UC Davis
UC Davis Previously Published Works

Title

Identifying energy savings opportunities in vacant commercial buildings using a semi-supervised sensor fusion model

Permalink

https://escholarship.org/uc/item/27t442nq

Authors

Kornbluth, Kurt
Slaughter, Lisa
Gul, Sadia
et al.

Publication Date

DOI

Peer reviewed

Identifying energy savings opportunities in vacant commercial buildings using a semi-supervised sensor fusion model

Kurt Kornbluth [a], Lisa Slaughter [a], Sadia Gul [a, *], Samanvith Reddy Pamireddy [b], Alan Meier [a]

[a] Energy and Efficiency Institute, University of California Davis, Davis, California, 95616 USA

[b] Mechanical and Aerospace Engineering, College of Engineering, University of California Davis, Davis, California, 95616 USA

[*] Corresponding Author. Email: gul@ucdavis.edu, Energy and Efficiency Institute, University of California Davis 95616 USA

Abstract:

Recent events have forced building managers to examine energy use during vacant periods and revealed miscellaneous electrical loads (MELs) as an opportunity for savings. This paper addresses a key step in unlocking these savings, specifically the reliable identification of when a building is vacant. A Vacancy Inference Engine (VIE), using sensor fusion, was developed to identify vacant periods based on outputs from common sensors, historical building vacancy patterns, and expert knowledge. The VIE calculates the confidence that a building is vacant, allowing building managers to balance the capture of energy savings with the possibility of complaints due to powering down MELs. The VIE has the advantage over logistic regression and other models in that it does not require a full set of ground truth for the training process.

The VIE successfully predicted vacancy in an office building using input data streams of instantaneous electricity demand, indoor carbon dioxide concentrations, and the number of active Wi-Fi connections. The VIE's ability to predict vacancy was compared to that of logistic regression using a metric based on the Complaint Opportunity Rate and found to be nearly identical (0.94 versus 0.95, respectively).

Keywords: Building energy consumption, vacancy modeling, plug loads, miscellaneous loads, green buildings, energy efficiency

1. Introduction

Two intersecting trends in building energy usage have made electricity consumption during periods of no occupancy an attractive opportunity for energy savings. The first trend is the steadily increasing electricity use of l devices that operate even when few or no people are present. These devices include plug loads, network equipment, security systems, elevators, and certain, uncontrolled components in HVAC systems. These miscellaneous electrical loads are responsible for over 30% of electricity use in U.S. commercial buildings and appear to be still growing [1,2]. The second trend is the rising percentage of time when buildings are empty or intermittently occupied. Now, in addition to nights, weekends, and holidays, building managers must "plan" for closures due to smoke and even pandemics. The combination of higher miscellaneous electrical loads used during vacant periods and more vacant hours suggests the existence of potential energy savings.

This paper addresses an essential step in realizing these savings, specifically, reliably identifying when a building is vacant. The paper begins by documenting the two trends and then presents and tests a method for inferring vacancy.

1.1. Miscellaneous Electrical Loads

Miscellaneous electric loads (MELs) are typically defined as any electric load outside of a building's core functions of HVAC, lighting, water heating, and refrigeration [3,4]. This definition by exclusion results in ambiguity as to which end uses are MELs and which are not. Nevertheless, the lack of a consistent definition should not hinder the development of strategies to reduce their energy use.

Miscellaneous electric loads are responsible for a large fraction of electricity consumption in both residential and commercial buildings [5]. In the United States, MELs are responsible for 46% of total building electricity consumption [6]. This category of energy use is disaggregated by individual MEL type in Figure 1, which shows the contributions by various devices in the residential and commercial sectors [7]. The identified contributors illustrate the diversity of loads in the MELs category and the large, "Unspecified" slices hint at an even greater diversity.
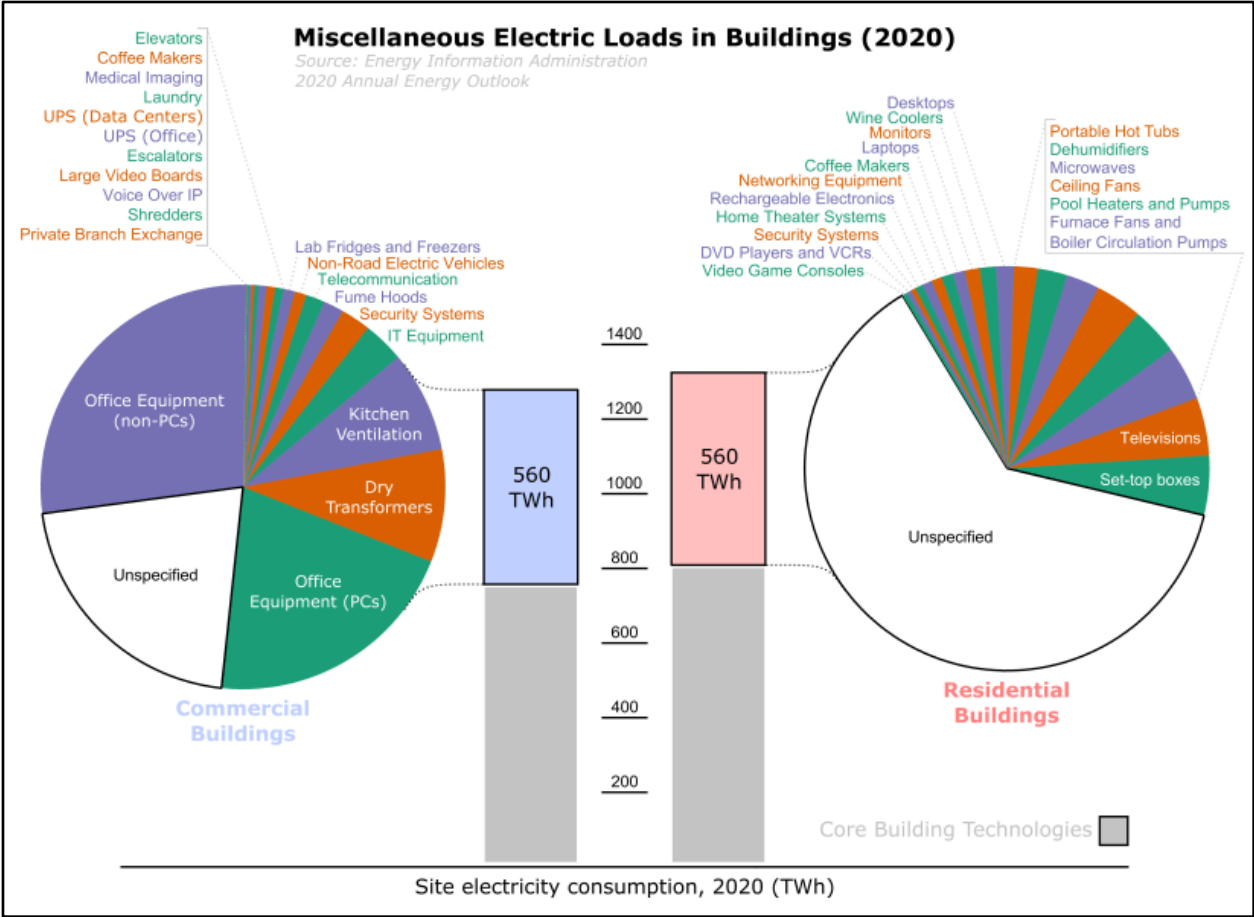
**Figure 1.** Electrical consumption of miscellaneous electric loads (MELs) in US building stock

Investigations of individual buildings support the high MEL consumption estimates shown in Figure 1. For example, a study by Hafer of 220 buildings found that, on average, plug loads were responsible for 32% of the total building electricity consumed [8]. Other studies found similar levels in North America, Malaysia, and Europe [6,9–11].

