacenter\ source\ energy} \tag{6}$$

6.2 Economy-related metrics

Total cost of ownership (TCO) calculates the real costs of the building, owning, and operating the facility, as expressed by Eqn. (7) [4]. $\mathrm{Cost}_{capital}$ is the initial investment for the heat reuse solution, $\mathrm{Cost}_{operate}$ is the operating cost, $\mathrm{Cost}_{replace}$ is the cost for component replacement, and $\mathrm{Cost}_{residual}$ is the residual value after its life cycles [4].

$$\mathrm{TCO} = \mathrm{Cost}_{capital} + \mathrm{Cost}_{operate} + \mathrm{Cost}_{replace} - \mathrm{Cost}_{residual} \tag{7}$$

Net present value (NPV) quantifies the total net profit during the project life, i.e., the difference between the operational profit and the capital investments. Eqn. (8) illustrates the calculation of NPV, where CF_i represents the net profit in the i_{th} year, *d* is the discount rate, and *N* indicates the lifespan of the system.

$$\mathrm{NPV} = \sum_{i=1}^{N} \frac{CF_i}{(1+d)^i} - \mathrm{Cost}_{capital} \tag{8}$$

Discounted payback period (DPBP) describes the period of time that is required to refund the capital investments. It can be calculated based on the NPV, as shown by the equation below.

$$\mathrm{NPV}|_{\mathrm{N=DPBP}} = 0 \longrightarrow \mathrm{DPBP} \tag{9}$$

The internal rate of return (IRR) is a discount rate that makes the NPV of all cash flows from a particular project equal to zero, as shown by Eqn. (10) [12].

$$\mathrm{NPV}|_{\mathrm{d=IRR}} = 0 \longrightarrow \mathrm{IRR} \tag{10}$$

The economic value of heat (Value_{heat}) quantifies benefits of heat recovered from datacenters [13]. It is calculated as the cost of 1 kW·h heat produced by fossil fuel combustion ($\mathrm{Cost}_{heat,fossil}$) divided by the cost of 1 kW·h heat from datacenter ($\mathrm{Cost}_{heat,datacenter}$).

$$\mathrm{Value_{heat}} = \frac{\mathrm{Cost_{heat,fossil}}}{\mathrm{Cost_{heat,datacenter}}} \tag{11}$$

6.3 Environment-related metrics

Carbon usage effectiveness (CUE) assesses total greenhouse gas emissions of a datacenter, relative to its IT energy usage [15]. For datacenters with electricity as the only energy source, CUE is calculated by Eqn. (12).

$$\mathrm{CUE} = \frac{CO_2\ emitted\ (kgCO_2e)}{Unit\ of\ energy\ (kWh)} \times PUE \quad (12)$$

7. CHALLENGES AND FUTURE WORK

The existing metrics evaluate the datacenter performances either from the perspective of energy consumers (e.g., PUE) or from the perspective of energy producers (e.g., ERE, ERF). Metrics that evaluate the datacenters as prosumers to consider the overall performances are still lacking. Future work is needed to develop such global metrics for more appropriate quantification of the datacenter performances.

Different ways of connecting datacenter with the upstream renewable energy systems or the downstream waste heat utilization facilities can lead to very different overall system energy efficiencies. Future work is needed to develop proper ways of datacenter connections (e.g. the ancillary components, connection topology) in the urban energy systems to increase the overall datacenter energy efficiencies.

Also, development of proper supervisory controls is needed to optimize the operation of datacenters globally. The existing controls optimizes the datacenter operation to match its demand with the renewable energy generations, while neglecting the impacts on the production of waste heat and its impacts on the waste heat utilization systems. Global controls that can simultaneously manage the upstream renewable production, datacenter operation and waste heat generation, and downstream waste heat utilization are needed.

8. SUMMARY

This study has conducted a comprehensive review of the datacenters as energy prosumers in the urban energy systems. The datacenter technologies in aspects of cooling systems, integration with renewable energy, waste heat recovery and reuse for district heating network, and metrics, have been investigated and discussed. By presenting a full picture of the datacenter in the urban energy systems, this study provides new opportunities for improving datacenter overall energy efficiency and reducing carbon emissions.

Future work is needed to develop new metrics that can evaluate the datacenter overall performances, to develop new system connection topology that can maximize the energy efficiency, and to develop advanced control methods that can coordinate the operation of the whole systems.

ACKNOWLEDGEMENT

The authors would like to acknowledge the financial support from the European Regional Development Fund through the project Energiinnovation.

REFERENCE

[1] Oró E, Depoorter V, Garcia A, Salom J. Energy efficiency and renewable energy integration in data centres. Strategies and modelling review. Renew Sust Energ Rev. 2015;42:429- 45.

[2] Sheme E, Holmbacka S, Lafond S, Lučanin D, Frashëri N. Feasibility of using renewable energy to supply data centers in 60° north latitude. Sustainable Computing: Informatics and Systems. 2018;17:96-106.

[3] Aksanli B, Venkatesh J, Zhang L, Rosing T. Utilizing green energy prediction to schedule mixed batch and service jobs in data centers. Proceedings of the 4th Workshop on Power-Aware Computing and Systems (p. 5). . ACM.2011.

[4] Oró E, Taddeo P, Salom J. Waste heat recovery from urban air cooled data centres to increase energy efficiency of district heating networks. Sustain Cities Soc. 2019;45:522-42.

[5] ASHRAE. Thermal Guidelines for Data Processing Environments: ASHRAE; 2015.

[6] Davies G, Maidment G, Tozer R. Using data centres for combined heating and cooling: An investigation for London. Appl Therm Eng. 2016;94:296-304.

[7] Evans T. The different technologies for cooling data centres. . Schneider electric white paper 59.

[8] Ebrahimi K, Jones GF, Fleischer AS. A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities. Renewable and Sustainable Energy Reviews. 2014;31:622-38.

[9] Bowers MB, Mudawar I. High flux boiling in low flow rate, low pressure drop mini-channel and micro-channel heat sinks. Int J Heat Mass Tran. 1994;37:321-32.

[10] Marszal AJ, Heiselberg P, Bourrelle JS, Musall E, Voss K, Sartori I, et al. Zero Energy Building–A review of definitions and calculation methodologies. Energ Buildings. 2011;43:971-9.

[11] Bruck M, Sandborn P, Goudarzi N. A Levelized Cost of Energy (LCOE) model for wind farms that include Power Purchase Agreements (PPAs). Renew Energ. 2018;122:131-9.

[12] Murphy AR, Fung AS. Techno-economic study of an energy sharing network comprised of a data centre and multi-unit residential buildings for cold climate. Energ Buildings. 2019;186:261-75.

[13] Zimmermann S, Meijer I, Tiwari MK, Paredes S, Michel B, Poulikakos D. Aquasar: A hot water cooled data center with direct energy reuse. Energy. 2012;43:237-45.

[14] He Z, Ding T, Liu Y, Li Z. Analysis of a district heating system using waste heat in a distributed cooling data center. Appl Therm Eng. 2018;141:1131-40.

[15] Grid TG. Carbon usage effectiveness (CUE): a green grid data center sustainability metric. 2010. p. White Paper #32.

