Journal of Parallel and Distributed Computing 165 (2022) 156–166

Contents lists available at ScienceDirect

Journal of Parallel and Distributed Computing

www.elsevier.com/locate/jpdc

An energy and carbon-aware algorithm for renewable energy usage maximization in distributed cloud data centers

Daming Zhao, Jiantao Zhou *

College of Computer Science, Inner Mongolia University, Hohhot 010021, China

A R T I C L E   I N F O

Article history:
Received 25 November 2020
Received in revised form 1 February 2022
Accepted 2 April 2022
Available online 12 April 2022

Keywords:
Carbon emission
Energy consumption
Renewable energy
Virtual machine placement

A B S T R A C T

The vigorous development and the increasing popularity of cloud computing highlight the necessity of reducing data center energy consumption and the environmental impact of carbon dioxide emissions. For geographically distributed data centers, cloud servers are connected to the conventional power grid and in addition they are supported by an attached renewable energy source. Since the carbon footprint rate of energy consumption has dynamic differences in space, reducing energy consumption does not mean decrease carbon emission, which indicates that energy consumption and carbon footprint need to be synergistically optimized. In this paper, an energy and carbon-aware algorithm for virtual machine placement is proposed. The goal is to obtain a virtual machine allocation scheme that aims to achieve the trade-off between energy consumption and carbon emissions by improving renewable energy utilization. The experimental results show that the proposed approach is more energy-efficient and greener, which can also maximize the renewable energy utilization with 73.11% while ensuring the SLA violation with 0.2% in comparison to the baseline algorithms.

1. Introduction

Cloud computing can provide users with "on-demand" computing, storage, and communication resources as a new Internet application model, which is also a popular research topic [30]. In recent years, with the continuous promotion and maturity of cloud computing technology, more and more cloud consumers choose to host applications on cloud platforms. The energy consumption of these geographically distributed data centers (DCs) deployed in different regions of the world is increasing accordingly. Large companies (e.g., Google, Amazon and Apple) have built their own data centers that can accommodate thousands of servers and consume tens of megabytes of power [26]. Japan's METI forecasts that global IT energy consumption will increase five times by 2025 and 12 times by 2050 [11]. Besides the high energy consumption of data centers, the carbon emission associated with the energy is a severe problem as well. According to [35], the energy consumption of the US data center is predicted to grow up to 140 billion kWh in 2020, which will generate 150 million tons of carbon emission.

Since the carbon emission is derived from energy consumption, an effective approach is proposing energy-efficient solutions at the server level by using the virtual machine placement (VMP) approach, which allocates the workload on minimum possible physical machines (PMs) to increase the resource utilization and virtual machine (VM) density on the servers [18]. However, the need to make data centers efficient not only in regards to performance factors but also in both energy consumption and carbon emission reduction have motivated a flurry of research recently [33]. Heuristic and meta-heuristic, as two common VMP methods, enable VMs to be reallocated from high carbon emission data centers as well as minimize the number of servers by switching idle hosts to low-power mode to save energy consumption. Although VMP has been proved to be effective in minimizing energy consumption [4], traditional energy supply methods have limited carbon emissions reduction. Moreover, such kind of technique may lead to service level agreement (SLA) violations if desired quantity of resources is not available.

In order to overcome the problem of high carbon emissions caused by high energy consumption, IT giants have made efforts on powering data centers with renewable energy. For example, Apple has built a 20MW solar array for its data center in North Carolina [29]. Facebook is trying to power its data centers using self-generated solar energy or wind energy [10]. To this end, onsite renewable energy source (RES) is considered, which is popular in modern data centers and is able to reduce the use of brown energy that comes from conventional fossil fuels, such as oil or coal. However, RES is often intermittent, unstable, and dynamic. Wind energy is influenced by wind speed and solar energy is af-

* Corresponding author.
E-mail address: cszhoujiantao@qq.com (J. Zhou).

https://doi.org/10.1016/j.jpdc.2022.04.001

fected by temperature and radiation. Therefore, it is challenging for the cloud provider to minimize carbon emissions by maximizing on-site renewable energy. An effective way to achieve this goal is to determine the destination site based on different criteria upon the reception of VM requests and variations of renewable energy supply, which is also viewed as a bin packing problem. But VM placement in distributed data centers is a complex decision-making process due to the spatial-variability of carbon footprint rate (CFR), the time-variability workload, and the intermittent nature of renewable energy. Besides, carbon emission and carbon footprint are assumed to have the same meaning in this paper and are interchangeable.

In this paper, a novel algorithm named EFP is proposed to achieve the trade-off between Energy consumption and carbon Footprint with Predictive renewable energy while taking into account both temporal-variability workload and spat- ial-variability CFRs for VM placement. The followings are the main contributions:

- Design a combined renewable energy generation prediction method based on Ensemble Empirical Mode Decomposition (EEMD) and Temporal Convolutional Network (TCN) to improve the prediction accuracy of RES based on Root Mean Squared Error (RMSE), which makes the resulting energy consumption matches the RES generation to decrease the consumption of brown energy.
- Propose a novel algorithm named EFP which solves the VMP problem through two successive heuristic algorithms. It is fulfilled to achieve minimal overall energy consumption by a modified power aware best first decreasing method in advance. And then, a carbon footprint-aware algorithm is presented to achieve a desirable trade-off between energy consumption and carbon footprint with the maximal exploitation of RES, which also takes the predictions of available renewable energy and VMs' workload into account.
- Compare the performance of EFP to five baseline algorithms. The superiority of EFP is demonstrated by experimental results, which show that it can effectively reduce brown energy consumption and carbon footprint with lower SLA violations while maximizing RES utilization.

The rest of this paper is organized as follows. Section 2 discusses the related works and highlights their limitations. The renewable energy generation, energy consumption, and the prediction model are built in Section 3. Section 4 describes the proposed VM placement algorithm. Section 5 shows the experimental results and compares them with other methods. Section 6 concludes the paper.

2. Related work

Over the last few years, VM placement has received a lot of attention on reducing data center energy consumption and carbon footprint due to environmental concerns. The prior works concerning the energy consumption management are known as energy-aware. However, both these approaches use coal-based brown energy as the primary source of energy. In this regard, this section will also discuss some supplementary approaches introducing renewable energy sources, which regard the trade-off between brown energy and carbon footprint.

As virtualization is a core technology of cloud computing, VMP has become an important approach for improving energy efficiency and resource utilization in cloud infrastructures. Zhao et al. [43] also investigated the effectiveness of combining genetic and tabu search algorithms to achieve energy efficiency while maximizing load balance among various resources. Chen et al. [9] proposed a two-tier VM placement framework, which includes an LP-based variation-unaware VM placement algorithm with minimized energy consumption and a feasibility driven stochastic VM placement with satisfactory feasibility. Li et al. [27] developed a discrete differential evolution algorithm that aimed to minimize the energy consumption of running PMs while ensuring the lowest possible overloading risk of PM resources via a dynamic optimization model for VM placement. Regaieg et al. [32] modelized the VMP with a multi-objective integer linear program and solve it through two different methods, which aims to maximize the number of hosted VMs and minimize the amount of energy consumption in two DC architectures. Unlike these approaches for VM placement optimization, which always just take into account the reduction of the number of active hosts. This work concentrates on: (1) considering the impact of the carbon footprint on the environment; (2) introducing renewable energy to supply data centers so that the brown energy can be reduced.

Shifting from fossil fuels to renewable energy sources is essential to keep carbon and greenhouse emissions in data centers to a minimum. But VM requests vary according to the temporal and spatial, so it is difficult to match with renewable energy. In view of this, Grange et al. [13] adopted a new scheduling heuristic algorithm agnostic of the electrical infrastructure, which takes advantage of the intermittent nature of available RES to reduce the need for brown energy. Camus et al. [8] proposed an energy-efficient cloud management approach that is able to exchange renewable energy between data centers and thus, optimize the cloud's self-consumption and further consume the locally produced renewable energy. Khosravi et al. [25] investigated parameters that have the biggest effect on energy and carbon footprint cost and then proposed an efficient VM placement approach with access to renewable energy sources. Xu et al. [37] introduced a workload shifting algorithm which is derived from the best fit decreasing algorithm. It is essentially a two-stage algorithm that checks the availability of green energy and then considers the average response time while allocating requests to reduce brown energy and ensure the quality of service (QoS). Although the aforementioned methods incorporate RES to make the DC green, there are still challenges existed such as: (1) ignoring the variation of RES in the future, which may not only cause the wastage of renewable energy but also lead to the increment of brown energy consumption; (2) neglecting the workload changes within a period while making decisions on energy and carbon optimization, which may result in the VMs' requests exceed PM's capacity in some cases.

