Accelerating the Energy Efficient Design of Traditional Data Centers Through Modeling*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Dr. Tathagata Bhattacharya
Department of Computer Science
Auburn University at Montgomery
Montgomery , USA
tzb0063@auburn.edu

2nd Dr. Xiaopu Peng
Department of Mathematics and Computing
Lander University, Greenwood
SC, USA
xzp0007@auburn.edu

3rd Taha Takreeti
Computer Science and Software Engineering
Auburn University
Auburn , USA
tka0002@auburn,edu

4th Dr. Jianzhou Mao
Computer Science and Software Engineering
Auburn University
Auburn, USA
jzm0122@auburn.edu

5th Dr. Ting Cao
Computer Science and Software Engineering
Auburn University
Auburn, USA
tzc0028@auburn.edu

6th Dr. Xiao Qin
Computer Science and Software Engineering
Auburn University
Auburn, USA
xqin@auburn.edu

7th Mostafa Rahgouy
Computer Science and Software Engineering
Auburn University
Auburn, USA
mzr0108@auburn.edu

2022 IEEE International Conference on Networking, Architecture and Storage (NAS) | 978-1-6654-5408-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/NAS55553.2022.9925343

***Abstract*—Power Management Strategies and the impact of carbon dioxide emission from the data centers across the globe have drawn significant attention worldwide. The rapid growth of energy consumed in data centers has lead to 1) huge costs 2) depletion of non-renewable resources such as coal and petroleum, and 3) emission of greenhouse gases like CO2 in the atmosphere. These greenhouse gas significantly contributes to the climate change of the earth. To tackle this challenge, our research deals with modeling the energy resources of data centers, thereby offering insights to reduce global carbon footprint and energy cost. In our model, we prioritize green energy consumption by eliminating the brown energy resources. In this process, we devise an algorithm that can determine the amount of CO2 emission in the atmosphere per hour by different energy resources. We create an energy model for data centers by incorporating the support vector regression algorithm. Our model is adroit at projecting energy consumed in data centers powered by green energy. Our experimental results confirm that our model consistently delivers high prediction accuracy in terms of energy usage in data centers. The model is expected to facilitate data analytic venues to optimize energy efficiency and sustainability for the development of future data centers.**

***Index Terms*—energy modeling, power management, green energy, brown energy, carbon footprint**

I. Introduction

Aiming to battle the global carbon footprint, we propose a novel approach of modeling the carbon footprint of modern data centers govern by energy-saving techniques. This research is inspired by the state-of-the-art studies anchored in (1) energy-efficiency of the data centers (Section II-A), (2) sustainable computing (Section I-B), and (3) carbon emission reduction (Section I-C).

A. Energy-Efficient Data Centers

According to reports by the U.S. Department of Energy's Lawrence Berkeley National Lab, data centers in the U.S consume 70 billion Kilowatt-hours of power each year which costs around 7 billion. Data centers throughout the globe are growing rapidly as the cloud computing infrastructures are expanding. Data centers worldwide are the reason for tremendous IT growth, extensive cloud infrastructure, and a shift to virtualization from paper-based model [1]. From recent online banking transactions to the growth in e-commerce and multimedia; from the usage of electronic medical data to the

978-1-6654-5408-7/22/$31.00 ©2022 IEEE
Authorized licensed use limited to: Auburn University. Downloaded on November 17,2026 at 23:18:02 UTC from IEEE Xplore. Restrictions apply.

cryptocurrency and building firewall; energy demand is never ending [2], [3]. Recent reports of data centers claim that these data centers consume over 1000 W/m2 of power — more than 10 times of that required by a typical commercial office space [4]. Another study finds out that a typical data center consumes energy as large as that of 25,000 households and 100 to 200 times of electricity consumed by standard office space.

Though all these cutting-edge technologies are prominent in managing the energy consumption in data centers, most of the existing schemes overlook modeling the clean energy resources (e.g, solar and wind). Since renewable energy resources are less expensive [5], it is beneficial to cut down brown energy sources by incorporating green energy to brings down the cost for energy consumption of the data centers.

B. Sustainability Computing

To ensure earth's sustainability, researchers have been attempting to blend green computing in all spheres. It has become extremely necessary to use the non renewable and renewable energy resources wisely ensuring their availability for the next generation [6]. But on a contradiction , electric grids burn a hefty amount of brown energy resources to meet the never-ending power needs of the data centers [7]. This tremendous power consumption leads to increased emission of greenhouse gases which is causing global pollution [8] . Due to extensive consumption of non-renewable resources like coal, petroleum, these resources are getting depleted rapidly [9]. This huge energy consumption will result in the destruction of the non-renewable energy resources by 2030 [10]. This issue sparks us to come up with energy models aiming to help system administrators judiciously manage the energy resources of the data centers.

C. Carbon Emission Reduction

Data centers are responsible for approximately two percent of the global carbon footprint [11]. A massive amount of energy consumed by the data centers leads to high carbon emission [12]. It is desirable to reduce the data center's carbon footprint by adopting new technologies, thereby launching future green data centers [13]. With the advancement of parallel and cloud computing powered by virtualization, capping the carbon footprint of data centers becomes possible [14]. For example, workload scheduling and geographical load balancing are common techniques to address the issue of carbon emission in the air [15]. Recently, development has offered the green slots to increase the solar energy consumption, which reduces the CO2 elimination by 39% [16]. Alongside, smart grids are expected to cause a green revolution for the future data centers managing their carbon footprint [17].

