JOURNAL OF
EMERGING INVESTIGATORS

Article

# Quantifying kitchen appliances' phantom loads using shifted gamma distribution model

**Albert Gan[1], David Klee[2]**
[1] Hopkins School, New Haven, Connecticut
[2] Northeastern University, Boston, Massachusetts

## SUMMARY

**Global energy consumption has gone up and will continue to increase into the future as the population increases and more energy is required to provide resources for the growing population. One way to reduce electric power consumption is to reduce phantom loads, which occur when electronic devices draw energy, even when not in use. This seemingly innocuous source of energy consumption has gained attention for its impact on residential bills and the environment. We hypothesize that kitchen appliances' phantom loads, based on online statistics, account for more than 10% of their total electric power consumption. To test this hypothesis, we propose a shifted gamma distribution model to estimate the phantom loads and apply the model to a public dataset of household electric power consumption. Our results showed 10.9% phantom loads for kitchen appliances supporting our hypothesis. Our findings suggest that implementing strategies to mitigate phantom loads becomes imperative to reduce electricity bills and save energy.**

## INTRODUCTION

Phantom loads refer to the electric power consumption of electronic devices when they are turned off (1). While phantom loads may seem insignificant, according to the National Resources Defense Council more than $19 billion in energy is wasted annually by phantom loads in the United States alone (1). To put this colossal waste into perspective, consider the financial burden placed upon the average North American household: $200 annually, representing at least 10% of the household's electrical bill (2). This often-needless energy waste not only drains financial resources but also exacerbates the environmental crisis–further adding carbon emissions into the atmosphere that are released by mainstream energy generation methods such as burning fossil fuels or natural gases (2). Although transitioning to renewable energy sources has the potential to reduce carbon emissions stemming from energy wasted through phantom loads, it does not address the root of the issue, which is the inherently wasteful nature of phantom loads. Ultimately, unless phantom loads are addressed directly, electronics will continue to draw energy even when not in use.

Dawson *et al.* studied the phantom loads of common appliances such as televisions, microwaves, printers, laptops, speakers, desktop computers, gaming consoles, and coffee makers (2). Their study used two measurement components to collect electricity consumption data from a 50-unit apartment with 100 residents. The first component collected electricity use data of appliances at a 15-minute interval over 24 hours. The second component is a data logger that measured ambient light levels to determine whether kitchen appliances were in normal use by obtaining when the kitchen light was turned on or off. They found that the phantom loads of different appliances account for 4% to 44% of total energy consumption. Laptops have the lowest phantom load, which accounts for 4% of total power consumption. Large televisions have the highest phantom load, which accounts for 44% of the total power consumption. The phantom loads of microwaves account for 18% of total power consumption (2).

In this paper, we hypothesized that the phantom loads of kitchen appliances (e.g., dishwashers, ovens, microwaves) account for more than 10% of total electric power consumption (2). We used a mathematical method to estimate the phantom loads of kitchen appliances from the power consumption data of these appliances that do not contain information regarding the state of the appliance. The household electric power consumption dataset from UCI Machine Learning Repository is suitable for our study because it contains electric power usages of kitchen equipment per minute over a long period (3). This dataset was also used in a study of hybrid voltage control in distribution networks (4). We created a simple statistical model and used the method of moments to estimate the model parameters. The method of moments is a method used to estimate parameters of a probability distribution by equating sample moments with theoretical moments. Our results indicated that the phantom loads of kitchen appliances account for 10.9% of the total energy consumption.

Our results are comparable to the results obtained by Dawson *et al.* (2). However, we did not use special measurement equipment to collect data. Instead, we used a simple statistical model to estimate the phantom loads from data without state information. The results show that our model works well and can produce accurate estimations.

## RESULTS

Here we present the results of applying the method of moments to the electric power consumption data to estimate the phantom loads of kitchen appliances. The kitchen appliances include a dishwasher, an oven, and a microwave, which are assumed to not run continuously.