To further decrease carbon emission, decision-making based on a statistical analysis of historical workload and RES data is an effective method to improve the energy efficiency of data centers. Specifically, Xu et al. [38] developed a prototype system to evaluate the performance of the proposed self-adaptive approach, which includes a multiple layers perspective model for interactive and batch workloads involving renewable energy. Goiri et al. [12] proposed the Greenslot model which is a greedy workload scheduling method. By predicting the output of solar energy and ensuring each work's execution can meet their deadlines, it aims to schedule as many jobs as possible in the queue to minimize the data center overhead. Kaur et al. [22] designed a two-phrased multi-objective optimization algorithm to encourage the sustainability of geo-distributed cloud DCs, which uses Boruta to classify workload and then formulate VM placement that is based on an enhanced heuristic approach to improve RES utilization while ensuring SLA. Khosravi et al. [25] designed an online VM migration solution with limited future knowledge regarding solar or wind power availability, which decides the time to migrate a VM to another data center with excess renewable energy to minimize brown energy consumption. Different from these efforts, this paper covers the perspectives as: (1) none of them put the efforts to enhance the accuracy of RES prediction; (2) none of them consider the effect of

Table 1
Comparison of related work.

Reference	Energy-Aware	Carbon-Aware	Multi-CFR	Metrics					
				RES prediction	Energy	Brown	Carbon	RES utilization	SLA
Zhao et al. [43]	✓				✓				
Chen et al. [9]	✓				✓				
Li et al. [27]	✓				✓				✓
Regaieg et al. [32]	✓				✓				✓
Grange et al. [13]	✓				✓	✓		✓	✓
Camus et al. [8]	✓				✓	✓		✓	
Khosravi et al. [24]	✓	✓	✓		✓	✓	✓		✓
Xu et al. [37]	✓	✓	✓		✓	✓	✓		✓
Xu et al. [38]	✓	✓		✓	✓	✓			✓
Goiri et al. [12]	✓	✓		✓	✓	✓		✓	✓
Kaur et al. [22]	✓	✓		✓	✓		✓	✓	✓
Khosravi et al. [25]	✓	✓		✓	✓	✓	✓	✓	✓
This paper	✓	✓	✓	✓	✓	✓	✓	✓	✓

geographically varied carbon footprint rates that will cause huge carbon emission even if the brown energy consumption of the DC is small.

Table 1 summarizes the relevant studies. The approach that is proposed in this paper is listed at the bottom of the table. To be best of the knowledge, it is the first time that EEMD and TCN are combined to improve the accuracy of renewable energy prediction so that the wastage caused by resource mismatch can be reduced in a hybrid-powered data center. In addition to this, the proposed two-stage VM placement approach considers variations in CFRs of geographically distributed DCs and workloads of different PMs to minimize carbon emissions.

3. System model

In this section, a typical IaaS cloud system is being considered, wherein renewable energy is adopted to supply the data center so that the carbon footprint can be minimized through VM placement. The overall system architecture of the proposed algorithm is shown in Fig. 1. For the proposed solution, the energy consumption module firstly calculates and predicts the energy consumption of IT devices based on workload data to obtain the energy consumption of DCs by the value of Power Usage Effectiveness (PUE). Then, the generation of renewable energy is predicted according to the traces of wind speed collected from wind farms. EEMD and TCN are combined here to improve accuracy. Finally, the energy consumption module, together with the output of the renewable generation module, will be conveyed to the energy adaptation module that aims to maximize the utilization of renewable energy so that carbon emission can be minimized. Details of the energy adaptation module will be introduced in the next section. Table 2 defines the symbols that are used throughout this paper. Moreover, energy consumption and power consumption will be used interchangeably in this paper.

3.1. Energy consumption model

The basic mathematical model is established to depict the typical geographically distributed data center system in the real world, such as Amazon's EC2. n DCs, shown as $D = \{D_1, D_2, \ldots, D_n\}$, are powered by brown energy and renewable energy to provide cloud services. Each DC contains m physical servers with different configurations, shown as $S = \{S_1, S_2, \ldots, S_m\}$. In each discrete time $t\epsilon\{1, 2, \ldots T\}$, cloud user requests arrive at a DC, which are executed by VMs that are represented as $VM = \{VM_1, VM_2, \ldots VM_h\}$.

Moreover, each geo-distributed DC has a PUE value that is expressed as $PUE = \{PUE_1, PUE_2, \ldots, PUE_n\}$, which is a metric of

Table 2
Description of symbols.

Notation	Description
D	Set of data center sites
S	Set of physical servers in a data center
VM	Set of VM requests
PUE	Set of PUE values
$P_k(t)$	Total power consumption of DC k at time t
$P_k^{IT}(t)$	Power consumed by IT devices in DC k at time t
$\beta_{jk}(t)$	CPU utilization of the jth server in DC k at time
$CFR_k(t)$	Carbon intensity of kth DC at time t
$CF_k(t)$	Carbon footprint of kth DC at time t
$v_k(t)$	Wind speed of the kth DC at time t
v_{in}	Cut-in speed of wind turbine
v_{out}	Cut-out speed of wind turbine
v_r	Rated wind speed of wind turbine
P_r	Rated power of wind turbine
$PT_{lk}(t)$	Wind power output of the lth turbine in kth DC at time t
M_k	The number of wind turbine in DC k
$R_k(t)$	Wind power generation of DC k at time t
$IMF_k^a(t)$	ath IMF of DC k at time t
$RS_k^b(t)$	Residual component of DC k at time t
$G_k(t)$	White noise sequence of DC k at time t
η	Window size of prediction
$v_k^p(t+\eta)$	Predicted wind speed of DC k at time $t+\eta$
$R_k^p(t+\eta)$	Predicted wind power generation of DC k at time $t+\eta$
$\mu_k(t+\eta)$	Predicted server utilization in DC k at time $t+\eta$
$P_k^p(t+\eta)$	Predicted total power consumption of DC k at time $t+\eta$

overhead power. In general, the value of PUE changes with time and the environment outside DC. The corresponding research has also been conducted in [24], but it is not the focus of this paper. In this research, PUE is considered as a constant like most researches do [25,23,22], which is equal to the ratio of the total power consumption of various equipment in a data center to the power consumed by IT devices:

$$PUE_k(t) = \frac{P_k(t)}{P_k^{IT}(t)} \tag{1}$$

where $P_k^{IT}(t)$ is the sum of the power consumption of servers in kth DC and it can be formalized as:

$$P_k^{IT}(t) = \sum_{j=1}^{m} P_{jk}^{IT}\left(\beta_{jk}(t)\right) \tag{2}$$

where $\beta_{jk}(t)$ is the CPU utilization of the jth server in the kth DC. As modeling a precise analytical model for P_{jk}^{IT} is quite a difficult job, the results of SPECpower benchmark from a real power con-