Many recent studies have examined the correlation between Wi-Fi connections and occupancy [12–15] and energy consumption [16,17]. For example, in a study at a university campus in California, total building electricity consumption was tracked along with Wi-Fi connections as Wi-Fi connections are known to be a good proxy for occupancy. Figure 2 shows Wi-Fi for one week that includes a holiday.

32
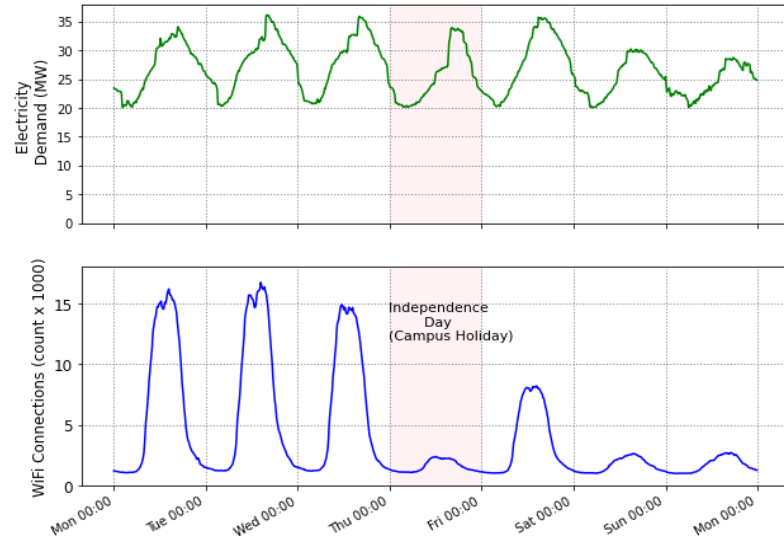
Figure 2. Campus-wide electricity demand and Wi-Fi connections for Monday, July 1 to Monday, July 8, 2019.

Red, shaded regions indicate a holiday.

For the campus, heating and cooling energy is centrally supplied and most of the lighting is centrally or sensor controlled so the electricity consumption shown in Figure 2 is primarily of MELs. It shows that electricity use fell only 24% during the holiday even though occupancy levels were only about 29% of previous days [18]. Many buildings had periods of zero occupancy (inferred by Wi-Fi connection) but also exhibited only modest falls in electricity use [19].

Despite their diversity, most MELs have commonality in that they are small loads (less than a kilowatt), have some standby power consumption, and are connected to communication networks (because so many provide Information Technology services). MELs are also typically not centrally controlled by building management systems. This lack of central control became evident during the pandemic when buildings' electricity use decreased only slightly even though nobody was working in them. A national survey of office buildings found only a 21% reduction in electricity use during the pandemic [20]. In the Empire State Building (New York City), electricity use fell only 28% even though it was almost completely vacant [21].

Miscellaneous electrical loads are projected to grow in the near term. For example, in the United States, MELs are forecasted to increase to nearly 54% of all residential and commercial electricity purchased in 2050 [6].

4

1.2. Vacancy: An Opportunity for Savings

Researchers have extensively investigated the impact of occupancy such as working hours and occupancy schedules on energy consumption [22–24] and their results inform HVAC operation and optimization. On the other hand, few studies have explicitly considered periods of vacancy and its impact on energy use in the buildings [25–28].[1] Some of the detailed studies of occupancy indirectly reveal vacancy information when they report zero Wi-Fi counts or no activity from occupancy sensors. Although a number of studies have documented occupancy patterns in diverse commercial buildings revealing significant periods of zero occupancy (weekends and holiday periods were ignored), the number of vacant hours and duration were not treated with similar detail [15,26,29–36]. Nonetheless, commercial buildings are vacant a significant number of hours each year and vacant long enough to justify special energy-saving strategies such as in the present study.

Most commercial buildings are vacant for at least a few hours per year, but the number of hours depends on building functions, size, cleaning schedules, security procedures, culture, and other factors. Few measurements have been made on the frequency and duration of vacant hours. Most information is anecdotal; for example, every weekday security guards lock university buildings at 23:00 and custodians re-open them at 07:00 in the morning [26]. This simple schedule implies the building is vacant for 33% of the year. Operating procedures for weekends, holidays, and special situations (such as a pandemic) are not always clearly documented but only increase the proportion of vacant hours. On the other hand, certain staff may still be able to enter the building during the locked periods, which will reduce the vacant hours. A study of 24 buildings on one university campus estimated that the buildings were vacant 29% of the year [19].

During the COVID-19 pandemic, campus buildings experienced a 90% reduction in occupancy, but electricity use fell only 15% [37]. In the past, a study of six commercial buildings showed that during unoccupied times of the weekend the buildings were consuming on average 23% of the building's weekly energy [38]. Situations like those described above suggest that reducing electricity consumption when buildings are vacant or sparsely occupied is an untapped opportunity with significant potential for electricity savings [39,40].


[1] We call buildings "vacant" when they are unoccupied for brief periods – minutes, hours, days. This differs from the insurance definition, which considers "vacant" to refer to a building that contains no personal or corporate property.

1.3. Overview of Vacancy Inference Approach

To unlock the savings described above, vacancy must be determined with high confidence. Completely unoccupied buildings – even if only for a few hours – are the most attractive target for energy savings because nobody is inconvenienced by reduced services. However, truly aggressive control strategies cannot be implemented unless there is high confidence that nobody is in the building. Thus, a "Vacancy Inference Engine" is proposed to combine data sources to identify vacant periods and then signal this condition to MELs. The logical flow of the Vacancy Inference Engine is shown in Figure 3.
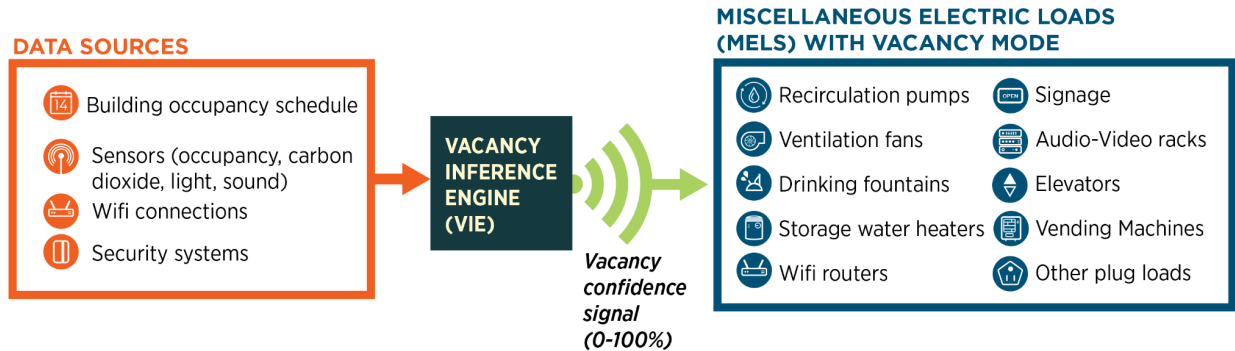
Figure 3. Proposed method to reduce MELs during periods of building vacancy.

Ultimately, the MELs need to be modified to receive a signal and enter a "vacancy mode" when the situation is appropriate. Some devices are already equipped to receive a signal in order to respond to real-time electricity prices and to provide other grid services [41]. This paper focuses on the VIE and its ability to infer vacancy from common sensors and then applies it to a case study.

1.3.1. Occupancy Sensing vs Vacancy Inference

The initial targets for energy savings are MELs operating during periods while a building is vacant. For this reason, a determination of vacancy is required before these devices can be powered down. The state of being vacant exhibits features that make it very different from the occupied state. For instance, the occupancy information can be categorized into presence, location, count, activity, direction, and identity of the person based on their spatial and temporal properties [42]. Moreover, the training data for vacancy inference is likely to be more similar across a wide range of

93 buildings than occupancy data for two reasons. First, the vacant condition is essentially identical across all buildings,

94 that is, zero people are in the building. Second, near-vacant conditions, that is, when only a few occupants are present,

95 are also likely to be similar across many types of buildings. For example, near-vacant conditions correspond to when

96 a custodian enters to clean the building, a security guard makes the rounds, or a person enters to fetch an item during

97 off-hours. On the other hand, occupancy-based models are building specific, and the occupancy profile depends on

98 the building type and its occupant characteristics [43].