D. Contributions

All the aforementioned breakthroughs advocate for extensive technological developments at huge expenses. In this research, we delve into a novel approach to ramping up a carbon footprint predictor for energy usage of data centers fueled by renewable and traditional energy. Additionally, we build a model to holistically estimate the energy demand of integrated energy conservation techniques in data centers. In particular, our model builds an overall energy usage pattern by incorporating clean energy and eliminating most of the brown energy resources. Also, we prove that our model embraces accuracy and efficiency strengths in terms of power modeling.

In contrast to the pre-existing models, our power model is based on the overall load consumption of a data center.Instead of keeping track of the energy consumption of individual data center components (e.g., chillers, UPS, virtual nodes) at the micro-level, our macro-level model treats a mixture of energy-conservation schemes and modules as a seamlessly integrated system. Additionally, our power consumption model applies to modern data centers substituting green energy for brown energy to mitigate carbon footprint issues.

We summarize our contributions to this research field as follows:

- We design a carbon footprint calculator for the tradiitonal data center for estimating the carbon emission (see sec III-C)of the data center powered by heterogeneous energy resources.
- We design an energy consumption model for data centers that balances clean and non renewable energy resources.
- We evaluate our prediction accuracy and results rigorously.
- We discuss the drawbacks of our model and compare with pre-existing models.

II. Background Study

Considering the importance of building energy-efficient data centers, researchers and engineers have come up with several models to curb power consumption and CO2 production of the data centers. This section presents the background of energy-efficient computing that brings solidarity and uniqueness to our research.

A. Energy-Efficient Computing

Data centers are mostly powered by electric grids, which are a huge consumer of brown energy. Data centers have multiple layers of abstraction, each of which is very complex in architecture, concept, and technology. Feature extraction is a very common technique to model the energy optimization of data centers [18]. Measuring the raw power consumption through hardware and software has been intrinsic research towards power optimization [19]. Since these techniques are not portable, current data centers depend on PDUs to optimize energy efficiency [20].

B. Energy Consumption Modeling

Data centers comprise a variety of components, the energy-consumption models of which are summarized in this subsection. Servers are very an essential piece in a data center as the servers correspond to significant load requirements. Hence, constructing a power model for servers is crucial in reducing the total energy consumption of data centers. Server power

Authorized licensed use limited to: Auburn University. Downloaded on November 17,2026 at 23:18:02 UTC from IEEE Xplore. Restrictions apply.

models depend on other collaborating components. Roy *et al.* devised a simple yet elegant server power model, which is a summation of the load consumption of CPU and memory [21]. Apart from that, Orangerie *et al.* orchestrated server energy as a function of the energy consumption of system boot and system halt [22]; Li *et al.* constructed a power model as a function of energy consumed by active virtual machines running on servers [23].

C. Contributions

With the advent of machine learning techniques, researchers have investigated supervised and unsupervised learning models for power consumption in the data center. For instance, Berral *et al.* demonstrated a way of modeling expected processor and disk usage based on the M5P algorithm [24]. Moreno *et al.* developed a MLP-based neural network to predict power from clients' historical data [25]. Recently, *iOverbook* - an autonomous, online, and intelligent overbooking strategy - has drawn attention in the field of power modelling [26]. All the aforementioned models have significantly contributed to capping the power consumption of data centers from the perspectives of cooling and server power consumption. It remains an open issue to the model power consumption of data centers powered by a diversity of energy resources. To fill this technological gap, we propose a power-consumption model for modern data centers that are powered by a mixture of green and brown energy to reduce the cost and carbon footprint.

D. Capping Carbon Footprints in Data Centers

Green computing has opened a new horizon for the energy-efficient computing community to cap carbon dioxide elimination from data centers. To tackle this issue, researchers have delved into the development of novel technologies like virtualization, Live VM migration, geographical load balancing, scheduling policies, and to name just a few. Virtualization by consuming less energy and creating negligible e-waste contribute towards reducing the carbon footprint from data centers [27]. Live VM migrations facilitate virtual-machine consolidation to boost energy efficiency and edging carbon emission [28] [29]. Unlike the above techniques, our solution has its root in a power model that efficiently substitutes brown energy resources with green energy to reduce the carbon elimination from data centers.

III. System Overview

We start this study by designing architecture as an underpinning of our model. In what follows, we outline the system architecture in Section III-A. In Section III-B, we discuss the methodology applied to the development of the model.

A. System Architecture

Fig. 1 depicts the framework of a typical *Data center*, which compromises a large number of components like physical nodes, UPS, Battery, chillers, and heterogeneous energy resources to power and maintain the operations of the data center.

More often than not, data centers consume a massive amount of energy from brown energy resources (see, for example, Coal and Petroleum at the bottom of Fig. 1), thereby releasing an enormous amount of heat and greenhouse gases in the air to cause a major adverse impact on the environment. Recent evidence confirms that data centers consume energy as large as that of 25,000 households and 100 to 200 times of electricity consumed by standard office space [30]. The high energy cost can be curtailed by deploying green energy like solar and wind (see the green energy box at the bottom of Fig. 1).