The statistical model proposed in this study is a shifted gamma distribution with three parameters: the phantom load, the shape parameter, and the scale parameter. We assumed that the total energy consumption $Y$ of the kitchen appliances is the sum of phantom load $p$ and the energy consumption $X$

https://doi.org/10.59720/24-011

of the appliances when they are used to perform their primary functions:

$$Y = p + X \qquad \text{[Eqn 1]}$$

In addition, we assumed that $X$ is a random variable that follows the gamma distribution with two parameters $\alpha$ and $\theta$ because the histogram of the daily data is right-skewed (**Figure 1**). The probability density function of the gamma distribution we used is:

$$f(x) = \frac{x^{\alpha-1}e^{-x/\theta}}{\theta^{\alpha}\Gamma(\alpha)}, x > 0 \qquad \text{[Eqn 2]}$$

Applying the method of moments to the daily data produced the following estimates of the model parameters:

$$\hat{\alpha} = 1.64, \hat{\theta} = 1097.52, \hat{p} = 219.9 \qquad \text{[Eqn 3]}$$

Where the cap notation above the symbol indicates that the parameter is estimated from the data. Since the method of moments is sensitive to outliers, we removed the top 0.5% of data points before applying the estimation method. Our results show that the phantom load of kitchen appliances is 219.9 watts per day. The shifted gamma distribution model fits the data well (**Figure 2**). The daily phantom loads over the 47 months (from December 2006 to November 2010) are near the bottom of the daily power consumption (**Figure 3**).

The average daily electric power consumed by kitchen appliances to perform their primary functions is denoted as:

$$E[X] = \alpha\theta = 1.64 \times 1097.52 = 1794.84 \text{ watts} \qquad \text{[Eqn 4]}$$

The percentage of phantom loads of the kitchen appliances is:

$$R = \frac{p}{p+\alpha\theta} = \frac{219.9}{219.9+1794.84} = 10.9\% \qquad \text{[Eqn 5]}$$

The result support our hypothesis that the phantom loads of kitchen appliances account for more than 10% of the total energy consumption of kitchen equipment. Since the estimates of the parameters contain errors, we applied the bootstrap to estimate the standard error and the confidence interval of the percentage of phantom loads $R$ (5). The standard error of $R$ produced by the bootstrap with 10,000 resamples is 4.24% and the 90% confidence interval is (3.87%, 17.67%). The percentage of phantom loads estimated from the whole data is in the middle of the bootstrap distribution (**Figure 4**).

## DISCUSSION

In this study, we measured the phantom loads of appliances in the kitchen. The dataset we used for this study contained measurements for household appliances such as refrigerators, water heaters, and air conditioners. However, it is difficult to measure phantom loads of refrigerators, water heaters, and air conditioners as these devices run continuously. Compared to those appliances, kitchen appliances do not run continuously. The mathematical model we proposed to measure the phantom loads is a shifted gamma distribution model. The gamma distribution assigns different probabilities to a variable (e.g., daily electric consumption) taking values from zero to infinity. If the appliances have phantom loads, then the daily electric consumption will be, at the least, the phantom loads. In the shifted gamma distribution model, the shift corresponds to the phantom loads.

Our results showed that the simple statistical model produces results that are consistent with prior work. The method of moments worked well to obtain meaningful estimates of the model parameters. According to the model, the percentage of phantom loads of kitchen appliances is around 10.9%. This number is close to that found by Dawson *et al.* (2), who found that microwaves have a phantom load percentage of 18%. However, they employed special measurement equipment to collect the data while we did not use this special equipment. Instead, we used the total electric power consumption data of the kitchen appliances. We derived the phantom loads from the total electric power consumption data by a statistical model.

Our study has some limitations. An implicit assumption of our model is that kitchen appliances are not in use every day. If the kitchen appliances are used every day for a minimum amount of time, then the electric power consumption for the minimum amount of time will be added to the phantom loads. In this case, our model will overestimate the phantom loads. This is a limitation of the data as we cannot tell from the data whether the people are using kitchen appliances. Another limitation of our study is related to the limitation of the method of moments. The method of moments is sensitive to outliers because outliers will lead to overestimated parameters.

Our results show that phantom loads of kitchen appliances account for a significant percentage of daily electric consumption. Understanding the impact of phantom loads is

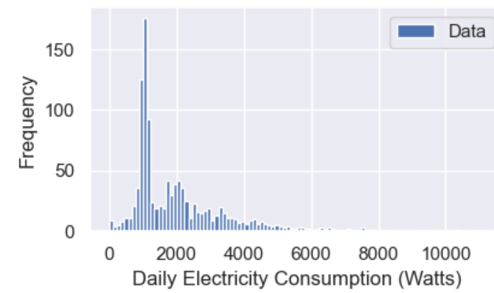


**Figure 1: Daily household electric consumption data from December 2006 to November 2010.** 100 bins containing the frequency of daily energy consumption. The width of each bin is 111.77 watts.