99 No sensors exist to detect vacancy; nevertheless, the absence of signals from occupancy detectors is a crude inference

100 of vacancy. The inference of vacancy is uncertain for three principal reasons. First, all sensors are susceptible to noise.

101 The noise can lead to a sensor falsely interpreting a vacant condition as occupied. The opposite type of error is also

102 possible where noise filtering is applied; there, the signal generated by occupant presence is filtered out and ignored.

103 Some sensors, such as a PIR (passive infrared) sensor, fail to detect stationary persons, thereby reporting an occupied

104 space as vacant. The third source of uncertainty arises because occupancy detectors are typically designed to maintain

105 an 'on' output for a fixed period after ceasing to detect a presence. This delay translates into longer periods of apparent

106 occupancy than actually occurred and a bias towards shorter durations of vacancy.

107 Modeling occupancy and vacancy differs because the presence of persons can be detected whereas their absence can

108 only be inferred. When testing for occupancy, a binary output would be described as [44]

109   1: *Occupied*

110   0: *Not occupied*

111 whereas, in the case of vacancy inference, this would become

112   1: *Vacant*

113   0: *Not vacant*

114 Here, the default case is that one or more persons are present, defined as "0", and changes to "1" only when reasonable

115 doubt exists. This reasonable doubt is quantifiable as a confidence level that is not solely a function of sensor accuracy,

116 exposing the need for an output that describes this confidence in the form of a percentage.

1.3.2. A Sensor Fusion Approach

Conventional sensors in buildings detect occupancy through a variety of physical means, such as a movement, a sound, a change of temperature, an increase in carbon dioxide, camera footage, and Wi-Fi connections [45]. Single sensors such as Wi-Fi however, have been proven to be unreliable [30], especially at very low occupancy levels [35]. A "sensor fusion" method can be employed to minimize these shortcomings and improve the accuracy of the result [46,47]. In this approach the information from various sensors is combined to infer vacancy, resulting in a lower overall error than a single sensor would achieve alone. This assumes the sensor outputs are independent [30] and results in a higher signal-to-noise ratio. Sensor fusion results in a single, actionable number that can be used in subsequent operations [48].

Occupancy sensing algorithms have been generally categorized into several methods, including statistical and classical learning, kernel-based, data mining and clustering, probabilistic graphical, and neural networks [49]. These methods require building-specific ground truth occupancy information for training data, which is time consuming and expensive [50]. For example, for obtaining the ground truth, all the building entries and exits should be recorded, and it can be done using camera logs, keycard events, sensor technologies or human observations.

The proposed model, referred to as the Vacancy Inference Engine (VIE), is trained using building vacancy patterns obtained from building experts, occupants, and general knowledge. Only a few studies have inferred occupancy probabilities without using any ground truth data; even then, their focus is not on vacancy [51,52] and moreover, assessment of various probabilistic models also highlight the fact that these methods lack the ability to predict only vacancy timeslots [43]. While on the other hand the method proposed in this study eliminates the need for ground truth and uses only the vacancy patterns to label the training data. The training data set is further simplified in that it only contains information about the vacant state and does not require knowledge of the occupied state, as is needed for other approaches. Since data used to train the VIE does not contain both possible states, we classify the model in the present study as a semi-supervised sensor fusion method.

The output of the model is a number between 0 and 1 which represents confidence in the determination of vacancy. The signal communicating this confidence is expected to be transmitted to and monitored by end-use devices that process it and trigger themselves into low- or no-power modes, hereafter referred to as vacancy modes. Since the

143 output is a continuous variable rather than a binary output, it allows for the consideration of different risk profiles

144 across devices. For example, the consequences of mistakenly powering down an elevator while someone is inside are

145 more severe than the consequences of shutting down a water cooler while someone is thirsty. Thus, a higher confidence

146 of vacancy would be required for placing the elevator into vacancy mode versus the water cooler.

2. Methodology: Model Development

148 The creation of the vacancy inference model and its usage occurs in three phases: model training, vacancy inference,

149 and device action. The process flow for these phases is depicted in Figure 4. The parameters determined during the

150 one-time model training phase (μ and s) are used during the inference phase to convert raw sensor input (x) into

151 intermediate confidence of vacancy (c). These intermediate predictions are fused at a single node in the inference

152 phase to provide overall confidence of vacancy for the target area (C). During operation, an end-use device converts

153 the overall confidence of vacancy into a binary output *(O)* by comparing to its specific threshold value to initiate

154 action through its controller.
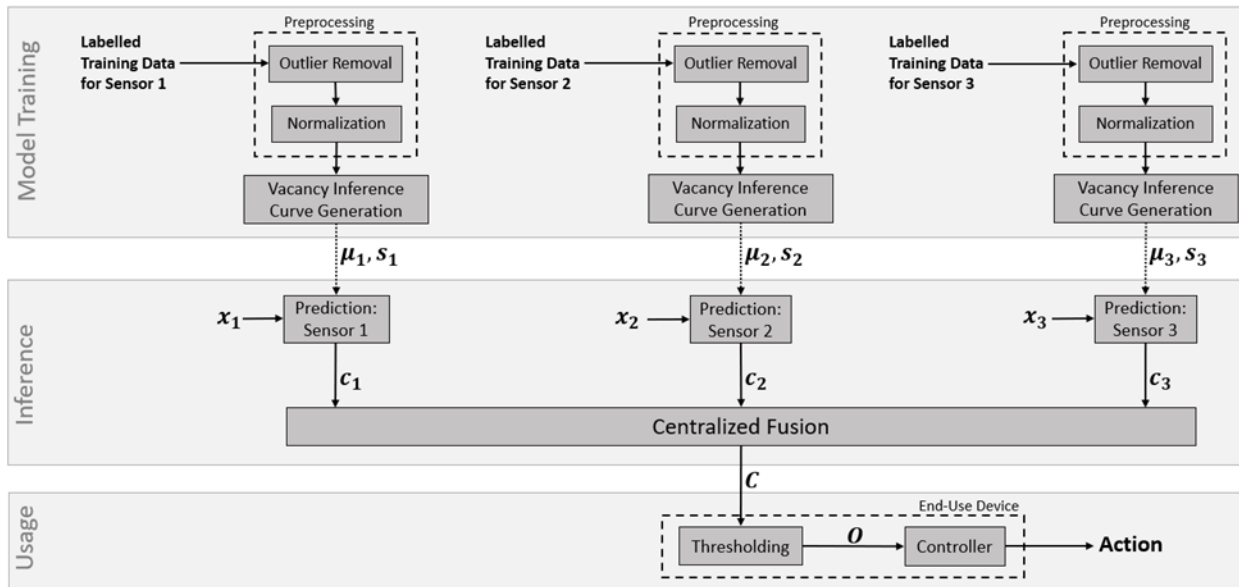
Figure 4. Logic flow for VIE during its development and usage phases

157 The following assumptions were adopted to formulate the VIE and to process data produced by the individual sensors:

158 • Increased occupancy will result in an increase in the sensor value.

- Once a sensor value rises significantly beyond those typical of vacancy, it can reasonably be assumed to indicate occupancy. This allows for the rejection of data from known occupied periods during the training process.
- Sensor inputs are independent of each other.
- The relationship between sensor outputs and confidence of vacancy can be approximated by a logistic probability distribution function

2.1. Model Training Phase

The VIE must first be trained to recognize vacant periods. This is accomplished by feeding it sensor values from periods where the space is known to be vacant. In this example, the sensor values were manually inspected, so confidence of vacancy was extremely high (though not 100%). This means the data can be used as training data with a label of 1 which represents 100% confidence of vacancy. The model training phase involves two main steps, preprocessing and vacancy curve generation.