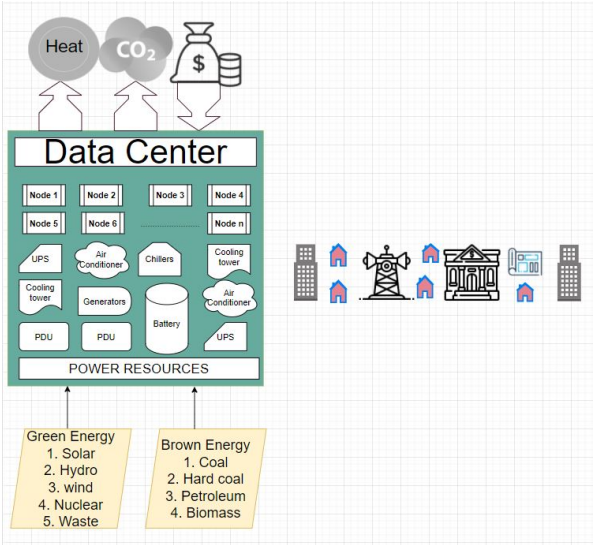

Fig. 1. The framework orchestrates a diversity of green and brown energy sources to maximize the energy efficiency of data centers.

Instead of using massive coal, petroleum and fossil fuel energy, we advocate for balancing renewable(solar, nuclear, wind, and hydro energy) and non renewable energy resources(coal, petroleum etc.)to optimize the hefty energy demand and diminish the operational cost of the data centers. It is expensive and time consuming to transform a traditional data center into a green data center deliberately. But a proper balance of energy resources can not only reduces energy costs but also fulfills the UN sustainability goals by lowering greenhouse gas emissions in the air thus ensuring sustainability of the world. Our data set shows that solar, wind and hydro energies are notable sources of power to reduce brown energy usage.

B. Methodology

During the course of designing a power model for data centers, our first and foremost task is to investigate the popular modeling approaches to obtain a rich data set of energy usage per hour. After harnessing the data set, we portray the energy usage in graphical forms to come up with a clear picture of a data center's energy consumption and carbon footprint at any particular time of the day. This analysis strategy helps us in building a theoretical underpinning of our model. In what follows, we articulate our research methodology.

Authorized licensed use limited to: Auburn University. Downloaded on November 17,2026 at 23:18:02 UTC from IEEE Xplore. Restrictions apply.

- **Support Vector Regression.** After constructing a linear model from our collective data set, we test the model's accuracy and efficiency at a particular time instance. We notice that linear model does not fit during peak hours. Therefore, we decide to implement a support vector regression model(because of it's flexibility and non linearity) [31] to examine its efficacy to predict energy efficiency.
- **Other Models** In a future study, we will compare the SVR model against the other modeling methodologies (e.g., Ridge regression, polynomial regression) in terms of accuracy and efficiency. Also we will tune the parameters γ, ϵ and cost.

TABLE I. Symbol and Annotation

Symbol	Annotation
$\alpha 1$	hydro pump energy
$\alpha 2$	solar energy
$\alpha 3$	nuclear energy
$\alpha 4$	wind energy
$\beta 1$	coal energy
$\beta 2$	biomass energy
$\beta 3$	fossil gas
$\beta 4$	oil energy
$\gamma 1$	carbon footprint of green energy
$\gamma 2$	carbon footprint of brown energy
δ	total carbon footprint
Δ	total energy input
α	green energy
β	brown energy

C. Carbon Footprint Calculator

We devise a carbon footprint calculator to estimate the carbon footprint of data centers. To reduce the carbon dioxide emission from data centers, our first challenge is to gauge the carbon footprint of various energy resources associated with the data centers. We address this challenge by designing an algorithm to quantitatively compute a data center's carbon footprint, which leads to the total carbon dioxide emissions.

Let $t = \{\alpha, \beta\}$ denote a set of energy types, in which α and β represent green energy and brown energy, respectively. The calculator presented in Algorithm 1 measures carbon footprint according to a given energy type in set t.

Algorithm 1 detects the carbon footprint from both green and brown energy resources in a data center. The data for carbon footprint for the brown energy resources are available at [32]. The values that for green energy carbon footprint can be obtained from [33], [34], [35], [36]. We do our little bit of calculations and play with these value and brought to a specific unit pound-per-hour(lbph). Then we used simple arithmetic operations to find the amount of lbph of CO2 emissions per KWh burning of the different fuels. Though we include nuclear energy in this algorithm, one may opt for pruning the nuclear energy usage from the model. Nuclear energy resources contribute a very less carbon footprint but such resources are highly radioactive. It is evident that nuclear energy's radioactivity causes cell damage, impotency, and cancer [37]. Because we are focusing on sustainability, we ensure that nuclear energy only plays a supporting role in our model. Also, it was initially considered that biomass is a zero-carbon energy resource [38]. Recent studies, however, indicate that burning biomass produces a carbon footprint double the amount of that produced by coal energy. For this reason, we treat biomass as brown energy in the model though it is traditionally classified as renewable energy [36].

Algorithm 1: Carbon Footprint Calculator

Input : Δ(KWh)
Output: δ(lbph)
if Δ *is KWH* **then**
  go to next step
end
else
  convert Δ in KWH
end
if $t == \alpha$ **then**
  return $\gamma 1$= *0.04078552* * $\alpha 1$ + *0.110231* * $\alpha 2$ + *0.0264555* * $\alpha 3$ + *0.0242508* * $\alpha 4$
end
else if $t == \beta$ **then**
  return $\gamma 2$ = *2.21* * $\beta 1$ + *(2 * 2.41)* * $\beta 2$ + *0.92* * $\beta 3$ + *2.11* * $\beta 4$
end
δ = $\gamma 1$ + $\gamma 2$

IV. Observations and Trends

A. Data Collection and Data Handling

We investigate a panel dataset by a traditional data-center. This dataset elaborately mentions the energy usage per hour of the data center over a year. After obtaining the panel data set, we perform multiple machine-learning algorithms to exploit the entire scenario. Also, the carbon footprint of this data center was unavailable. Henceforth, we devise the carbon footprint calculator to estimate the carbon emission per hour from the data center. We segmented the data set into weekday-weekend, monthly, quarterly energy usage of the data center to experiment with the energy data of the data center.