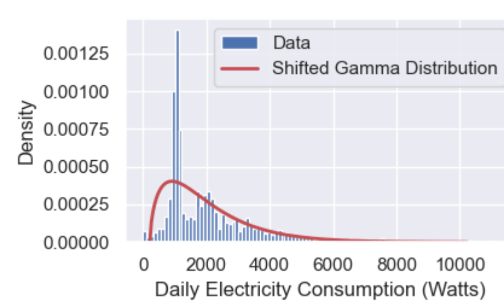


**Figure 2: Daily household electric consumption data with the shifted gamma distribution fitted to the data.** The red curve shows the shifted gamma distribution fitted to the data. The fitted parameters are $\alpha$=1.64, $\theta$=1097.52, $p$=219.9.

https://doi.org/10.59720/24-011

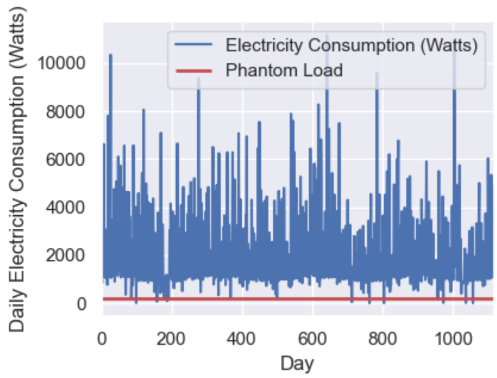


**Figure 3: Daily household kitchen electric power consumption with the estimated phantom loads.** The red line shows the estimated phantom loads of each day.

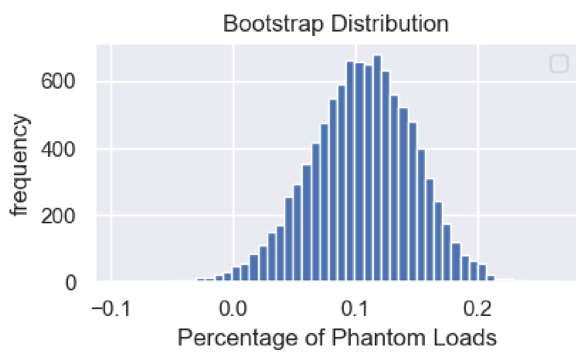


**Figure 4: Bootstrap distribution.** The distribution of the estimates of the percentage of the phantom loads from 10,000 resamples of the data.

crucial in our efforts to optimize energy usage and reduce electricity bills. With this knowledge in mind, implementing strategies to mitigate phantom loads becomes imperative. One effective method is encouraging the habit of unplugging kitchen appliances when they're not actively in use. This simple practice prevents phantom loads as appliances do not have access to power. Additionally, technologies that adapt to power usage to automate the plugging and unplugging of appliances can also help, especially at scale. Ultimately, devices that use energy in more efficient ways will mitigate the effects of phantom load.

## MATERIALS AND METHODS

We used the household electric power consumption dataset from UCI Machine Learning Repository (3) to perform the study. The data set contains 2,075,259 measurements collected in a house over 47 months between December 2006 and November 2010. Measurements were taken every minute in watt-hours. The dataset contains three sub-metering measures, measuring the respective energy drawn corresponding to the kitchen, laundry room, and climate control appliances. The kitchen measurement contains three main appliances: a dishwasher, an oven, and a microwave. In the kitchen where the data was collected, the stove (hot-plates) was gas not electric and therefore the oven data does not include stovetop use. The raw dataset contains the per-minute power consumption of the kitchen appliances (**Table 1**).

| Date | Time | Sub_metering_1 |
|---|---|---|
| 5/1/2007 | 20:09:00 | 1 |
| 5/1/2007 | 20:10:00 | 1 |
| 5/1/2007 | 20:11:00 | 1 |
| 5/1/2007 | 20:12:00 | 1 |
| 5/1/2007 | 20:13:00 | 1 |
| 5/1/2007 | 20:14:00 | 2 |

**Table 1: A subset of the household electric power consumption dataset.** Sub_metering_1 is the watt-hour of electricity consumed in the kitchen, which contains mainly a dishwasher, an oven, and a microwave.