The input data must be preprocessed and made free of outliers and anomalous data. For example, faulty sensors must not pass negative or zero values to the VIE. Instances of outliers could occur because of malfunctioning sensors, the systems that support them, undesired noise, or real-world occurrences. The values outside of three standard deviations from the mean of the training data are rejected from consideration. Note that using standard deviation assumes the normal distribution, but this study assumes the input values to be distributed according to logistic distribution. This disparity is expected to have little effect on the results as the two distributions are very similar to each other.

Normalization of training data removes bias and scales all values to between 0 and 1. The normalized value of every data point in the training set is given by:

$$x' = \frac{x - min(X)}{max(X) - min(X)} \ldots .. \text{ Eq } 1$$

where X is the set of all x in the training set and x' is the normalized value.

The next step, the generation of vacancy inference curves, is achieved by performing a curve fit to empirical data that is referred to as the training data. When training a supervised model such as logistic regression, it is required that

183   information about all possible outcomes be included in the training set. In contrast, the set of training data used in

184   this treatment is simplified to contain only information about conditions during the "vacant" state; it does not include

185   any data taken during the "not vacant" state. This selective generation of training data causes this treatment to fall

186   under the umbrella of semi-supervised training. This approach is made possible by two features of the problem being

187   solved. Firstly, this treatment is uninterested in determining when a building is occupied, only when it is vacant.

188   Second, input data values remain relatively static during vacancy and any significant upward deviation indicates

189   presence.

190   The data during times of vacancy are distributed according to a logistic probability distribution function (PDF), and

191   thus a representation of the underlying logistic cumulative distribution function (CDF) can be coaxed out of the data.

192   For convenience, we will define "vacant" = 1, "not vacant" = 0, so lower sensor values correspond to higher

193   probabilities of vacancy. Subtracting the logistic CDF from 1 yields Eq 2:

194
$$CDF(x_i) = 1 - \frac{1}{1+e^{-\frac{x_i-\mu}{s}}} \qquad \text{.... Eq 2}$$

195   where μ is the location parameter which is equivalent to the mean of the training set, s is the scale parameter which is

196   proportional to the standard deviation of the training set, and x_i is the raw sensor input value.

197   To determine the parameters μ and s, the CDF shown in Eq 2 is fit to an empirical CDF built from the set of normalized

198   training data, X'. When normalized such that its maximum value is 1 and it's minimum is zero, the output of the CDF

199   represents the confidence of vacancy[2] indicated by a given x_i:

200
$$c(x_i) = CDF(x_i) \qquad \text{.... Eq 3}$$

201   An empirically derived vacancy inference curve for each sensor input was generated from the cleaned input data as

202   follows:

203   1.   Calculate the mean μ_0 and standard deviation s_0 of the unnormalized training data. Retain these for later use,

204   and sort the normalized training data in ascending order according to its value, ignoring any timestamp

205   information.

---

[2]The output of Eq 3 represents a true probability when both vacant and occupied states are considered. In our case, the VIE output does not represent a true probability because data from occupied periods have been eliminated. More precisely, the output represents a percentage of occurrences.

206 2. Create a cumulative distribution by calculating the definite integral of the set generated in Step 1.

207 Computationally, this is equivalent to calculating the sum of all values in the set that are less than or equal to

208 x_i', at each x_i'. This is represented mathematically by:

209 $$z_i = \sum_{i=1}^{i} x_i' \qquad .... \text{ Eq 4}$$

210 where x_i' is the ith value from the sorted, normalized training data set.

211 3. Generate the CDF for the vacant state by normalizing the results of Eq 4 according to Eq 5.

212 $$c_i = \text{CDF}(z_i) = 1 - \frac{z_i}{max(z)} \quad .... \text{ Eq 5}$$

213 where z is the data set generated by Step 2. This re-normalizes the data and flips it across the vertical axis.

214 The result is an empirically obtained vacancy inference curve, which can be approximated by the CDF in Eq

215 2.

216 4. Create a vacancy inference curve that maps raw sensor values to a confidence of vacancy by fitting the CDF

217 from Eq 2 to the empirical vacancy inference curve and extract for later use the values that correspond to

218 μ and s in Eq 2. In performing this fit, use the sorted unnormalized training data for the x-axis values, use

219 the corresponding inference curve values from Step 3 for the y-axis values, and use the mean μ_0 and standard

220 deviation s_0 obtained in Step 1 as the initial guesses of μ and s, respectively. Once the fitted values of μ and

221 s are known, they can be plugged into Eq 2 and the result used to infer the confidence of vacancy from new

222 data received from the sensors during the operational phase.

223 The steps above are performed to create a vacancy inference curve for each sensor. A minimum two-week data set

224 was used in order to capture the weekly occupancy cycle.

225 The three vacancy inference curves created by this study are shown in Figure 5 for the sensor inputs of indoor carbon

226 dioxide concentration, electricity demand, and the number of active Wi-Fi connections.
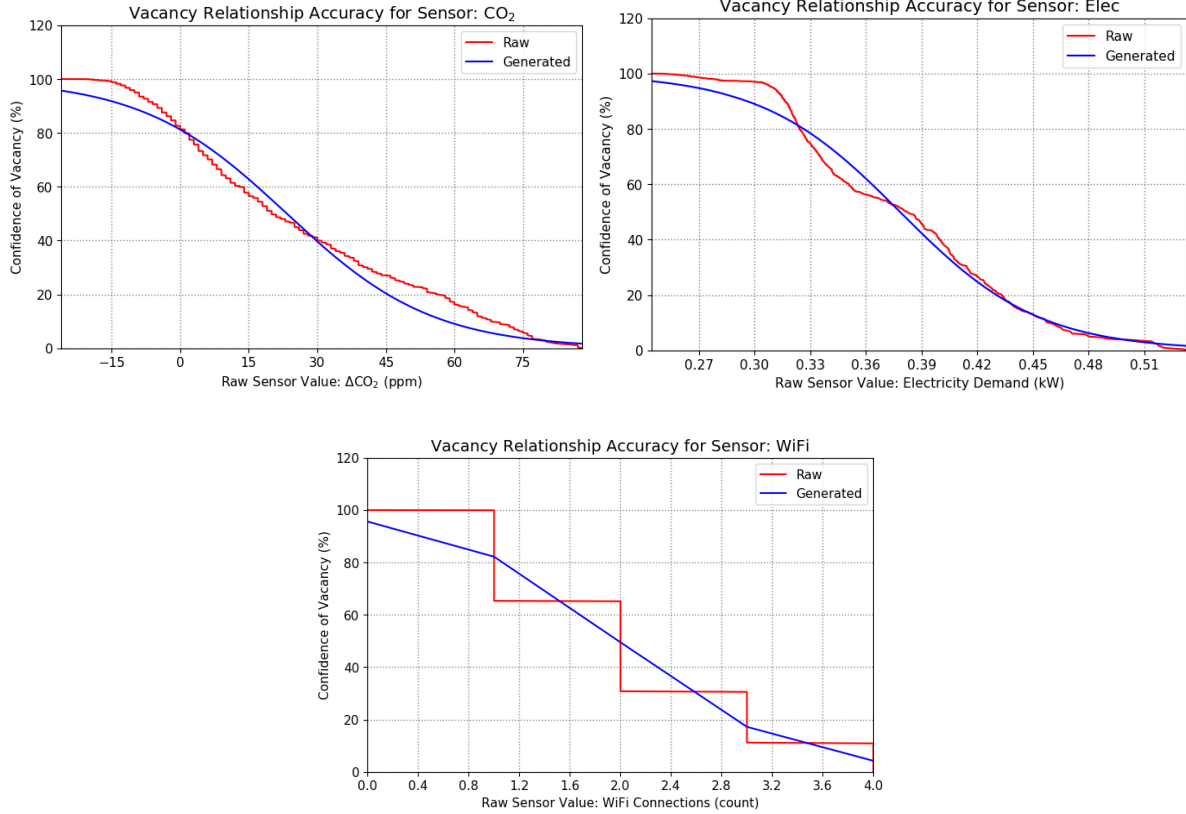
Figure 5. Vacancy inference curves for CO_2 (top left), electricity demand (top right), and active Wi-Fi connections (bottom). Red lines indicate actual data with best-fit curves indicated in blue.