B. Monthly Energy Usage

Figs. 2, 3, plot the monthly brown energy usage and monthly green energy usage during one year. In Figs. 2 and 3 the boxes show the maximum Energy utilization range whereas the bars parallel to the y-axis at the top and bottom shows the range of their energy usage, For example in Fig. 2 the range of the energy usage for February is 4800 KWh to 20500 KWh. The max utilization of brown energy in this plot for February ranges from 7000 KWh to 12000 KWh and the average brown energy usage is 9000 KWh.

The statistical data plotted in Figs. 2 and 3 intuitively demonstrate dynamic energy usage in the data center. An intriguing observation drawn from Figs. 2 and 3 is that except for a few months (e.g., February and October), brown energy

Authorized licensed use limited to: Auburn University. Downloaded on November 17,2026 at 23:18:02 UTC from IEEE Xplore. Restrictions apply.

has predominantly been used by the data center where green energy usage is minimum.

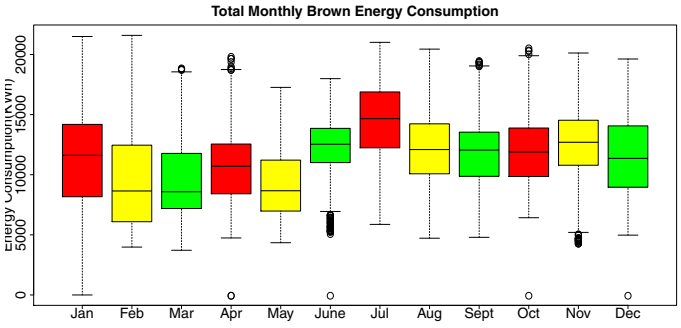

Fig. 2. The monthly brown energy utilization for the data center from January to December.

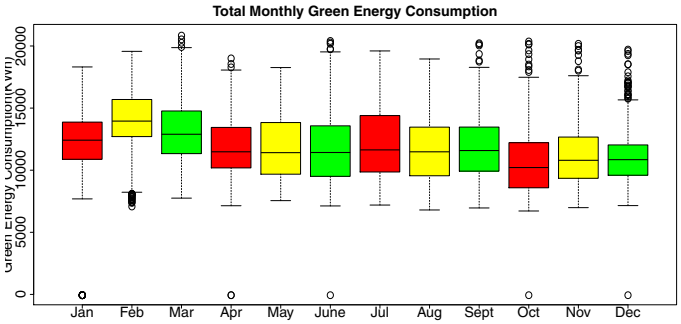

Fig. 3. The monthly green energy utilization for the data center from January to December.

The results illustrate that the average monthly energy consumption for the data center ranges from 25000 to 35000 KWh, which is responsible for high heat dissipation and global carbon footprint. The data are evidence that there is a pressing need to cut back brown energy while fostering green energy usage.

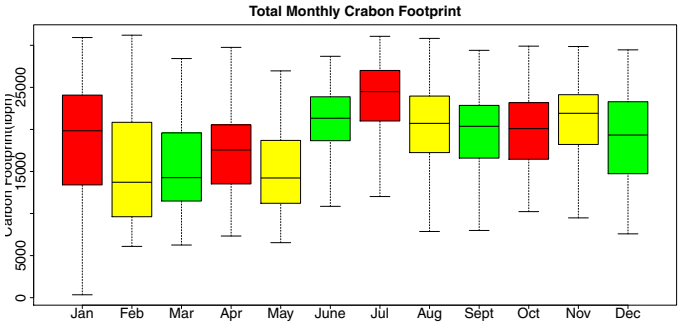

Fig. 4. Monthly carbon footprint of the data center from January to December.

Fig. 4 elaborates the monthly carbon footprint of the data center. Similar to the previous figures, Fig. 4 shows the minimum, maximum, and average carbon footprint of each month. We conclude that the carbon footprint of the data center is proportionately related to renewable and non-renewable energy usage. For example, carbon footprint in July reaches the peak level due to the excessive amount of brown energy usage.

C. Daily Energy Usage

Now we dive into data analysis with the perspective of daily energy consumption. We plot the two energy usage data in Fig. 5 and Fig. 6 to intuitively visualize the brown and green energy usage in a data center with regard to daily trends. Here, the Y axis symbolizes energy usage and the x-axis represents time. In each figure, we plot a total of 24 points for the data center's energy consumption in the unit of hours.