The electric consumption in the kitchen was measured in watts and rounded to integers. As a result, the measurements contain zeros and have rounding errors. Since the measurements were collected every minute, they are not independent as the appliances can run for several minutes continuously. To quantify the phantom loads of the kitchen appliances, we aggregated the raw data into a low-frequency interval so that the aggregate data are approximately independent and contain a small number of zeros. The minute data was aggregated into daily data to achieve this goal (**Table 2**). We removed the first day and the last day of the raw data because the two days do not contain full 24-hour measurements. We also removed the zeros when the electric power consumption on these days was not collected.

| Date | Sub_metering_1 |
|---|---|
| 12/17/2006 | 2033 |
| 12/18/2006 | 1063 |
| 12/19/2006 | 839 |
| 12/21/2006 | 1765 |
| 12/22/2006 | 3151 |
| 12/23/2006 | 2669 |

**Table 2: A subset of the aggregated household electric consumption dataset.** Sub_metering_1 shows the electric consumption in watts of a day.

To quantify the phantom loads of the kitchen appliances, we propose to use the following statistical model: $Y=p+X$, where $Y$ denotes the total electric power consumption by the kitchen appliances in a day, $p$ denotes the phantom loads of the kitchen appliances in a day, and $X$ denotes the daily electric power used by the appliances when they are switched on to perform their primary functions. The histogram of $Y$ shows that the observations are skewed to the right (**Figure 1**). By looking at the histogram, we can assume that $X$ follows a gamma distribution with parameters (6). In this case, $Y$ follows a shifted gamma distribution with three parameters: $p$, $\alpha$, and $\theta$. Since we have the observations of $Y$, we can use the method of moments to estimate the three parameters by matching the first three raw moments (5). Since the method of moments is sensitive to outliers, we need to remove outliers before applying the estimation method (7). Let $y_1$, $y_2$, …, $y_n$ be $n$ observations after the outliers are removed. Then, we estimated the parameters by solving the following three equations:

https://doi.org/10.59720/24-011

$$E[Y] = E[p + X] = p + E[X] = p + \alpha\theta = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \text{[Eqn 6]}$$

$$E[Y^2] = p^2 + 2pE[X] + E[X^2] = p^2 + 2p\alpha\theta + \alpha(\alpha+1)\theta^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2, \quad \text{[Eqn 7]}$$

$$E[Y^3] = p^3 + 3p^2E[X] + 3pE[X^2] + E[X^3] = p^3 + 3p^2\alpha\theta + 3p\alpha(\alpha+1)\theta^2 + \alpha(\alpha+1)(\alpha+2)\theta^3 = \frac{1}{n}\sum_{i=1}^{n} y_i^3, \quad \text{[Eqn 8]}$$

where *E[Y]* and *E[X]* denote the expected values of *Y* and *X*, respectively (8). The above equations can be solved numerically by using fsolve in the Python package SciPy.

## ACKNOWLEDGMENTS

We would like to thank Guojun Gan for the discussion of the statistical models and the use of the method of moments.

**Received:** January 1, 2024
**Accepted:** March 24, 2024
**Published:** August 15, 2024

## REFERENCES

1. Jones, Jennifer. "How Much Does Your Phantom Load Affect Your Electric Bill?" *Citizens Utility Board*. www.citizensutilityboard.org/blog/2019/03/22/how-much-does-your-phantom-load-affect-your-electric-bill/. Accessed 25 Dec. 2023.
2. Dawson, E., et al. "Are Phantom Loads Haunting Your Energy Bill?." *Department of Architecture, University of Oregon.* pages.uoregon.edu/hof/W09HOF/19Phantom_ppr.pdf.
3. Hebrail, Georges and Alice Berard. Individual Household Electric Power Consumption*. UCI Machine Learning Repository*, 2012. https://doi.org/10.24432/C58K54.
4. Liu, H. J., et al. "Hybrid Voltage Control in Distribution Networks Under Limited Communication Rates." *IEEE Transactions on Smart Grid*, vol. 10, no. 3, 2019. https://doi.org/10.1109/TSG.2018.2797692.
5. Dodge, Yadolah. "The Concise Encyclopedia of Statistics." *Springer*, 2010.
6. "Gamma Distribution." *Wikipedia*, *Wikimedia Foundation*. en.wikipedia.org/wiki/Gamma_distribution. Accessed 25 Dec. 2023.
7. Dhruv, and Vasilis Syrgkanis. "Robust Generalized Method of Moments: A Finite Sample Viewpoint." *Advances in Neural Information Processing Systems*, edited by Koyejo, S., et al., vol. 35, Curran Associates, Inc., 2022, pp. 15970–15981.
8. Ross, Sheldom. "A First Course in Probability". *Pearson Education Limited*, 2020.