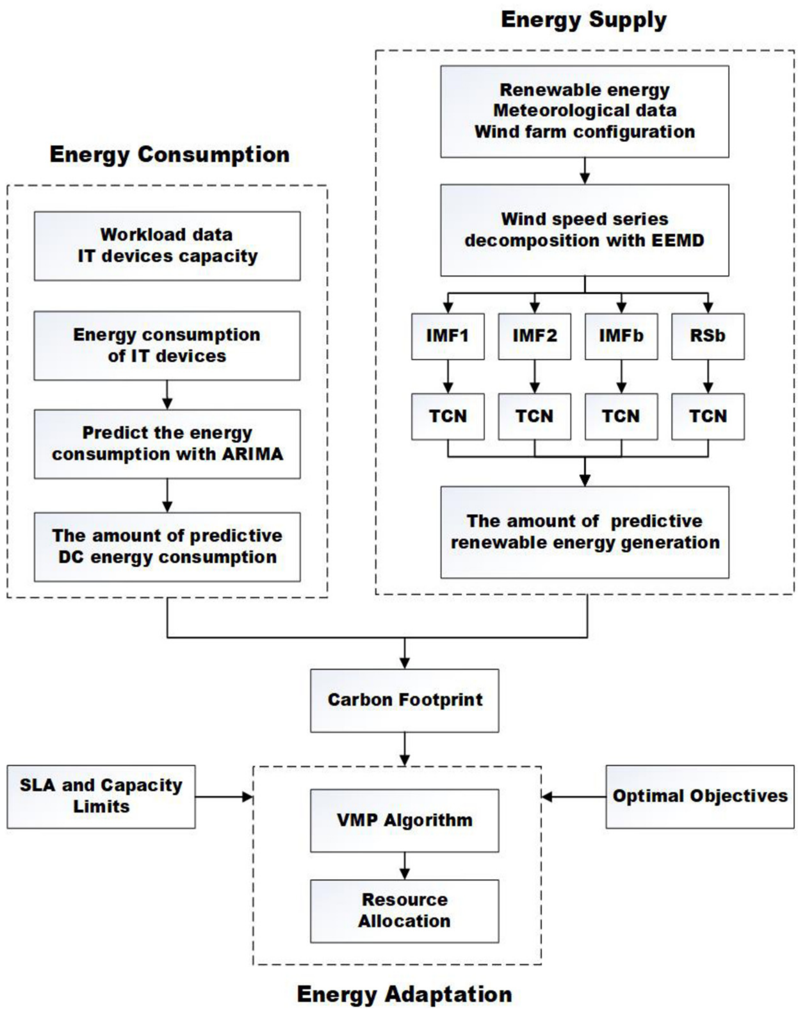

Fig. 1. System architecture of the proposed algorithm.

Table 3
Power consumption against CPU utilization of servers under consideration at various loads in Watts.

Servers	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
HP ProLiant G4	86	89.4	92.6	96	99.5	102	106	108	112	114	117
HP ProLiant G5	93.7	97	101	105	110	116	121	125	129	133	135

sumption data [44] is adopted to evaluate P_{jk}^{IT} which is decided by the value of $\beta_{jk}(t)$, as shown in Table 3.

3.2. Renewable energy generation model

As the cost of deploying renewable energy power generators continues to decrease, a growing number of DCs are gradually becoming partly or wholly supplied by renewable energy. It is assumed that each DC uses on-site wind energy to power its infrastructures and servers, and the wind power output depends on environmental conditions, such as wind speed, wind direction, atmospheric pressure, air density, etc. In the existing research [41], it has been proved that wind speed is the most crucial impact factor of wind power. Furthermore, wind turbines generation can be modeled as a function of wind speed, as mentioned in [10,42]. Let $v_k(t)$ be the wind speed of the kth DC at time t, then the wind power output of the lth turbine in kth DC at time t can be approximated as follows:

$$PT_{lk}(t) = \begin{cases} 0 & v_k(t) < v_{in}, v_k(t) > v_{out} \\ P_r \times \frac{v_k(t) - v_{in}}{v_r - v_{in}} & v_{in} < v_k(t) < v_r \\ P_r & v_r < v_k(t) < v_{out} \end{cases} \tag{3}$$

when the wind speed is lower than the cut-in speed v_{in} or higher than the cut-out speed v_{out}, then the power is set to 0. If the actual wind speed is between cut-in and rated, the wind power will increase as the wind speed increases. When the wind speed of the wind turbine is higher than the rated wind speed and lower than the cut-out speed, the wind power will equal to the rated power. If the kth DC contains M_k turbine without changing over time, and the wind turbine is assumed to have the same wind speed in the same wind field for the purpose of simplifying the model, then the wind power of this DC at time t can be calculated as:

$$R_k(t) = PT_{lk}(t) \times M_k \tag{4}$$

3.3. Carbon footprint model

Carbon footprint is caused by brown energy, which is emitted into the environment. In this paper, the CFR for green energy is assumed to be zero. CFR_k is the CFR of the kth DC, which is viewed as a constant and only change with the data center's location. At time t, the carbon emission of the kth data centers is formalized as follows:

$$CF_k(t) = P_k^b(t) \times CFR_k \tag{5}$$

where $P_k^b(t)$ is the actual power consumption affected by RES, which is calculated as:

$$P_k^b(t) = max\left(0, P_k(t) - R_k(t)\right) \tag{6}$$

3.4. Prediction model

The biggest challenge faced by wind power forecasting is its intermittency and uncertainty. This kind of fluctuation is the main factor causing prediction errors. As mentioned earlier, wind speed is the biggest factor affecting wind power. Moreover, the decomposition of nonlinear wind speed time series instead of modeling and prediction directly has been proved to be efficient. The whole process of wind prediction is shown in Fig. 1. The basic idea is to use the EEMD method proposed by Zhao and Huang [36,17] to improve the aliasing phenomenon. For a DC, a series of intrinsic mode functions (IMF) $IMF_k^1(t), IMF_k^2(t), \ldots, IMF_k^b(t)$ can be obtained from high to low and a residual sequence $RS_k^b(t)$ by taking the wind speed with Gaussian white noise as a whole and adaptively decomposes it. The process can be described as:

$$v_k(t) + G_k(t) = \sum_{a=1}^{b} IMF_k^a(t) + RS_k^b(t) \tag{7}$$

where $IMF_k^a(t)$ is the ath IMF obtained by EEMD decomposition at time t. $RS_k^b(t)$ is the residual component after decomposing and filtering out b IMFs, which usually represents the trend of the wind speed. $G_k(t)$ is the added white noise sequence. Then, the TCN is constructed to get the predicted values $v_k^1(t+\eta), \ldots, v_k^b(t+\eta), v_k^{b+1}(t+\eta)$, η is the window size which has an important effect on accuracy and it will be discussed in Section 5. Finally, the predicted value of those subsequences are comprehensively superimposed to complete the entire prediction process. The predicted wind speed can be calculated as:

$$v_k^p(t+\eta) = \sum_{a=1}^{b+1} v_k^a(t+\eta) \tag{8}$$

The reason for choosing TCN here is that compared with the LSTM [45] which is commonly used in most time series prediction research, there is a causal relationship between the convolution network layer, meaning that historical information or future data will not be missed. Even if the LSTM has a memory gate, it cannot remember all the historical information completely. On the contrary, TCN can improve accuracy and accelerate the prediction speed. Compared with ARIMA [40], TCN is better at dealing with multiple variables. For example, when considering the environmental factors that may affect the wind speed of each renewable energy generation center, such as temperature, humidity, and pressure on wind speed, TCN is a better choice. To the best knowledge, this paper is the first study to combine EEMD and TCN for wind speed prediction. The calculation process of predicted renewable energy output of kth DC is the same as (5) and (6), and its value is $R_k^p(t+\eta)$.

The resource utilization prediction model adopted in this paper is based on the ARIMA model, which is characterized by high prediction accuracy and is capable of one-step or multi-step prediction based on loads matching multiple patterns. Most importantly, it has been proved to be simple for forecasting resource utilization in a cloud environment [7]. The predicted CPU utilization of the jth server in the kth DC is computed as:

$$\beta_{jk}(t+\eta) = ARIMA\left(\beta_{jk}(t)\right) \tag{9}$$

Then, the predictive power consumption of kth DC at time $t+\eta$ can be calculated according to (1) and (2), which is regarded as $P_k^p(t+\eta)$.

3.5. Objective function

The optimization function of geographically distributed DC powered by renewable energy is minimizing the combination of brown energy consumption and renewable energy wastage, which eventually affects the carbon footprint as shown in Eq. (5). Therefore, the optimization problem becomes to minimize the total carbon footprint while satisfying the PM and VM capacity constraints, which then can be defined as follows:

$$CF = \sum_{k=1}^{n} CF_k(t) \tag{10}$$

For each time slot, the allocated CPU capacity of VM should be less than that of the hosted server's capacity to satisfy VM requests, which can be expressed as:

$$C_{jk}^{CPU}(t) \le S_{jk}^{CPU} \tag{11}$$

where S_{jk}^{CPU} is the CPU capacity of the jth server in the kth DC. $C_{jk}^{CPU}(t)$ is the total CPU capacity of VMs that are allocated on the jth server in the kth DC, which can be calculated as:

$$C_{jk}^{CPU}(t) = \sum_{i \epsilon VM} v_{ijk}^{CPU}(t) \times x_{ijk}(t) \tag{12}$$

where $v_{ijk}^{CPU}(t)$ is the CPU capacity of the ith VM of the jth server in the kth DC. A binary integer $x_{ijk}(t)$ with the value of 1 represents VM i is assigned to server j of the kth DC and 0 otherwise. Furthermore, each VM must be allocated to a PM, which can be then expressed as:

$$\sum_{j \epsilon S} \sum_{k \epsilon D} x_{ijk}(t) = 1 \tag{13}$$

4. VM placement approaches

In this section, VM placement is viewed as a bin packing problem, which is solved by the proposed EFP algorithm that consists of two successive heuristic algorithms. To be specific, the first algorithm aims to achieve minimal overall energy consumption by determining the destination PM for VM requests based on the existing PABFD method [5]. But the energy consumption in the near future is considered. After that, a novel algorithm named EFP is proposed to balance Energy consumption and carbon Footprint with Predictive renewable energy, which also takes VMs' workload into account. In addition, five variants of the proposed method are also designed as baseline algorithms, which include different factors that affect the carbon emission of data centers.