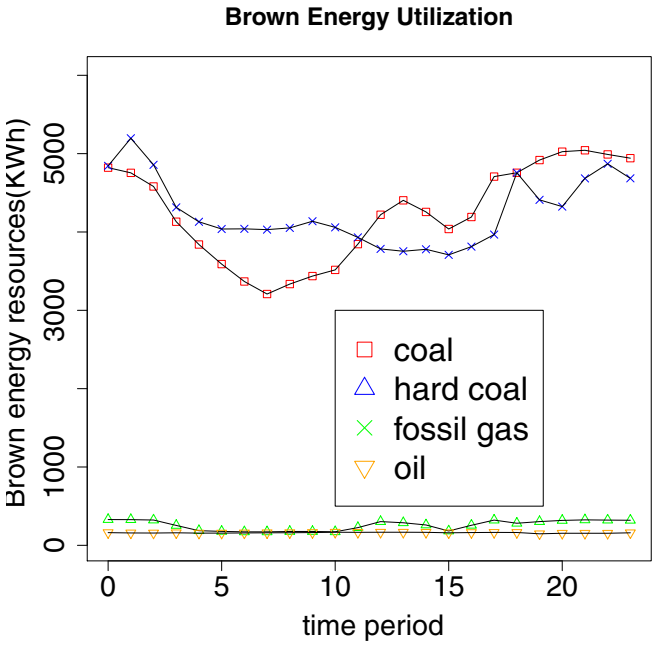

Fig. 5. The brown energy utilization of a data center within one day with a measurement interval of one hour.

Fig. 5 depicts that brown energy usage is a major player in comparison to of green energy to supply the power need of the data center. Fig. 5 reveals that among the brown energy resources used in the data centers, coal and hard coal are utilized in the maximum amount causing a heavy carbon footprint. Another observation indicates that these coal and hard coal are mostly used during peak time from 11 a.m. to midnight. On the flip side of the coin, Fig. 6 shows that there exists a small amount of green energy consumed to meet the energy demand of the data center. Among the pile of various resources, solar, hydro pump, and nuclear energy are widely applied to supply power to the data center. Compared to brown energy usage, the utilization of green energy is significantly low. The data center uses a significant amount of nuclear energy. Generally, though, the radioactive rays from expensive nuclear energy are detrimental to the health of people in neighboring areas.

After comparing Figs. 5 and 6, we conclude that these maximum brown energy usage by data centers is the key driver

Authorized licensed use limited to: Auburn University. Downloaded on November 17,2026 at 23:18:02 UTC from IEEE Xplore. Restrictions apply.

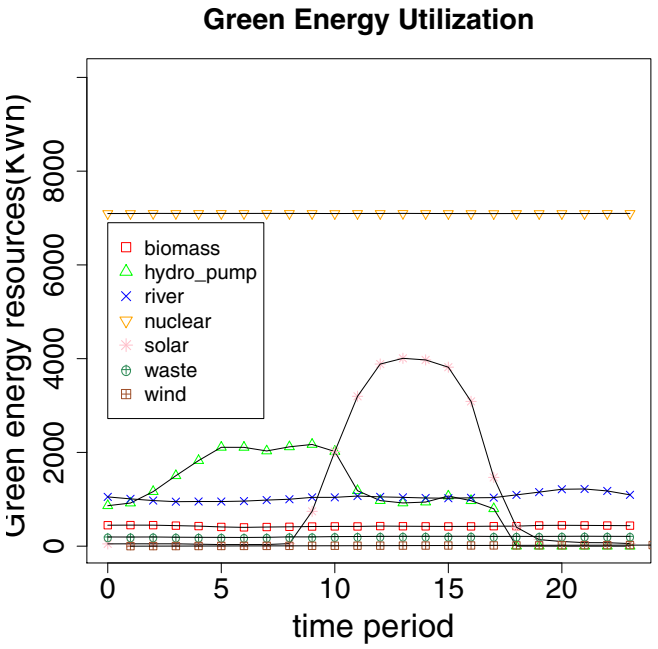

Fig. 6. The green energy utilization of a data center within one day with a measurement interval of one hour.

behind carbon emission from the data centers. Henceforth, we decide to come up with an energy model that aims to actively boost utilization of green energy to reduce the global carbon footprint of the data centers.

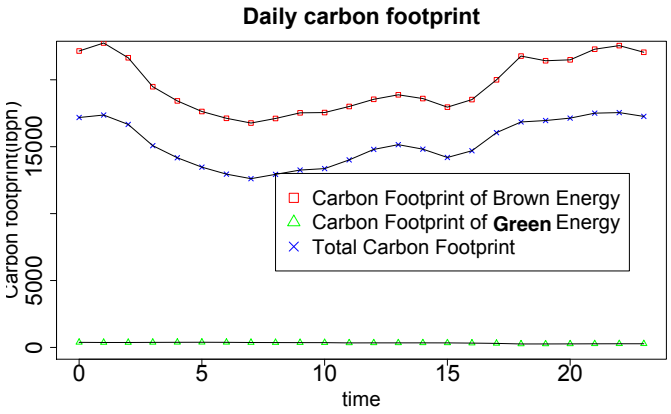

Fig. 7. The carbon footprint of a data center on each day with a measurement interval of one hour for brown energy, green energy and total load.

Fig. 7 unveils that during daytime the green energy carbon footprint is the least and brown energy carbon footprint is huge. The range of carbon footprint of green energy lies between 200 to 700 lbph whereas that for brown energy is around 15000 lbph causing a huge impact on the environment and sustainability. Also, it is worth mentioning that we have eliminated some of the brown energy and green energy resources as their contributions were not that significant. Hence, the total carbon footprint in Fig. 7 is not equal to the projected carbon footprint from the brown and green energy resources.

V. Modeling Energy Usage

In our prior study reported in [39], we have developed a linear- regression-based energy model for data centers. To capture data-enter energy usage, we extend the baseline model using *support vector regression* to enhance modeling accuracy as well as improving classification strategies. Unlike the other regression models that minimize the error between real and predicted values, the support vector regression technique (SVR) attempts to fit the best line within a threshold value. Such a threshold value is defined as the distance between hyper-plane and boundary line [40].