JOURNAL OF
EMERGING INVESTIGATORS
https://doi.org/10.59720/24-011

## APPENDIX

The Python code used to solve the equations is as follows:

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from scipy.optimize import fsolve
from scipy.stats import gamma
import seaborn as sns
from scipy.stats import bootstrap

data = pd.read_csv(household_power_consumption.txt', sep=";") #Loading data
dat1 = data[['Date', 'Time', 'Sub_metering_1']].replace('?', '0.0')
dat1 = dat1[~dat1['Date'].isin(['16/12/2006', '26/11/2010'])].reset_index()
dat1['Date'] = pd.to_datetime( dat1['Date'], dayfirst=True)
dat1['Sub_metering_1'] = pd.to_numeric( dat1['Sub_metering_1'] )
datd = dat1[['Date', 'Sub_metering_1']].groupby('Date').agg('sum').reset_index()
datd = datd[datd['Sub_metering_1'] > 0].reset_index()

def equations(vars, m1, m2, m3):
 alpha, theta, p = vars
 eq1 = alpha*theta + p - m1
 eq2 = p**2 + 2*p*alpha*theta + alpha*(alpha+1)*theta**2 - m2
 eq3 = p**3 + 3 * p**2*alpha*theta + 3*p*alpha*(alpha+1)*theta**2 + alpha*(alpha+1)*(alpha+2)*theta**3-m3
 return [eq1, eq2, eq3]

def getR(y):
 m1, m2, m3 = np.mean(y), np.square(y).mean(), np.power(y, 3).mean()
 alpha, theta, p = fsolve(equations, (2, 1, 0.1), args=(m1, m2, m3), maxfev=5000)
 return p/(p + alpha * theta)

yd = datd.loc[:,'Sub_metering_1']/1000
p99 = np.percentile(yd, 99.5) #Filtering data
yd = yd[yd < p99]
m1, m2, m3 = np.mean(yd), np.square(yd).mean(), np.power(yd, 3).mean()
alpha, theta, p = fsolve(equations, (2, 1, 0.1), args=(m1, m2, m3), maxfev=5000)

data = (yd,)
res = bootstrap(data, getR, confidence_level=0.9, n_resamples=10000)

sns.set_theme()
fig, axs = plt.subplots()

def plotBootstrapDist():
 axs.hist(res.bootstrap_distribution, bins=50)
 axs.set_title('Bootstrap Distribution')
 axs.set_xlabel('Percentage of Phantom Loads')
 axs.set_ylabel('frequency')

def plotHistogram():
 axs.hist(datd.loc[:,'Sub_metering_1'], bins=100, label='Data')
 axs.set_xlabel('Daily Electricity Consumption (Watts)')
 axs.set_ylabel('Frequency')

def plotHistogramGamma():
 axs.hist(datd.loc[:,'Sub_metering_1'], bins=100, density=True, label='Data')
 x = np.linspace(0, 10000, 200)
 pdf = gamma.pdf(x, alpha, scale=(theta*1000))
 plt.plot(x+p*1000, pdf, linewidth=2, label='Shifted Gamma Distribution', color='r')
 axs.set_xlabel('Daily Electricity Consumption (Watts)')
 axs.set_ylabel('Density')
```

https://doi.org/10.59720/24-011

## APPENDIX Cont'd

```
def plotScatter():
 ax = datd['Sub_metering_1'].plot(label='Electricity Consumption (Watts)')
 ax.hlines(y=p*1000, xmin=1, xmax=datd.shape[0], linewidth=2, color='r', label='Phantom Load')
 ax.set_xlim([0, datd.shape[0]])
 ax.set_xlabel('Day')
 ax.set_ylabel('Daily Electricity Consumption (Watts)')

#plotBootstrapDist(); plotHistogram(); plotHistogramGamma()
plotScatter()
fig.tight_layout(); plt.legend(); plt.show()
```