The best fit lines, shown in blue by Figure 5, take the form of the modified logistic cumulative distribution function from Eq 2. These vacancy inference curves describe, for each sensor, a mapping between the raw input data and the confidence of vacancy that will be generated by the model.

2.2. Model Operation

Raw sensor input (x) is fed into the functional form of the modified logistic CDF in Eq 2, using values of μ and s as determined during the model training phase. The output of this is the confidence of vacancy as predicted from each individual sensor stream, c_j. This is referred to as the intermediate confidence of vacancy for each sensor stream. The set of c_j for all sensor streams is denoted by vector $\vec{c}$.

240 Outlier detection is not performed on incoming data because outliers do not adversely affect the model during the

241 inference phase. This is convenient because the true mean and standard deviation of incoming data is unknown for

242 each sensor stream. Outliers could still result from malfunctioning sensors, calibration errors, or sensor failure. Often,

243 sensors are set to send unreasonably high values when they fail so that their malfunction becomes apparent to

244 operators. Since a vacant state is defined only by the lowest readings for each sensor, the outliers in the upper boundary

245 of the input are not influential to confidence estimation.

246 If a failing sensor erroneously produces a negative value or a value of zero, the model will output a nearly 100%

247 confidence of vacancy for that data stream. If erroneous data is input to the VIE, the sensor fusion step will mitigate

248 the effect of the data stream on the model output. Since this will affect model performance, building data systems need

249 to be designed to differentiate real data from that produced by faulty sensors.

2.2.1. Fusing Data Streams

251 During the centralized fusion step, the intermediate confidences of vacancy, $\vec{c}$, obtained from the various sensor input

252 streams are fused to produce a single, overall confidence of vacancy, C:

$$C = f(\vec{c}) \qquad \text{.... Eq 6}$$

254 To determine the form of the fusion function $f(\vec{c})$ eight candidate functions for sensor fusion were studied for the

255 given data. These are the maximum, product, root sum square, arithmetic mean, standard deviation weighted mean,

256 geometric mean, harmonic mean, and root mean square. The overall confidence of vacancy obtained from this

257 decision-level fusion was compared against the actual vacancy pattern of the office space and evaluated in terms of

258 the potential for complaint versus missed opportunities for energy savings. This work is detailed by a separate study

259 which concluded that Root Mean Square, shown by Eq 7, performed superior to the other tested methods in

260 determining the overall probability of vacancy [18].

$$f(\vec{c}) = \left(\frac{1}{n} \sum_{j=1}^{n} c_j^2 \right)^{\frac{1}{2}} \qquad \text{.... Eq 7}$$

262 A notable benefit of fusing data in this manner (as opposed to regression-type approaches) is that it allows the

263 intermediate confidence from any number of sensors to be used at any given moment in time without the need to

264 retrain a fusion model. This makes the fusion step more robust and resilient to sensor drop-out.

2.3. Device Action: Conversion of the VIE Output to Device Control

End-use devices are expected to utilize the VIE output to decide if they should enter vacancy mode. Since the VIE output is a continuous number, it must be transformed into a binary value of vacant (1) or not vacant (0) by the device's controller. This is performed by thresholding as shown in the Device Action stage of Figure 4.

Each piece of equipment is administered a defined threshold based on its risk profile. This threshold is compared with the confidence of vacancy from the VIE. If the confidence is found to be greater than the value of the threshold, the device considers it vacant (1) and enters vacancy mode. Conversely, if the confidence is found to be less than or equal to the value of the threshold, the device sees the space as occupied (0) and remains fully powered.

To choose the threshold, a building manager must first understand the tradeoffs between risk and benefit that will manifest at each possible setting between 0 and 1. When comparing the VIE output to actual truth, there are four possible outcomes, as summarized in the confusion matrix shown in Figure 6.

		Ground Truth	
		Vacant (1)	Occupied (0)
Predicted	Vacant (1)	True Positive-TP	False Positive-FP (Complaint Opportunity)
	Occupied (0)	False Negative-FN (Missed Opportunity)	True Negative-TN

Figure 6. Confusion matrix comparing VIE output to ground truth

The usual concept of accuracy is characterized by Vacancy Inference Accuracy (VIA), which describes the rate at which vacancy is correctly inferred by the model. VIA is defined as the number of true positives divided by the number of occurrences of actual vacancy, as shown by Eq 8:

$$VIA = \frac{TP}{TP+FN} \quad .... \text{ Eq 8}$$

Missed opportunities for energy savings occur when the model wrongly infers presence although the space is actually vacant. In this case, action is not taken based on the VIE output and because of it, energy is used to provide services

284 that no-one is around to enjoy. These are defined as Missed Opportunity Rate (MOR) described mathematically by

285 Eq 9:

286 $$MOR = \frac{FN}{TP+FN} \qquad .... \text{Eq } 9$$

287 Opportunities for complaint arise in the reversed case. When these occur, taking action based on the VIE output

288 switches off devices when people are actually present, causing potential problems for the occupants. The Complaint

289 Opportunity Rate (COR) is expressed as

290 $$COR = \frac{FP}{FP+TN} \qquad .... \text{Eq } 10$$

291 The consequences of these errors vary by end-device and application and can range from mild inconvenience to

292 possible harm to the undetected occupants. A building manager will seek to minimize disruption to occupants while

293 capturing energy savings and thus will choose a vacancy confidence threshold that tailors this trade-off to anticipated

294 occupant tolerances.

3. Model Application: Inferring Vacancy in Real Buildings

296 The vacancy inference method proposed in this study is tested at an office in Davis, California, USA, the floorplan of

297 which is illustrated in Figure 8.

3.1. Input Data Selection

299 The proposed method can accept a wide variety of inputs to infer vacancy, but not all candidates are suitable for the

300 task. Some inputs contain little information about the state of vacancy and, if used, will effectively inject noise into

301 the system and drastically reduce the VIA of the model. Many supervised methods exist for quantifying the correlation

302 of input to vacancy. Unfortunately, these tools cannot be used for this treatment because it explicitly avoids ground

303 truth requirements in the problem definition. Building managers who seek to apply the proposed vacancy inference

304 method are not expected to collect ground truth data. Thus, qualitative methods must be used to determine input

305 suitability. The most accessible way to establish a relation between the sensor inputs and vacancy is to observe the

306 raw sensor data in accordance with the expected occupancy patterns of the building. It is assumed that the value of the

307 sensor input will generally increase with an increase in occupancy.
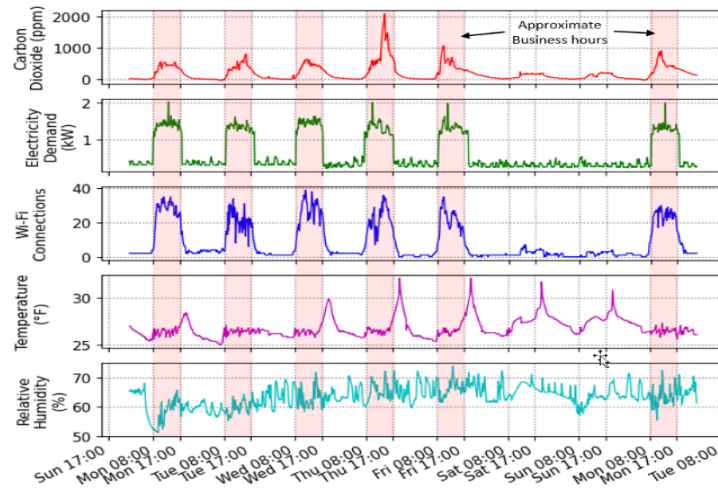
308