Algorithm 1: Predictive PABFD-Based VM Placement Algorithm.

Input: DC_List: the list of distributed DCs, $Server\_List$: servers in DCs, VM_List: VMs requests arriving at time t
Output: VM Placement

1 Sort VM_List according to their CPU capacity by descending order
2 **for** vm *in* VM_List **do**
3 $minPower \leftarrow MAX$
4 **for** $dc \in DC\_List$ **do**
5 **for** $pm \in dc$ **do**
6 **if** *pm can accommodate vm without exceeding its capacity and* $P_k(t) + P_k^p(t+\eta) < minPower$ **then**
7 $minPower \leftarrow P_k(t) + P_k^p(t+\eta)$
8 $S_{des} \leftarrow pm$
9 $DC_{des} \leftarrow dc$
10 **end**
11 **end**
12 **end**
13 Allocate vm to S_{des} belongings to DC_{des}
14 **end**

4.1. Energy and carbon footprint-aware with predictive RES (EFP)

At the initial phrase, Algorithm 1 sorts the VMs in descending order of their CPU capacity and servers according to their available CPU capacity (Line 1). Then, Lines 2-13 apply the modified PABFD algorithm to search for the destination DC and server that satisfies VM requests, which allocates VMs by decreasing power consumption with the consideration of servers' capacity limits (Line 6). Among them, Line 3 assumes that the initial power consumption is large enough. The difference from the traditional PABFD algorithm is that the consideration of predictive energy consumption (Lines 7-11). Therein, the data center will consume the least overall energy consumption for a period of time in the future. Moreover, compared to the well-known First Fit (FF) algorithm [19] and Best Fit (BF) algorithm [20], the adopted PABFD regards increased power consumption as a metric rather than just putting VMs on the PMs to accommodate them.

Algorithm Complexity Analysis: The time complexity of Algorithm 1 is dependent upon the complexities to carry out sorting VMs and PMs, and determining available PMs. It is assumed that h VMs, n data center sites, and m PMs within each data center as mentioned in the previous section. The sort operation executed in Line 1 takes $O(hlog(h))$. The modified PABFD can be done in $O(nm)$ to determine the destination for a VM. Thus, the final complexity of Algorithm 1 is $O(hlog(h) + hnm)$.

The second algorithm evaluates and adjusts the placement of VMs in the data center according to the carbon emission of each data center. The main idea can be roughly divided into three steps. The first step is to obtain the data center information (Line 1). Then, each data center calculates its power consumption and RES generation (Lines 2-12) to calculate the carbon footprint (Line 11). Specifically, DCs that consume more energy than RES generation will be added into DC_{brown} (Line 6). On the contrary, the DC with no carbon footprint indicates its renewable energy output is greater than power consumption, which will be regarded as DC_{green} (Line 9). Since CPU is the main factor that leads to high energy consumption, migrating VM with the largest CPU resource requirement is an effective way to avoid excessive brown energy consumption (Line 14). Finally, the last step considers two situations for VMs executed on DC_{brown} (Lines 15-41). On the one hand, the overall energy consumption after VM migration is less than the RES generation and the carbon footprint is decreased (Lines 25-30), which is beneficial to green data centers. Therefore, the destination PM S_{des} situated in DC_{green} with the least increased carbon footprint will be selected (Lines 34-36). On the other hand, the method proposed in Algorithm 1 is applied to fulfill the power-aware VM placement of Algorithm 2 (Lines 37-39).

Algorithm 2: Carbon Footprint-aware VM Placement Algorithm.

Input: DC_List: the list of distributed DCs, $Server\_List$: all servers in DCs, VM_List: VMs requests arriving at time t
Output: VM Placement

1 Obtain DC's workload and RES generation
2 **for** dc *in* DC_List **do**
3 $powerConsumed \leftarrow$ Calculate power consumption at time t by (1)
4 $RESGeneration \leftarrow$ Get renewable energy generation by (4)
5 **if** $powerConsumed > RESGeneration$ **then**
6 add dc into DC_{brown}
7 **end**
8 **else**
9 add dc into DC_{green}
10 **end**
11 $carbonFootprint \leftarrow$ Calculate carbon footprint of dc by (5)
12 **end**
13 $DC_{green} \leftarrow$ Sort DCs in DC_{green} by descending predictive available renewable energy generation
14 Sort VMs according to their CPU capacity by descending order
15 **for** vm *in* VM_List **do**
16 $minCarbon \leftarrow 0$
17 $flag \leftarrow False$
18 **if** $vm \in DC_{brown}$ **then**
19 **for** dc *in* DC_{green} **do**
20 $availGreen \leftarrow$ Calculate the available renewable energy generation of dc
21 **for** pm *in* dc **do**
22 **if** *pm can accommodate vm without exceeding its capacity* **then**
23 $increasedEnergy \leftarrow$ Calculate the increased energy consumption if vm is allocated to pm
24 $increasedCarbon \leftarrow$ Calculate the carbon footprint increased if vm is allocated to pm
25 **if** $increasedEnergy \le availGreen$ *and* $increasedCarbon \le minCarbon$ **then**
26 $S_{des} \leftarrow pm$
27 $DC_{des} \leftarrow dc$
28 $minCarbon \leftarrow increasedCarbon$
29 $flag \leftarrow True$
30 **end**
31 **end**
32 **end**
33 **end**
34 **if** $flag$ **then**
35 Allocate vm to S_{des} belonging to DC_{des}
36 **end**
37 **else**
38 Determine the destination PM S_{des} and destination DC DC_{des} according to Algorithm 1 to achieve the minimum energy consumption
39 **end**
40 **end**
41 **end**

Algorithm Complexity Analysis: The overall complexity of Algorithm 2 in detail is as follows: Lines 2-12 take $O(n)$ to divide data centers into two sets. The time to sort DCs in Line 13 takes $O(nlog(n))$ if all DCs are regarded as green in the worst case. Line 14 takes $O(hlog(h))$ to sort VMs. Since the time to find the destination PM in a cleaner DC with sufficient renewable energy for a VM takes $O(nm)$, the total running time for all VMs is $O(hnm)$. Therefore, the final complexity of Algorithm 2 is $O(nlog(n) + hlog(h) + hnm)$.

4.2. Energy and carbon footprint-aware with non-predictive RES (EFNP)

The EFNP algorithm is different from the EFP algorithm because it does not consider the renewable energy generation and workload changes of the destination data center in a future period when allocating VM. In Algorithm 1, the server with minimum energy consumption increased at the current time slot will be selected to accommodate the arrived VM (Lines 6-11). Then, if the energy consumption after adjusting VM placement is higher than the RES generation at the current time slot, Line 38 of Algorithm 2 in the EFP should be modified accordingly. The result of this method is expected to be inferior to the EFP algorithm because EFP improves renewable energy efficiency in the next period while increasing the utilization of RES in the current situation. Assuming that the SLA violation (SLAV) means that the resource requirement of VMs can not be satisfied by the hosted PM, which leads to more PMs suffer with more than 100% of CPU utilization. Therefore, it is obvious that EFP will be more in line with the SLA, while EFNP has to take some time to adjust VMs again to avoid SLAV.

4.3. Energy and carbon footprint-aware (EF)

EF is a variant of EFP that does not consider the availability of renewable energy while making VM re-allocation decisions as shown in Algorithm 2. Note that all the other algorithms assume that data centers have available renewable energy, VMs should be allocated to these data centers to reduce carbon footprint. Therefore, this benchmark is designed to demonstrate the necessity of using renewable energy to make DCs green. In this regard, Algorithm 2 will not divide DCs into two sets, and Lines 1-13 will be omitted accordingly. Furthermore, Line 25 only needs to consider the variation of carbon footprint after VM allocation.