To start with support vector regression, we decide to play with eps regression and different kernel functions such as linear, polynomial, sigmoid, and radial. A kernel detects the hyper-plane in the higher dimensional space while keeping a constant computational cost [41]. Usually, the computational cost rises when the data dimension increases. When we are unable to find a separating hyper-plane in a given dimension, we ought to move in a higher dimension.

- We split the data set randomly into training and testing data. First, we experiment with 75 percent training and 25 percent testing data
- we perform support vector regression or eps regression with four different kernels, namely, the linear, polynomial, radial, and sigmoid kernels on this data set.
- We run each of these kernels on the testing data and measure the efficiency of this model on the testing data by calculating the R-squared value of the predicted and actual value.
- Once figuring out the best possible modeling strategy, we design the power model of the data center.

We evaluate our model with four parameters: (1) Mean-Squared-Error(MSE), (2) R-Squared Value on training data set after cross-validation, (3) R-squared value on testing data set, and (4) the number of support vectors. Mean Squared Error or MSE detects the average squared error between estimated and the original value; R squared error is a parameter for estimating the accuracy of the model or to demonstrate how good the model fits; support vectors define the points that are closest to the hyper-plane.

TABLE II Evaluation Parameters

Parameter	Linear	Polynomial	Radial	Sigmoid
MSE	2252429	3249210	1048172	39E+10
R-Squarred(Train)	0.8943514	0.8469862	0.9506667	0.6137753
R-squarred (Test)	0.8256411	0.7710267	0.9163271	0.6467810
Support vectors	4887	5016	3960	6527

Tables II, project the Mean Squared Error (MSE) and accuracy (R- squared Value) for the training and testing data. During the course of the experiments, we figure out that the MSE and R squared values change minutely when we tune the training and testing dataset ratio or number of cross-validations on the training data set. We set this aside for our next experiment. Table II suggests that the radial kernel performs the best in terms of MSE and R-Squared values.

Authorized licensed use limited to: Auburn University. Downloaded on November 17,2026 at 23:18:02 UTC from IEEE Xplore. Restrictions apply.

Hence, we conclude after a thorough analysis that the radial kernel is the most appropriate candidate for this model thanks to its high accuracy value (R-squared value) and low MSE. Therefore, we build the energy model using a Gaussian radial basis function kernel or *RBF*. The equation (Eq. 1) for the energy model incorporates the RBF kernel as follow.

$$E_{Energy-outcome} = \beta + (k \times (time + temp + hum + w_{speed} + E_{fossil-coal} + E_{fossil-oil} + E_{hydro-run} + Ehydro - storage + E_{water-reservoir} + E_{solar} + E_{wind})) \quad (1)$$

In Eq. 1 k is the kernel function (RBF kernel) in designing the energy model for data centers. β, which refers to the cost of the function, is chosen to be 1 in this case. Here $time$ $temp$, hum and w_{speed} defines time ,temperature , humidity, and wind speed of the hour, respectively. $E_{fossil-coal}$, $E_{fossil-oil}$, $E_{hydro-run}$, $E_{hydro-storage}$, $E_{water-reservoir}$, E_{solar}, E_{wind} denote energy from fossil coal, energy from hydro run of water bodies, energy from water storage, energy from water reservoirs, solar energy and wind energy, respectively. If we dig deep into Eq. 1, it is visible that this energy equation accommodates more renewable energy resources with a very few brown energy counterparts. Even, these brown energy resources (e.g., fossil coal, fossil oil) captured in our power model have lower carbon footprint than the other types of brown energy resources such as hard coal, lignite, and petroleum. Henceforth, we claim that our energy model reveals the insights that green energy resources significant reduce heat dissipation as well as carbon emission while maintaining the energy supply.

Now, we are positioned to solve the kernel equation Eq. 1. The RBF kernel function can be represented as

$$k(x_i, x_j) = \exp(\gamma \times \mod (x_i - x_j)^2) \quad (2)$$

x_i and x_j in Eq. 2 represent the distance between two decision boundaries. Here, we measure the distance as 0.2 (ϵ=0.1). We set up the default γ value to 0.090909. We obtain Eq. 3 below after substituting x_i, x_j, and γ values into Eq. 2.

$$k(x_i, x_j) = \exp(0.090909 \times \mod (0.1 + 0.1)^2) \quad (3)$$

Substituting the k value from Eq. 3 in Eq. 1, we derive the energy model as.

$$E_{Energy-outcome} = 1 + (\exp \times 0.00363636 \times (time + temp + hum + w_{speed} + E_{fossil-coal} + E_{fossil-oil} + E_{hydro-run} + Ehydro - storage + E_{water-reservoir} + E_{solar} + E_{wind})) \quad (4)$$

Next, we plot the graph between the actual and predicted value in Fig. 8, where our energy model is created by the virtue of the SVR with RBF kernels.