309 **Figure 7.** Correlation between sensor inputs and estimated occupancy patterns for Monday, July 8 through Monday,

310 July 15, 2019. Transparent red regions indicate approximate working hours

311

312 For the given office space, Figure 7 shows the variation of sensor inputs compared to expected occupancy patterns.

313 The approximate business hours are indicated by the transparent red regions. It is clear that carbon dioxide, the number

314 of Wi-Fi connections, and electricity demand track roughly to the operational hours of the building. It is to be noted

315 that the electricity demand shown in Figure 7 is independent of the demand associated with the HVAC system. This

316 is done to eliminate HVAC control affecting energy demand, isolating electricity consumed by the end-use devices

317 that are indicative of an occupied or vacant state. Relative humidity and temperature are expected to increase with

318 occupancy, but for the given office space this does not occur to a recognizable degree. Thus, carbon dioxide, electricity

319 demand, and the number of active Wi-Fi connections to the office's access points are deemed suitable inputs. The

320 binary signals (key card entry, etc.) were not included in the analysis to identify vacant periods because the building

321 did not have any accessible binary data streams

3.2. Testing Scenario

323 The proposed vacancy detection method was tested in an office/lab area in Davis, California, USA, comprising the

324 entire first floor of a building on the UC Davis campus. This floor consists of two research offices and the remaining

325 upper floors consist of residential apartments for university students. The research offices in the study and residential

326 areas are physically independent spaces. No movement of people is possible between the two spaces without first

327 exiting the building, nor is HVAC/air shared between them. Relevant input data from pre-existing sensors was isolated

328 to the area of interest to the researcher's best ability.
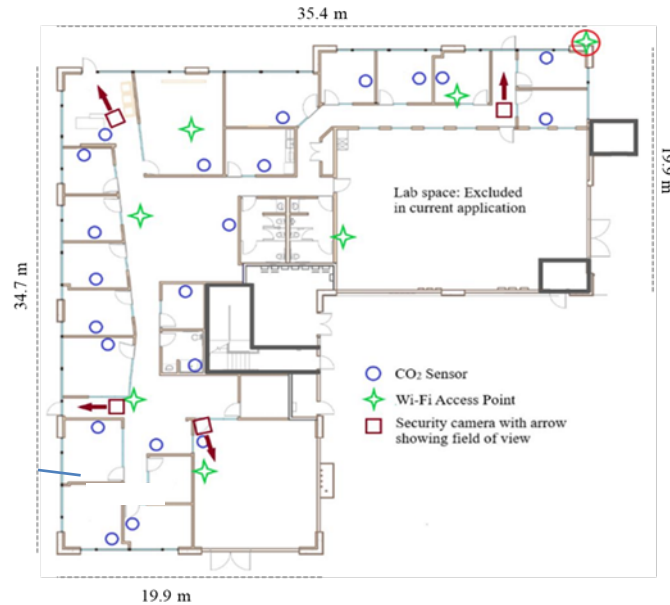
329

330 **Figure 8.** Layout of test office space with CO_2 sensors, Wi-Fi access point, and security camera

331 Carbon dioxide measurements were obtained using a pre-existing system of 22 CO_2 concentration sensors that gather

332 ambient environmental data in 5-minute intervals. In the rare instance that a sensor does not have a value reported in

333 the 5-minute window, the sensor's previous value is used to ensure a value exists at every 5-minute interval. To

334 combine the reading of all 22 sensors in the office space, the maximum of all sensor data is used. This preserves

335 smaller signals that would otherwise be suppressed by an overall average, thus allowing for a single sensor detecting

336 a single occupant in an otherwise vacant space. The outdoor carbon dioxide readings are then subtracted from this

337 maximum value so that room CO_2 levels are expressed as a deviation from outside concentrations. The outdoor CO_2

338 concentration was obtained from a sensor at a low-congestion area of the UC Davis campus. This sensor was found

339 to be a good indicator of outdoor levels, with no sudden fluctuations due to cars or people.

340 Concentration of carbon dioxide in buildings typically depends on the air exchange rates. Fortunately, the location

341 targeted by this study has no operable windows and the sensor units are not located next to any doorways. The

342 ventilation for the area is controlled by schedule according to business hours but is also able to be controlled manually

343   by occupants. CO_2 concentration measurements are also prone to temporal lag subject to placement of the sensors and

344   environmental rates of effusion. This varies by scenario and is dependent on multiple factors such as airflow rates, air

345   density, carbon dioxide concentration, and distance between the emitter and the sensor. Since the lag is not consistent

346   across buildings, correction for it was not attempted.

347   Electrical current readings are recorded at the control panel. These readings are converted from analog to digital and

348   routed to data logger modules. These modules sample readings every minute, uploading them to a cloud.

349   The Wi-Fi access points of the UC Davis campus-wide networks are grouped according to the building they are

350   intended to serve and are programmed to report the number of active connections to a central server maintained by

351   the campus. The total number of reported Wi-Fi connections data is updated and archived every 10 minutes.

352   Since the time interval for each sensor stream differed, the data for carbon dioxide and electricity demand were down

353   sampled such that their intervals matched that of the Wi-Fi sensor stream, resulting in an overall data collection interval

354   of 10 minutes. This interval is typical of building data collection systems.

355   A separate investigation was undertaken to determine the reliability of the Wi-Fi counts. Staff and students typically

356   configure their phones and laptops to automatically connect to the network, but the same cannot be assumed for

357   visitors' devices. To confirm these expectations, interviews were performed to determine if employees' mobile devices

358   are set to automatically connect to the network. This was found to be the case for all respondents. This survey also

359   revealed that the cell signal is weak in the target area, which would further encourage others to use the local Wi-Fi

360   network. However, it was found that visitors typically do not configure their devices to connect to the network and

361   are likely not reflected in the connection count. The custodial staff is also not typically captured. This decreases the

362   VIA for Wi-Fi alone as it fails to detect these groups of people. It is also possible that a single person accounts for

363   multiple observed Wi-Fi connections due to ownership of more than one smart device configured to automatically

364   connect to either network. Connections unrelated to occupancy might also originate from Wi-Fi printers, smart

365   televisions, and other "Internet of Things" devices. This investigation demonstrates why Wi-Fi connections alone

366   cannot be used to infer vacancy.

367   **Table 1.** Sensor inputs with data acquisition for the office space

Sensor Input	Sensor	Count	Data Acquisition Interval
Indoor Carbon Dioxide Concentration (ppm)	T6713 Miniature CO_2 Sensor Module	22	Each sensor reports its data asynchronously at 1-minute intervals, so data is grouped into 5-minute intervals for each sensor and the value closest to the 5-minute timestamp was extracted
Instantaneous Electricity Demand (kW)	Accu-CT Split Core Current Transformer	1	Continuous analog readings are converted to digital and routed to Wi-Fi data loggers, which sample readings every 1-minute
Number of Wi-Fi connections (Number)	Wi-Fi Network Access Point	7	Number of active connections every 10th minute - Instantaneously
Doorway Video (Ingress/Egress Count)	Video Camera	4	Update on change of value

368 As described in section 3.1, Wi-Fi, electricity demand, and carbon dioxide levels were selected for the study because

369 they are well-correlated to vacancy. Example data outputs from each of these sources are shown in Figure 7, where

370 they are compared with vacancy patterns over one week.

371 3.3. Ground Truth

372 The performance of the model for the given building was evaluated by comparing its predictions against the true state

373 of vacancy, known as ground truth. The ground truth for this investigation was extracted by manually reviewing the

374 recorded motion-sensing based videos from security cameras that monitor each exterior doorway of the target area.