4.4. Energy-aware with predictive RES (EP)

The EP algorithm makes a decision based on power consumption. For DC_{green}, EP orders DCs by increasing power consumption instead of renewable energy generation in Line 13 of Algorithm 2. In addition, Line 24 of Algorithm 2 should be removed and Line 25 will ignore the carbon footprint while determining the destination DC and PM. The experimental results of the EP are expected to be superior to EFP in the metric of energy consumption. However, due to a lack of consideration of spatial-varied CFRs of geographically distributed data centers, there is a possibility that EP will produce more carbon emissions.

4.5. Carbon footprint-aware with predictive RES (FP)

The strategy of the FP algorithm is similar to EFP. For the initial VM placement in Algorithm 1 (Lines 6-11), FP will select the destination server with minimum carbon footprint increased. Then, Line 38 of Algorithm 2 turns to find the server with the least carbon footprint increased. The expected carbon footprint of this algorithm is the best VM placement algorithm besides EFP. Because compared to EFP, FP only focuses on carbon footprint. Although it adopts the predicted RES and combines it with the improved Best Fit Decreasing (BFD) algorithm [21,6] to effectively reduce carbon footprint compared with other strategies, the utilization of renewable energy is not as good as EFP because of ignoring matching the power consumption. The experiments in Section 5 also proved this point of view.

4.6. Energy-aware with non predictive RES (ENP)

To verify the necessity of Algorithm 2 and its improvement in energy consumption and carbon footprint, Algorithm 1 is renamed as ENP.

5. Experimental evaluation

This section aims to evaluate the performance of the proposed EFP algorithm, which is realized by using PyCharm 3.3 as a simulator on a PC with Intel(R) Core i7-8750H processor with 2.2 GHz CPU and 16 GB RAM. Renewable energy estimation and utilization, energy consumption, carbon footprint, and SLA violation are considered as performance metrics. In the following, details of the data centers configuration, workload, renewable energy traces, carbon footprint rates, and PUEs are explained. Finally, the effectiveness of EFP is validated by comparing it with baseline algorithms.

Table 4
Server/VM types and capacity.

Name	CPU	Core
HP ProLiant G4	1860 MIPS	2
HP ProLiant G5	2660 MIPS	2
VM 1	2500 MIPS	1
VM 2	2000 MIPS	1
VM 3	1000 MIPS	1
VM 4	500 MIPS	1

5.1. Experimental settings

5.1.1. Data centers configuration

In order to simulate the impact of environmental change on renewable energy generation in different time zones, four data centers located in Arizona, California, Oregon, and Louisiana of the United States are taken into account. The main reason for selecting them is that they are distributed in different time zones, which indicates that wind speed will be affected by the climate and the variation of CFR is obviously among them. Specifically, each data center consists of a different number of heterogeneous physical servers, half of which are HP ProLiant ML G4 servers, and the other half are HP ProLiant ML110 G5 servers, which are also used in most researches [31,2]. Table 3 shows how these two types of servers' power consumption against CPU utilization of servers under consideration at various loads by the results of SPECpower benchmark. Since CPU consumes more energy than memory and storage, the capacity of a server is represented in the form of CPU frequency by MIPS, with the HP ProLiant G4 server at 1860 MIPS and the HP ProLiant G5 server at 2660 MIPS, which is shown in Table 4. VMs allocated to PMs are configured as single-core, which is assumed to be executed all time unless there is no resource requirement on it. Besides, the types and capacities of four VMs are also presented in Table 4.

5.1.2. Workload

The daily active users from different locations based on Google Cluster Data (GCD) [28] dataset are estimated to simulate the number of requests. GCD is used because it provides a wide range of applications and has users from various locations. In addition, most researchers adopt it when work in a simulation environment [2,1,34,16]. Since this paper only deals with the placement of the VM requests and allocation of their required resources, it does not need to know the type of application running within the instantiated VM. Furthermore, a VM is assumed to be active as long as the required CPU capacity in the request passed to the VM is 0.

5.1.3. Wind energy

The RES data traces come from the National Renewable Energy Laboratory (NREL) [29] that provides time-series data from 30000 measurement points worldwide at 10-minute intervals. In experiments, the records predict the last 5-day wind speed information from April 1st to April 30th, 2016 that is 720 hours totally. To be specific, the data of the first 20 days includes 480 hours are used as the training sample, which is composed of four input variables including environmental temperature, environmental humidity, wind speed, and time. The next five days consist of 120 hours are adopted as test samples and the last five days are used as the verification samples. Therefore, wind power generation can be then predicted. Furthermore, unlike researches in [3,14] that consider solar and wind energy simultaneously, this paper assumes

Table 5
Parameters of wind turbine.

Parameter	v_{in} (m/s)	v_{out} (m/s)	v_r (m/s)	P_r (kW)
Value	2.5	35	11	3

Table 6
Carbon Footprint Rate (Tons/MWh).

	Nevada	Arizona	California	Hawaii
CFR	0.909	0.658	0.350	0.858
Servers	300	100	200	400

Table 7
Wind speed prediction RMSE.

	Nevada	Arizona	California	Hawaii
LSTM	0.958	1.212	0.548	0.570
Hybrid	0.545	0.986	0.377	0.520

that NE-3000 wind turbines [15] are installed at each data center to fulfill the demand of cloud DCs. The parameters of the wind turbine are shown in Table 5.

5.1.4. Carbon footprint rate and PUE

The carbon footprint rate ($Tons/MWh$) of the four data centers are obtained from the US Department of Energy Electricity Emission Factors, which contains the carbon intensity data, as reported in Table 6. These values are used to calculate the carbon footprint, which is different between these data centers. California has the least carbon footprint when it consumes the same power compared with the other three sites. On the contrary, the PUE value of all these DCs can reach as low as 1.2 by using the state-of-art infrastructure. Although PUE is an uncertain value that varies between [1,2], it will not change in this research. Besides, the number of servers in each DC is also included in Table 6.

5.2. Experimental results

Experiments are conducted to evaluate the performance of the proposed algorithm according to energy consumption, renewable energy utilization, carbon emission, and SLA. All experiments are repeated 10 times.

5.2.1. Renewable energy estimation

In order to verify the accuracy of the prediction model, this paper predicts the wind speed by comparing the hybrid prediction model with the Long Short-Term Memory (LSTM) model which is often used as a prediction algorithm in relevant studies and has high accuracy [39]. In specific, the η mentioned in Section 3 is set to 10 minutes in accordance with the measurement intervals of NREL data traces. Furthermore, the value of b that occurs in Section 3.4 is uncertain, which is decided by the variation of wind speed sequence and obtained by the package of TensorFlow. It will be large if the wind speed fluctuates wildly. The prediction results of the wind speed in four locations are evaluated by the RMSE, as shown in Fig. 2 to Fig. 5. The detailed value of the predictive result is presented in Table 7. Accordingly, the renewable energy generation can be estimated by Eq. (3).

As can be seen from Fig. 2 to Fig. 5, both methods can fit the variations in wind speed, but the hybrid method performs better according to Table 7, especially in the data centers where wind speeds vary significantly, the prediction error of LSTM is more apparent. Moreover, although the prediction errors improved by the hybrid prediction models are small, it can be foreseen from Eq. (4) that the error will be amplified in wind energy prediction. Ob-

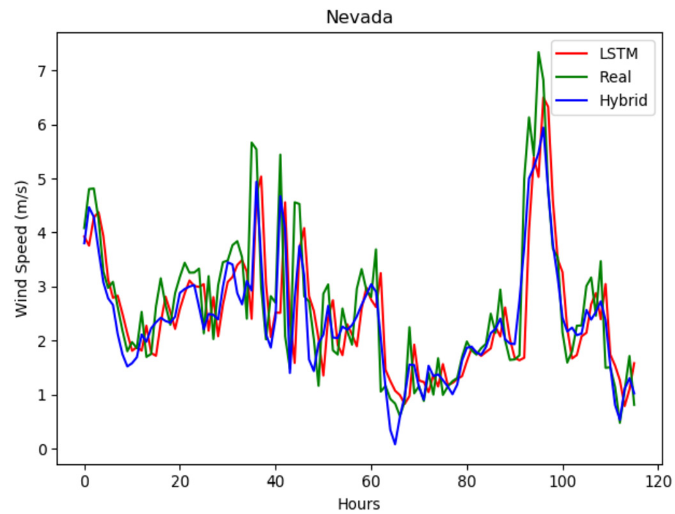

Fig. 2. Wind speed prediction of Nevada.