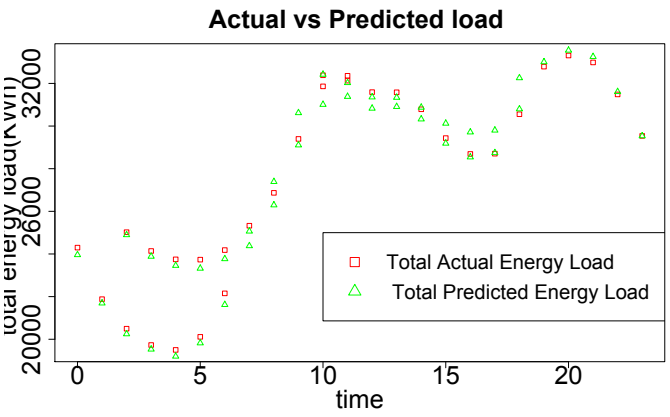

Fig. 8. The results demonstrate the discrepancy between actual and predicted energy values

Fig. 8 depicts the accuracy of our proposed model. The results intuitively show the accuracy of the predicted values compared to the actual ones. Unlike the results obtained from our previous study (see the details in [39]) where the efficiency of the model decreased to 70% from 90% during a peak time, Fig. 8 unravels that our model is quite efficient during the peak time as well. Most of the actual and predicted values lies close to each other and; hence, we confirm that SVR delivers a superb job in modeling energy usages to offer insights helping to cap the CO2 emission from data centers. Also, we claim since natural resources of energy (eg.solar, wind, hydro energies) are renewable and consumption of these energies is absolutely free, this model will be effective in reducing the operational cost of data centers thereby ensuring sustainability and a better financial strategy.

We want to mention that we have many similar x axis values in Fig. 8. This is because we have taken multiple energy data of each day over a period of one year. On the contrary, we can mention that in Fig. 8 during the non peak hours the actual and predicted values are not accurate. This shows the model has no over-fitting problems. As mentioned previously , we have already produced another model that aptly predicts the total energy load during the non peak hours of the day. Hence the data center managers can use that model for predicting the energy load.

VI. Conclusions and Future Directions

In this study, we aim to construct a power model to facilitate data analytic venues for optimizing energy efficiency of green data centers. Thanks to the Gaussian RBF kernel of the support vector regression approach (SVR), we made use of SVR to model the energy usage in green data centers. More specifically, we designed an algorithm calculating the amount of CO2 emission in the atmosphere per hour by various energy resources. We conducted extensive experiments to demonstrate that our model is conducive to predicting energy usages in data centers powered by green energy. Our experimental results confirm that our model consistently delivers high prediction accuracy in terms of energy usage in data centers.

Authorized licensed use limited to: Auburn University. Downloaded on November 17,2026 at 23:18:02 UTC from IEEE Xplore. Restrictions apply.

As a future research direction, we intend to experiment with a variety of statistical learning models such as lasso regression. In a second future direction, we have a keen desire to create a carbon footprint model for data centers. The novel carbon footprint model is expected to efficiently predict vital measures in cutting back carbon footprint and improving sustainability for future data centers.

REFERENCES

[1] L. Krug, M. Shackleton, and F. Saffre, "Understanding the environmental costs of fixed line networking," in *Proceedings of the 5th international conference on Future energy systems*, 2014, pp. 87–95.

[2] K. T. Rao, P. S. Kiran, and L. S. S. Reddy, "Energy efficiency in datacenters through virtualization: A case study," *Global Journal of Computer Science and Technology*, 2010.

[3] H. Yuan, C.-C. J. Kuo, and I. Ahmad, "Energy efficiency in data centers and cloud-based multimedia services: An overview and future directions," in *International Conference on Green Computing*. IEEE, 2010, pp. 375–382.

[4] J. Stein, "More efficient technology will ease the way for future data centers," in *Proceedings*, 2002.

[5] A. Ciarreta, M. P. Espinosa, and C. Pizarro-Irizar, "Is green energy expensive? empirical evidence from the spanish electricity market," *Energy Policy*, vol. 69, pp. 205–215, 2014.

[6] M. EVANS, "What is environmental sustainability," 2020, last accessed 21 July 2020. [Online]. Available: https://www.thebalancesmb.com/what-is-sustainability-3157876#citation-1

[7] F. Kong and X. Liu, "A survey on green-energy-aware power management for datacenters," *ACM Computing Surveys (CSUR)*, vol. 47, no. 2, pp. 1–38, 2014.

[8] D. Bouley, "Estimating a data center's electrical carbon footprint," *Schneider Electric White Paper Library*, pp. 14–22, 2011.

[9] J. A. Krautkraemer, "Nonrenewable resource scarcity," *Journal of Economic literature*, vol. 36, no. 4, pp. 2065–2107, 1998.

[10] M. McNerney, "The data center dilemma: Is our data destroying the environment?" 2019. [Online]. Available: https://www.datacenterknowledge.com/industry-perspectives/data-center-dilemma-our-data-destroying-environment

[11] M. R. Mishi, "Matching geographically distributed micro-datacenters' energy consumption with renewable energy," 2019.

[12] R. E. Brown, R. Brown, E. Masanet, B. Nordman, B. Tschudi, A. Shehabi, J. Stanley, J. Koomey, D. Sartor, P. Chan *et al.*, "Report to congress on server and data center energy efficiency: Public law 109-431," Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), Tech. Rep., 2007.

[13] P. Kurp, "Green computing," *Communications of the ACM*, vol. 51, no. 10, pp. 11–13, 2008.

[14] A. Khosravi and R. Buyya, "Energy and carbon footprint-aware management of geo-distributed cloud data centers: A taxonomy, state of the art, and future directions," in *Sustainable Development: Concepts, Methodologies, Tools, and Applications*. IGI Global, 2018, pp. 1456–1475.