375 The video clips were first cropped by using OpenCV [53] so that only the doorway was captured, and then clips

376 without people were deleted using a machine vision package called ImageAI [54] . The remaining video clips were

377 reviewed manually by the researcher, ensuring that the number of entries cancels the number of exits so that the final

378 occupancy at the end of each day is zero but 2 out of the 35 days ended at some value other than zero occupants. The

379   unprocessed footage was inspected, and it was found that the cameras' motion detection software missed those entry

380   or exit events. Thus, a contrived entry was added to the ground truth to ensure it matched reality as known from

381   historical vacancy patterns and occupant interviews.

382   These difficulties and uncertainties in obtaining the perfect ground truth further illustrate the need for a semi-

383   supervised method of inferring vacancy. It should be emphasized that the ground truth collected from security cameras

384   was used only for the evaluation of model performance; it was not used during model training.

385   4. Results and Discussion:

386   The confidence of vacancy obtained by the VIE for the test data was subjected to varying decision thresholds and

387   compared to the ground truth to obtain performance metrics. These metrics were also obtained using logistic regression

388   for vacancy inference and compared to those obtained for the proposed method to arrive at the following results. In

389   the following results, the VIE output (range: 0-1) is expressed as percentages (range 0-100%).

390   4.1. Receiver Operating Characteristic Curve

391   The Receiver Operating Characteristic (ROC) curve is a diagnostic that compares the outcome being maximized, VIA,

392   to the outcome being minimized, COR. The curve is built by varying the decision threshold across its full range and

393   plotting the resulting VIA on the y-axis and the COR on the x-axis.

394   The dashed line shown in Figure 9 represents the ROC curve of a perfectly predicting model i.e., ground truth, which

395   would yield a VIA of 1 at every threshold other than one and a COR of 0 at every threshold other than zero. The ROC

396   for the present treatment closely follows that of traditional logistic regression with slightly lower vacancy inference

397   accuracy than that of logistic regression for the same thresholds. A model that exhibits an area under this curve closer

398   to one is more desirable and the areas under the ROC curves for the proposed method and logistic regression are very

399   closely matched (0.94 vs. 0.95, respectively).
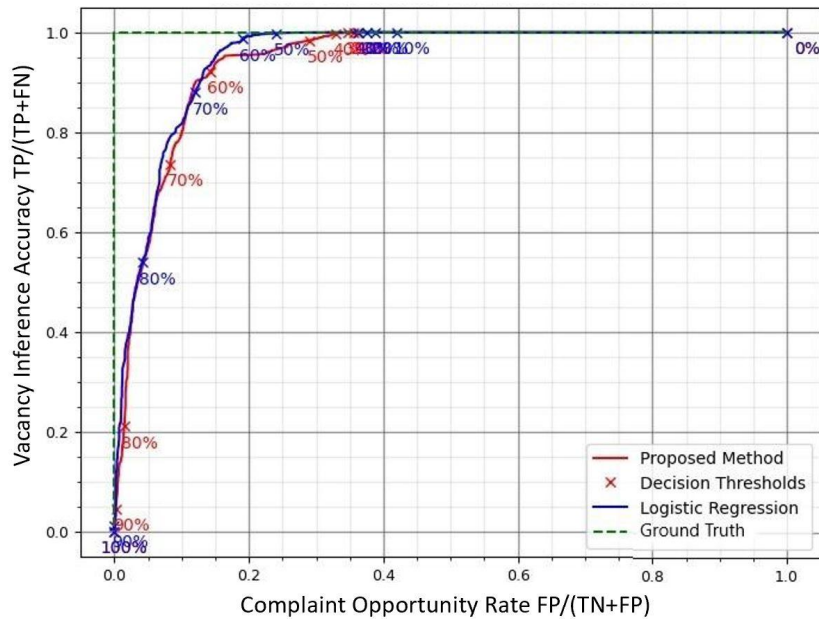
Figure 9. Receiver Operating Characteristic curve

4.2. COR-MOR Characteristic Curve

The COR-MOR Characteristic (CMC) curve is a diagnostic that plots the complaint opportunity rate (COR) against

the missed opportunity rate (MOR) at different decision thresholds, in a fashion similar to the ROC curve. The dashed

line in Figure 10 is indicative of a CMC curve for a perfectly predicting model, resulting in a COR of 0 at every

threshold other than zero and a MOR of 0 at every threshold other than one. Minimizing COR and MOR throughout

the range of all thresholds is highly preferred for an optimally working vacancy inference system, and a model with

area under this curve closer to zero is more desirable. The areas under the CMC curve for the proposed method and

logistic regression are very closely matched, at 0.059 vs. 0.051 (Figure 10).

22

410

411
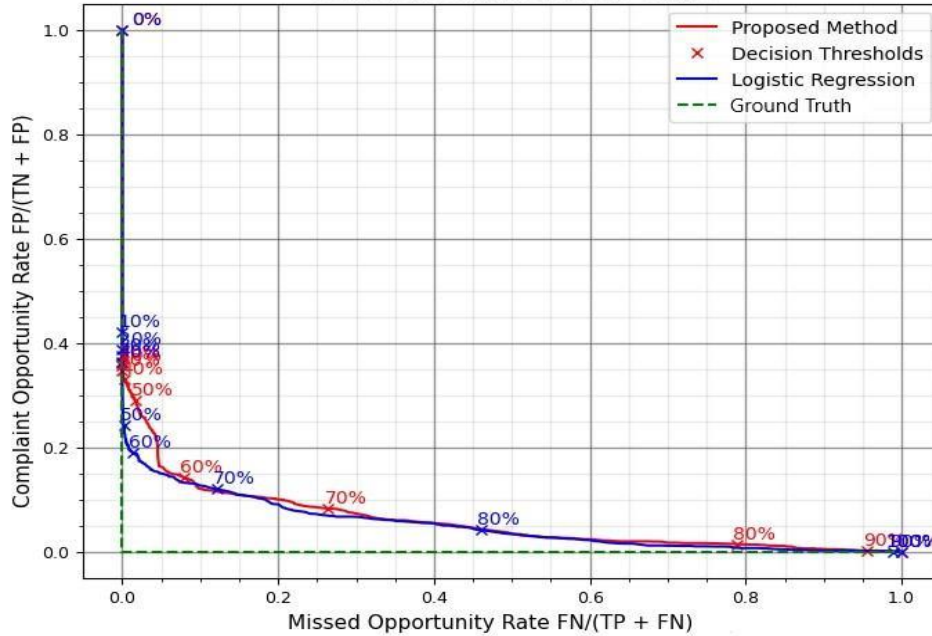
412 At decisions thresholds of about 63% for the VIE, COR and MOR are nearly equal, but for most applications building

413 managers would choose a decision threshold that sets the rate of complaint lower than the rate of missed energy

414 savings. For higher risk applications such as elevators, etc., a decision threshold of 80% is appropriate whereas for

415 lower risk applications such as water fountains 60% might be appropriate.

416 Note that at a COR of 0.2, logistic regression yields a MOR of 0.02, whereas the proposed method yields a MOR of

417 0.05. In this range, the proposed method performs worse than logistic regression, however as discussed, this is not a

418 typical operating range for most practical applications. Table 2 summarizes the results from Figures 9 and 10 for the

419 ideal, VIE, and logistic regressions.

420 **Table 2.** Comparison of overall model performance between VIE and logistic regression.

Method	Area under CMC Curve	Area under ROC Curve
Ideal values	0.0	1.0
VIE	0.059	0.94

23

Logistic regression	0.051	0.95

421

5. Model Limitations and Future Work

423   The limitations of this study can be divided into limitations of the Vacancy Inference Engine model design and
424   physical limitations of the model input data. These limitations also point to future work to improve the method.