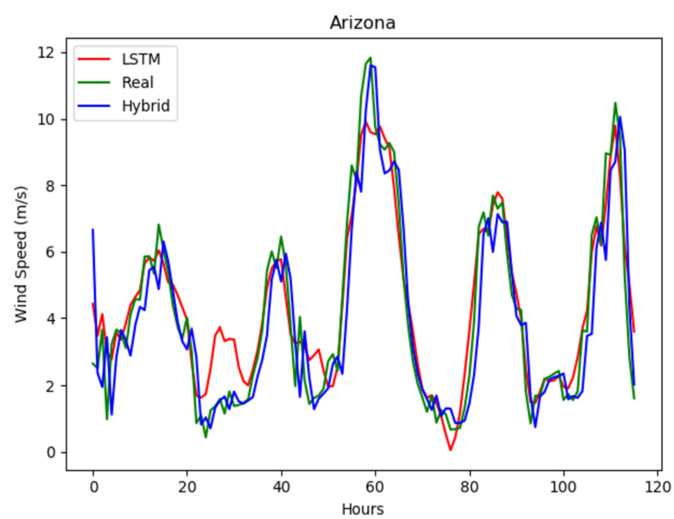

Fig. 3. Wind speed prediction of Arizona.

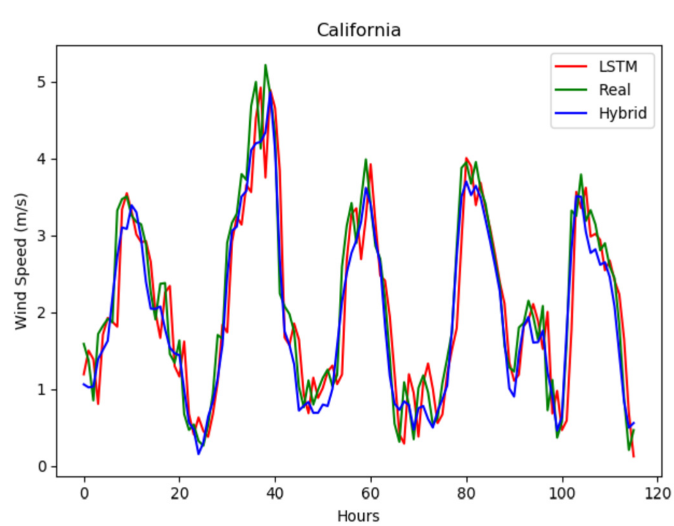

Fig. 4. Wind speed prediction of California.

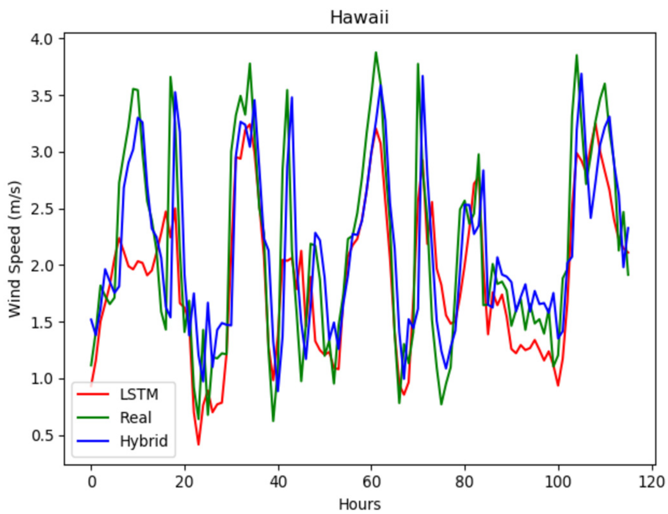

Fig. 5. Wind speed prediction of Hawaii.

Table 8
Energy consumption/Brown energy consumption of VMP algorithms.

VMP	Brown energy (W)	Total energy (W)
EFP	191506.1	322869.5
EFNP	191807.2	323167.2
EF	320108	320108
EP	192143.3	322573.3
FP	193010.1	324195.7
ENP	195964	322269

viously, the hybrid method can better represent the variation of actual wind energy.

5.2.2. Energy consumption

Fig. 6 is the comparison of the energy consumption between EFP and baseline algorithms. It can be found that EFP is not the most energy-aware. The reason is that EFP incorporates RES, which allocates VMs to DCs with the most available RES instead of DCs with the least energy consumption as EP and ENP do. On the contrary, FP is carbon-aware and thus consumes the most energy consumption. Furthermore, since EFNP does not consider the predictive RES, it will consume more energy caused by the mismatch of RES. Although EF performs best in the metric of energy consumption due to the omission of available RES, the main contribution to carbon emission is decided by brown energy. In this regard, Table 8 is introduced to present the detailed consumption of brown consumption, and EF is the worst with 320108 (W). Besides EF, ENP performs worst in brown energy, which proves the necessity of the carbon-aware method known as Algorithm 2 in this paper.

5.2.3. Renewable energy utilization

Fig. 7 is the utilization of wind energy, which represents the ratio of the wind energy used in the data center to the actual energy required by the data center. As can be seen from Fig. 7, EFP has the highest utilization rate of renewable energy, with a utilization rate of 73.11%. Therefore, it is not hard to explain why EFP consumes the least brown energy in Table 8. Although EFNP is close to EFP with 73.01% wind utilization, it is still inferior to EFP. The reason is that EFNP does not consider the predictive renewable energy, which results in that the server with a higher greenness in the near future will be not selected for the VM allocation. If more time slots are collected, the gap between them will be even larger. Similar to EFNP, ENP is also unaware of the future performance of VM allocation and RES generation and thus it performs worse

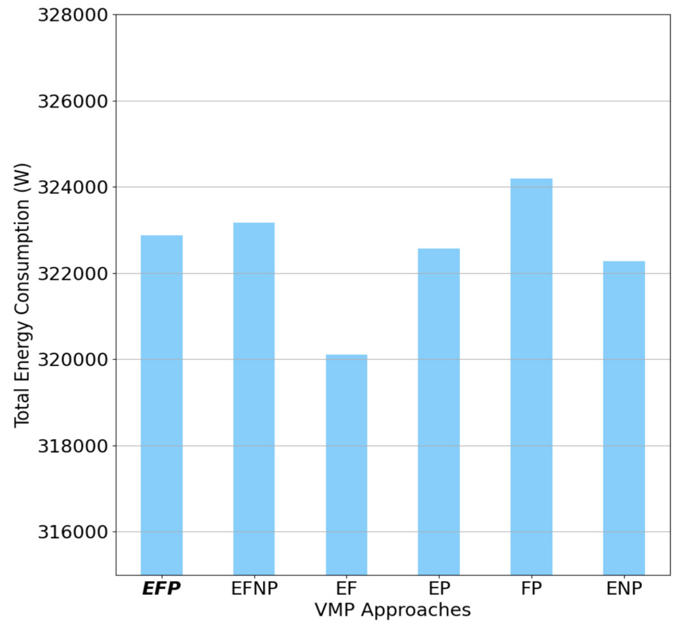

Fig. 6. Energy consumption of VMP algorithms.

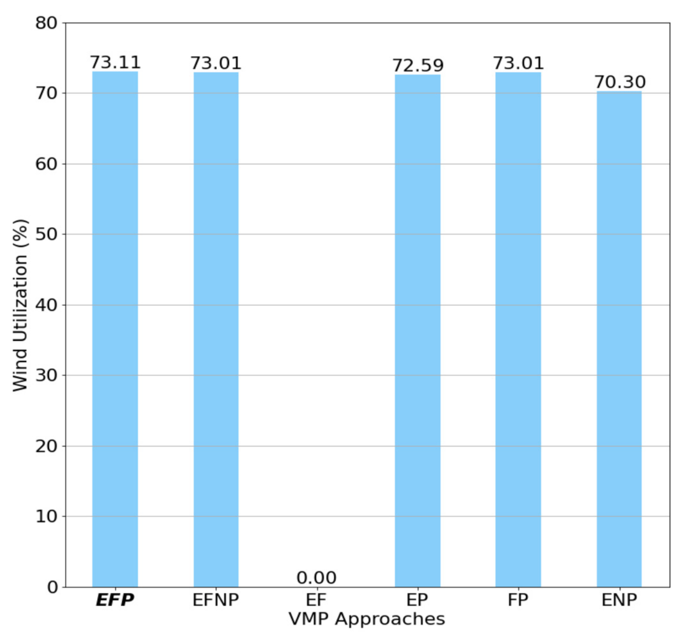

Fig. 7. Renewable energy utilization of VMP algorithms.

than the single-objective EP and FP. Since EF does not depend on renewable energy, its value is 0.