[15] C. Ren, D. Wang, B. Urgaonkar, and A. Sivasubramaniam, "Carbon-aware energy capacity planning for datacenters," in *2012 IEEE 20th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, 2012, pp. 391–400.

[16] Í. Goiri, M. E. Haque, K. Le, R. Beauchea, T. D. Nguyen, J. Guitart, J. Torres, and R. Bianchini, "Matching renewable energy supply and demand in green datacenters," *Ad Hoc Networks*, vol. 25, pp. 520–534, 2015.

[17] Z. Asad and M. A. R. Chaudhry, "A two-way street: Green big data processing for a greener smart grid," *IEEE Systems Journal*, vol. 11, no. 2, pp. 784–795, 2016.

[18] Y. Kodama, S. Itoh, T. Shimizu, S. Sekiguchi, H. Nakamura, and N. Mori, "Power reduction scheme of fans in a blade system by considering the imbalance of cpu temperatures," in *2010 IEEE/ACM Int'l Conference on Green Computing and Communications Int'l Conference on Cyber, Physical and Social Computing*. IEEE, 2010, pp. 81–87.

[19] W. L. Bircher and L. K. John, "Core-level activity prediction for multicore power management," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 3, pp. 218–227, 2011.

[20] J. W. Smith, A. Khajeh-Hosseini, J. S. Ward, and I. Sommerville, "Cloudmonitor: profiling power usage," in *2012 IEEE Fifth International Conference on Cloud Computing*. IEEE, 2012, pp. 947–948.

[21] S. Roy, A. Rudra, and A. Verma, "An energy complexity model for algorithms," in *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, 2013, pp. 283–304.

[22] A.-C. Orgerie, L. Lefèvre, and I. Guérin-Lassous, "Energy-efficient bandwidth reservation for bulk data transfers in dedicated wired networks," *The Journal of Supercomputing*, vol. 62, no. 3, pp. 1139–1166, 2012.

[23] Y. Li, Y. Wang, B. Yin, and L. Guan, "An online power metering model for cloud environment," in *2012 IEEE 11th International Symposium on Network Computing and Applications*. IEEE, 2012, pp. 175–180.

[24] J. L. Berral, R. Gavalda, and J. Torres, "Adaptive scheduling on power-aware managed data-centers using machine learning," in *2011 IEEE/ACM 12th International Conference on Grid Computing*. IEEE, 2011, pp. 66–73.

[25] I. S. Moreno and J. Xu, "Neural network-based overallocation for improved energy-efficiency in real-time cloud environments," in *2012 IEEE 15th International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing*. IEEE, 2012, pp. 119–126.

[26] F. Caglar and A. Gokhale, "ioverbook: intelligent resource-overbooking to support soft real-time applications in the cloud," in *2014 IEEE 7th International Conference on Cloud Computing*. IEEE, 2014, pp. 538–545.

[27] S. Agarwal and A. Nath, "Desktop virtualization and green computing solutions," in *Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012), December 28-30, 2012*. Springer, 2014, pp. 1439–1449.

[28] P. G. J. Leelipushpam and J. Sharmila, "Live vm migration techniques in cloud environment—a survey," in *2013 IEEE Conference on Information & Communication Technologies*. IEEE, 2013, pp. 408–413.

[29] A. Khosravi, S. K. Garg, and R. Buyya, "Energy and carbon-efficient placement of virtual machines in distributed cloud data centers," in *European Conference on Parallel Processing*. Springer, 2013, pp. 317–328.

[30] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 732–794, 2016.

[31] S. R. Gunn *et al.*, "Support vector machines for classification and regression," *ISIS technical report*, vol. 14, no. 1, pp. 5–16, 1998.

[32] How much carbon dioxide is produced when different fuels are burned? [Online]. Available: https://www.eia.gov/tools/faqs/faq.php?id=73&t=11

[33] Hydropower built back better and greener. [Online]. Available: https://iha-project.webflow.io/

[34] How much co2 does one solar panel create. [Online]. Available: https://www.treehugger.com/how-much-co-does-one-solar-panel-create-4868753

[35] Wind energy's carbon footprint. [Online]. Available: https://www.factcheck.org/2018/03/wind-energys-carbon-footprint/

[36] How can nuclear combat climate change? [Online]. Available: https://www.world-nuclear.org/nuclear-essentials/how-can-nuclear-combat-climate-change.aspx

[37] Questions about radiation. [Online]. Available: https://www.nrc.gov/about-nrc/radiation/related-info/faq.html

[38] L. Čuček, P. S. Varbanov, J. J. Klemeš, and Z. Kravanja, "Total footprints-based multi-criteria optimisation of regional biomass energy supply chains," *Energy*, vol. 44, no. 1, pp. 135–145, 2012.

[39] T. Bhattacharya and X. Qin, "Modeling energy efficiency of future green data centers," in *2020 11th International Green and Sustainable Computing Workshops (IGSC)*. IEEE, 2020, pp. 1–3.

[40] Machine learning basics: Support vector regression. [Online]. Available: https://towardsdatascience.com/machine-learning-basics-support-vector-regression-660306ac5226#:~:text=Overview%20of%20SVR,is%20referred%20to%20as%20hyperplane.

[41] Support vector regression tutorial for machine learning. [Online]. Available: https://www.analyticsvidhya.com/blog/2020/03/support-vector-regression-tutorial-for-machine-learning/

Authorized licensed use limited to: Auburn University. Downloaded on November 17,2026 at 23:18:02 UTC from IEEE Xplore. Restrictions apply.