5.2.4. Carbon footprint

Fig. 8 illustrates the carbon footprint caused by brown energy consumption. Compared to the traditional coal-based method EF, VM placement strategies involve RES can reduce significant carbon footprint. To be specific, since EFP has the highest renewable energy utilization as discussed in the above section, which means that EFP will consume more renewable energy instead of traditional grid power. In this regard, EFP could achieve the trade-off between energy-aware and carbon-aware so that it emits the least carbon footprint. In general, carbon-aware methods such as EFNP and FP consider the effect of geographically distributed CFRs, which leads to less carbon footprint. Therefore, it should be noticed that although EP produces less brown energy, it will generate

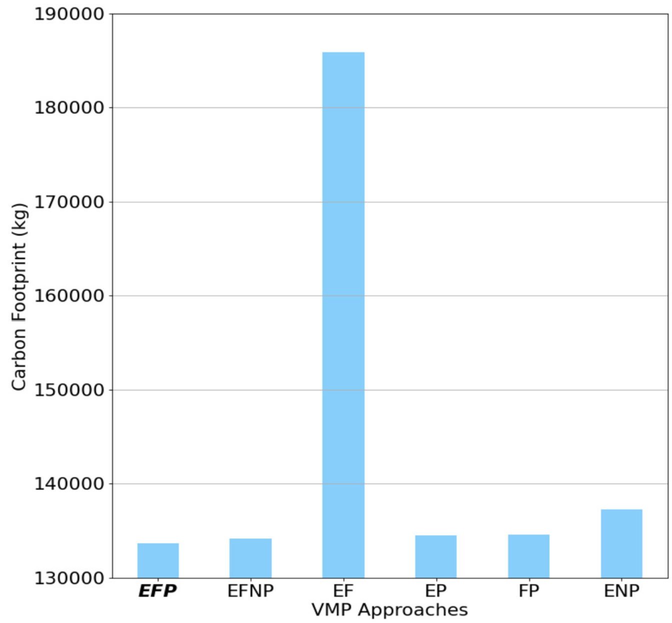

Fig. 8. Carbon footprint of VMP algorithms.

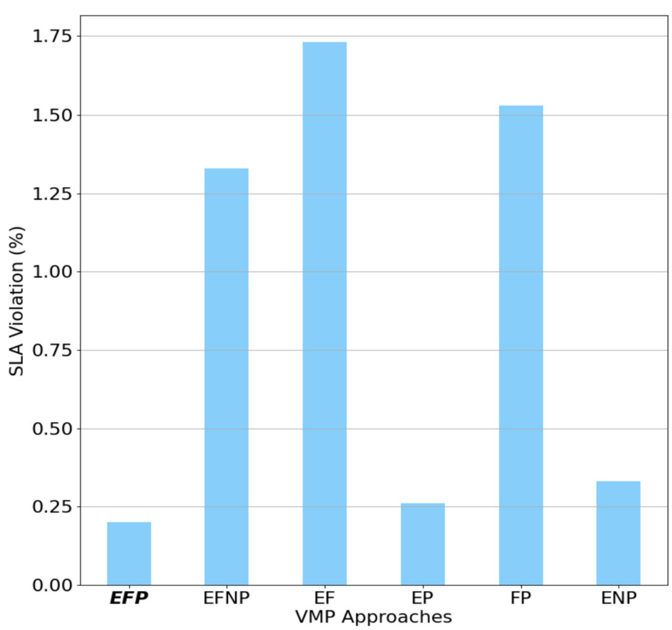

Fig. 9. SLAVs of VMP algorithms.

more carbon footprint than FP. In addition, EP performs better than ENP in this metric from the perspective of prediction.

5.2.5. SLA

The SLA violation rate is evaluated in the last metric as shown in Fig. 9, which ensures the cloud users' requests to be satisfied. SLA is defined as the ratio of the number of overloaded PMs that experience a level of 100% of CPU utilization at the current time or the future time slot to the number of active PMs. For example, if a PM's $C_{jk}^{CPU}(t)$ or $C_{jk}^{CPU}(t+\eta)$ surpasses its capacity S_{jk}^{CPU}, it is deemed as an SLA violation. Furthermore, the impact of workload prediction on the result is not mentioned. It should be emphasized here that all predictive algorithms predict the workload by default and allocate VMs accordingly. In this regard, predictive algorithms are superior to non-predictive algorithms except for FP. Specifically, EFP performs best and EP is superior to FP for the reason that it takes the energy consumption into account while processing VM allocation, which is directly related to CPU utilization. However, FP focuses on the temporal-varied CFRs and ignores the variation of CPU utilization. For the non-predictive algorithms, because EFNP needs to consider the impact of carbon emission on energy consumption, which may cause the situation that DCs with high CPU utilization could generate low carbon emission, and thus the possibility of SLA violation of EFNP is higher than ENP. Since EF does not divide DCs into two sets, it makes minimal changes to initial VM placement and it will have the highest SLA violation.

6. Conclusion

In this paper, the approach to balance energy consumption and carbon footprint in geographically data centers supplied by RES is investigated. To minimize the carbon footprint through maximizing RES utilization, temporal-variability wind speed and spatial-variability CFRs are considered. For this regard, a combined renewable energy generation prediction method based on EEMD and TCN is designed. Therefore, the wastage caused by the mismatch between the consumed energy and the generated RES can be reduced. In view of this, a novel VMP algorithm, named EFP, is proposed to achieve the trade-off between energy consumption and carbon footprint with predictive renewable energy. By using actual Google cluster traces and modeling the data centers located in Nevada, Arizona, California and Hawaii, the proposed EFP is compared with five baseline algorithms. The experimental results demonstrate that EFP is superior to those algorithms on brown energy and carbon footprint, which can maximize the renewable energy usage by 73.11% with the least SLA violation by 0.2%.

In future work, the proposed algorithm is expected to be evaluated in a real cloud infrastructure such as OpenStack. Furthermore, the energy cost will be considered as the optimization objective in the process of the carbon-aware algorithm.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (62162046), the Major Project of Inner Mongolia Natural Science Foundation (2019ZD15), Research and Application of Key Technology of Big Data for Discipline Inspection and Supervision (No. 2019GG372).

References

[1] F. Alharbi, Y.C. Tian, M. Tang, W.Z. Zhang, C. Peng, M. Fei, An ant colony system for energy-efficient dynamic virtual machine placement in data centers, Expert Syst. Appl. 120 (2019) 228–238.

[2] D. Alsadie, Z. Tari, E.J. Alzahrani, Online vm consolidation in cloud environments, in: 2019 IEEE 12th International Conference on Cloud Computing (CLOUD), IEEE, 2019, pp. 137–145.

[3] S. Aslam, S. Aslam, H. Herodotou, S.M. Mohsin, K. Aurangzeb, Towards energy efficiency and power trading exploiting renewable energy in cloud data centers, in: 2019 International Conference on Advances in the Emerging Computing Technologies (AECT), IEEE, 2020, pp. 1–6.

[4] Anton Beloglazov, Rajkumar Buyya, Managing overloaded hosts for dynamic consolidation of virtual machines in cloud data centers under quality of service constraints, IEEE Trans. Parallel Distrib. Syst. 24 (2013) 1366–1379.

[5] A. Beloglazov, R. Buyya, Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers, Concurr. Comput., Pract. Exp. 24 (2012) 1397–1420.

[6] K. Braiki, H. Youssef, Fuzzy-logic-based multi-objective best-fit-decreasing virtual machine reallocation, J. Supercomput. 76 (2020) 427–454.

[7] R.N. Calheiros, E. Masoumi, R. Ranjan, R. Buyya, Workload prediction using arima model and its impact on cloud applications' qos, IEEE Trans. Cloud Comput. 3 (2014) 449–458.
[8] B. Camus, A. Blavette, F. Dufossé, A.C. Orgerie, Self-consumption optimization of renewable energy production in distributed clouds, in: 2018 IEEE International Conference on Cluster Computing (CLUSTER), 2018, pp. 370–380.
[9] Y. Chen, X. Chen, W. Liu, Y. Zhou, A.Y. Zomaya, R. Ranjan, S. Hu, Stochastic scheduling for variation-aware virtual machine placement in a cloud computing cps, Future Gener. Comput. Syst. 105 (2020) 779–788.
[10] W. Deng, F. Liu, H. Jin, B. Li, D. Li, Harnessing renewable energy in cloud datacenters: opportunities and challenges, IEEE Netw. 28 (2014) 48–55.
[11] M. Esteban, J. Portugal-Pereira, Post-disaster resilience of a 100% renewable energy system in Japan, Energy 68 (2014) 756–764.
[12] Í. Goiri, M.E. Haque, K. Le, R. Beauchea, T.D. Nguyen, J. Guitart, J. Torres, R. Bianchini, Matching renewable energy supply and demand in green datacenters, Ad Hoc Netw. 25 (2015) 520–534.
[13] L. Grange, G. Da Costa, P. Stolf, Green it scheduling for data center powered with renewable energy, Future Gener. Comput. Syst. 86 (2018) 99–120.
[14] C. Gu, L. Fan, W. Wu, H. Huang, X. Jia, Greening cloud data centers in an economical way by energy trading with power grid, Future Gener. Comput. Syst. 78 (2018) 89–101.
[15] C. Gu, Z. Li, C. Liu, H. Huang, Planning for green cloud data centers using sustainable energy, in: 2016 IEEE Symposium on Computers and Communication (ISCC), IEEE, 2016, pp. 804–809.
[16] S.Y. Hsieh, C.S. Liu, R. Buyya, A.Y. Zomaya, Utilization-prediction-aware virtual machine consolidation approach for energy-efficient cloud data centers, J. Parallel Distrib. Comput. 139 (2020) 99–109.
[17] N.E. Huang, Z. Wu, A review on Hilbert-Huang transform: method and its applications to geophysical studies, Rev. Geophys. 46 (2008).
[18] A. Ibrahim, M. Noshy, H.A. Ali, M. Badawy, Papso: a power-aware vm placement technique based on particle swarm optimization, IEEE Access 8 (2020) 81747–81764.
[19] S. Jangiti, E.S. Ram, V.S. Sriram, Aggregated rank in first-fit-decreasing for green cloud computing, in: Cognitive Informatics and Soft Computing, Springer, 2019, pp. 545–555.
[20] S. Jangiti, et al. Hybrid best-fit heuristic for energy efficient virtual machine placement in cloud data centers, EAI Endorsed Transactions on Energy Web 7, 2020.
[21] M.A. Kaaouache, S. Bouamama, Solving bin packing problem with a hybrid genetic algorithm for vm placement in cloud, Proc. Comput. Sci. 60 (2015) 1061–1069.
[22] K. Kaur, S. Garg, G.S. Aujla, N. Kumar, A. Zomaya, A multi-objective optimization scheme for job scheduling in sustainable cloud data centers, IEEE Trans. Cloud Comput. (2019).
[23] M.I.K. Khalil, I. Ahmad, A.A. Almazroi, Energy efficient indivisible workload distribution in geographically distributed data centers, IEEE Access 7 (2019) 82672–82680.
[24] A. Khosravi, L.L. Andrew, R. Buyya, Dynamic vm placement method for minimizing energy and carbon cost in geographically distributed cloud data centers, IEEE Trans. Sustain. Comput. 2 (2017) 183–196.
[25] A. Khosravi, A. Nadjaran Toosi, R. Buyya, Online virtual machine migration for renewable energy usage maximization in geographically distributed cloud data centers, Concurr. Comput., Pract. Exp. 29 (2017) e4125.
[26] F. Kong, X. Liu, A survey on green-energy-aware power management for datacenters, ACM Comput. Surv. (CSUR) 47 (2014) 1–38.
[27] Z. Li, X. Yu, L. Yu, S. Guo, V. Chang, Energy-efficient and quality-aware vm consolidation method, Future Gener. Comput. Syst. 102 (2020) 789–809.
[28] Z. Liu, S. Cho, Characterizing machines and workloads on a Google cluster, in: International Conference on Parallel Processing Workshops, 2012.
[29] X. Lu, D. Jiang, G. He, H. Yu, Greenbdt: renewable-aware scheduling of bulk data transfers for geo-distributed sustainable datacenters, Sustain. Comput. Inform. Syst. 20 (2018) 120–129.
[30] P. Mell, T. Grance, et al., The nist definition of cloud computing, 2011.
[31] S. Mustafa, K. Bilal, S. Malik, S.A. Madani, Sla-aware energy efficient resource management for cloud environments, IEEE Access (2018) 1.
[32] R. Regaieg, M. Koubaa, Z. Ales, T. Aguili, Multi-objective optimization for vm placement in homogeneous and heterogeneous cloud service provider data centers, Computing (2021) 1–25.
[33] T. Renugadevi, K. Geetha, Task aware optimized energy cost and carbon emission-based virtual machine placement in sustainable data centers, J. Intell. Fuzzy Syst. (2021) 1–13.
[34] L. Wei, C.H. Foh, B. He, J. Cai, Towards efficient resource allocation for heterogeneous workloads in iaas clouds, IEEE Trans. Cloud Comput. 6 (2015) 264–275.
[35] J. Whitney, P. Delforge, Data center efficiency assessment–scaling up energy efficiency across the data center industry: Evaluating key drivers and barriers, Rep. IP, 14–08, NRDC and Anthesis, 2014.
[36] Z. Wu, N.E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (2009) 1–41.
[37] M. Xu, R. Buyya, Managing renewable energy and carbon footprint in multi-cloud computing environments, J. Parallel Distrib. Comput. 135 (2020) 191–202.
[38] M. Xu, A.N. Toosi, R. Buyya, A self-adaptive approach for managing applications and harnessing renewable energy for sustainable cloud computing, IEEE Trans. Sustain. Comput. (2020).
[39] R. Yu, J. Gao, M. Yu, W. Lu, T. Xu, M. Zhao, J. Zhang, R. Zhang, Z. Zhang, Lstm-efg for wind power forecasting based on sequential correlation features, Future Gener. Comput. Syst. 93 (2019) 33–42.
[40] K. Yunus, T. Thiringer, P. Chen, Arima-based frequency-decomposed modeling of wind speed time series, IEEE Trans. Power Syst. 31 (2015) 2546–2556.
[41] Y. Zhang, H. Sun, Y. Guo, Wind power prediction based on pso-svr and grey combination model, IEEE Access 7 (2019) 136254–136267.
[42] Y. Zhang, Y. Wang, X. Wang, Greenware: greening cloud-scale data centers to maximize the use of renewable energy, in: ACM/IFIP/USENIX International Conference on Distributed Systems Platforms and Open Distributed Processing, Springer, 2011, pp. 143–164.
[43] D.M. Zhao, J.T. Zhou, K. Li, An energy-aware algorithm for virtual machine placement in cloud computing, IEEE Access 7 (2019) 55659–55668.
[44] Z. Zhou, J.H. Abawajy, F. Li, Z. Hu, M.U. Chowdhury, A. Alelaiwi, K. Li, Fine-grained energy consumption model of servers based on task characteristics in cloud data center, IEEE Access 6 (2017) 27080–27090.
[45] Q. Zhu, H. Li, Z. Wang, J. Chen, B. Wang, Short-term wind power forecasting based on lstm. Power System Technology, 2017.

Daming Zhao received the B.S. degree in software engineering from the Harbin University of Science and Technology, Harbin, China, in 2016, and the M.S. degree in software engineering from Inner Mongolia University, Hohhot, China, in 2018, where he is currently pursuing the Ph.D. degree in computer science with the Department of Computer Science and Technology. His research interests include energy-aware for cloud computing, virtualization, and algorithm design.

Jiantao Zhou received the Ph.D. degree from the Tsinghua University in 2005. Since 1999, she has been on the faculty at Inner Mongolia University, China, where she is a professor now. Her research interests include formal methods, cloud computing and software engineering.