425   A key assumption of the VIE is that the sensor inputs are independent, but sensors employed in the present study were
426   not specifically tested for independence. A second limitation is that the VIE does not iteratively update. Thus, in order
427   to maintain accuracy, the model may need to be retrained over time, especially as new sensors are added. A third
428   limitation of the proposed VIE is that it must be retrained for each new building because the sensor reading during
429   vacancy will be different in each case.

430   This case study illustrates the physical limitations of the inputs. The building examined was only a zone of a larger
431   building (though physically separated) and the investigation did not capture variations beyond the scale of a few
432   weeks. Some of the sensors exhibited data dropout that required filtering and manual inspection, a level of attention
433   not feasible in routinely monitored buildings without automation. Furthermore, none of the data were collected and
434   processed in real-time, so the real-time application of the VIE has not yet been demonstrated. The approach described
435   here applies only to analog signals. Future work will include extension to binary sensor outputs, such as from card
436   key swipes and motion sensors. We believe that the sensor fusion approach can be extended to binary signals as long
437   as the sensor outputs are related to vacancy.

438   Finally, the model needs to be applied to different types of buildings, over longer time periods, and with different
439   combinations of sensor inputs. These investigations will contribute to a generalized model applicable to most
440   commercial buildings.

6. Conclusions

Commercial buildings often have frequent periods of vacancy, but electricity demand during these periods falls only slightly. This paper's first, new contribution is the recognition that complete building vacancy – when zero people are present – is a unique building condition. Furthermore, this vacant condition is more common than usually supposed and potentially involves a significant fraction of a commercial building's electricity consumption. If periods of vacancy can be confidently determined, then more aggressive energy-saving measures can be applied during those times by switching off miscellaneous electrical loads or greatly reducing their service levels.

This paper's second contribution is that the determination of vacancy is a unique problem requiring a new approach. Unlike occupancy – a condition that can be measured – vacancy must be inferred. Sensor fusion offers a means of creating an inference of vacancy from a collection of diverse sensors. The method is robust in the sense that it can accommodate the variable and intermittent characteristics of building sensors. The resulting output is a probability of vacancy. Each MEL can respond differently to the vacancy probability based on the services it provides. MELs providing less essential services can switch off during periods of low vacancy probability while MELs providing key services would only switch off at high probabilities. Thus, a pathway exists to save energy across a wide array of devices and services. Future research must still be undertaken to make devices responsive to vacancy levels. Fortunately, many of the components are already present in modern MELs, even if they have not been applied to precisely this goal.

This paper's third contribution deals with training a vacancy model. Only modest training is needed for vacancy inference because the relationships between the detected conditions and occupancy are similar in many buildings. In this way, training from one building can be transferred to another. This is in contrast to occupancy, where the relationships will often be unique to each building.

This study developed and tested a vacancy inference engine that employed multiple pre-existing sensors with near-zero ground truth requirements during model training. The sensors used in our test building were typical; electricity demand, room carbon dioxide concentrations, and number of active Wi-Fi connections. Although the method was tested only at a research office building, it should apply to any enclosed space. Further improvements are necessary;

466 however, this work identified a mostly-untapped opportunity for energy savings and a promising method to obtain

467 them.

Acknowledgments

468

472 **References:**

473 [1]   US Department of Energy, Quadrennial Technology Review 2015, Energy.Gov. (2015).

474   https://www.energy.gov/quadrennial-technology-review-2015 (accessed July 2, 2021).

475 [2]   Analysis and Representation of Miscellaneous Electric Loads in NEMS - Energy Information Administration,

476   (n.d.). https://www.eia.gov/analysis/studies/demand/miscelectric/ (accessed October 13, 2021).

477 [3]   K. Roth, K. Mckenney, C. Paetsch, R. Ponoum, U.S. Residential Miscellaneous Electric Loads Electricity

478   Consumption, Proc. ACEEE Summer Study Energy Effic. Build. (2008) 12.

479 [4]   S. Kwatra, Miscellaneous Energy Loads in Buildings, (n.d.) 95.

480 [5]   A. Kamilaris, B. Kalluri, S. Kondepudi, T. Kwok Wai, A literature survey on measuring energy usage for

481   miscellaneous electric loads in offices and commercial buildings, Renew. Sustain. Energy Rev. 34 (2014)

482   536–550. https://doi.org/10.1016/j.rser.2014.03.037.

483 [6]   J. Butzbaugh, R. Hosbach, A. Meier, Miscellaneous electric loads: Characterization and energy savings

484   potential, Energy Build. 241 (2021) 110892. https://doi.org/10.1016/j.enbuild.2021.110892.

485 [7]   W. Merrill, Building Electric Appliances, Devices, and Systems, Energy.Gov. (n.d.).

486   https://www.energy.gov/eere/buildings/building-electric-appliances-devices-and-systems (accessed July 2,

487   2021).

488 [8]   M. Hafer, Quantity and electricity consumption of plug load equipment on a university campus, Energy Effic.

489   10 (2017) 1013–1039. https://doi.org/10.1007/s12053-016-9503-2.

490 [9]   Q.J. Kwong, J.E. Lim, M.S. Hasim, Miscellaneous electric loads in Malaysian buildings - Energy

491   management opportunities and regulatory requirements, Energy Strategy Rev. 21 (2018) 35–49.

492   https://doi.org/10.1016/j.esr.2018.04.002.

493   [10]   H.A. Rahman, M.S. Majid, M.Y. Hassan, T.S. Lian, Energy Savings Through Power Management in the

494   Desktop Computer, (n.d.).

495   [11]   S.P. Borg, N.J. Kelly, The effect of appliance energy efficiency improvements on domestic electric loads in

496   European households, Energy Build. 43 (2011) 2240–2250. https://doi.org/10.1016/j.enbuild.2011.05.001.

497   [12]   I.P. Mohottige, T. Sutjarittham, N. Raju, H.H. Gharakheili, V. Sivaraman, Role of Campus Wi-Fi

498   Infrastructure for Occupancy Monitoring in a Large University, in: 2018 IEEE Int. Conf. Inf. Autom. Sustain.

499   ICIAfS, 2018: pp. 1–5. https://doi.org/10.1109/ICIAFS.2018.8913341.

500   [13]   Y. Wang, L. Shao, Understanding occupancy pattern and improving building energy efficiency through Wi-Fi

501   based indoor positioning, Build. Environ. 114 (2017) 106–117.

502   https://doi.org/10.1016/j.buildenv.2016.12.015.

503   [14]   Z. Wang, T. Hong, M.A. Piette, M. Pritoni, Inferring occupant counts from Wi-Fi data in buildings through

504   machine learning, Build. Environ. 158 (2019) 281–294. https://doi.org/10.1016/j.buildenv.2019.05.015.

505   [15]   M.M. Ouf, M.H. Issa, A. Azzouz, A.-M. Sadick, Effectiveness of using Wi-Fi technologies to detect and

506   predict building occupancy, Sustain. Build. 2 (2017) 7. https://doi.org/10.1051/sbuild/2017005.

507   [16]   J. Chen, C. Ahn, Assessing occupants' energy load variation through existing wireless network infrastructure

508   in commercial and educational buildings, Energy Build. 82 (2014) 540–549.

509   https://doi.org/10.1016/j.enbuild.2014.07.053.

510   [17]   S. Zhan, A. Chong, Building occupancy and energy consumption: Case studies across building types, Energy

511   Built Environ. 2 (2021) 167–174. https://doi.org/10.1016/j.enbenv.2020.08.001.

512   [18]   L.M. Slaughter, A Modular Semi-Supervised Sensor Fusion Method for Inferring Real Time Vacancy in

513   Buildings, M.S., University of California, Davis, 2019.

514   https://www.proquest.com/pqdt/docview/2384571533/abstract/4311B947E83F4844PQ/1 (accessed June 7,

515   2021).

516   [19]   A.J. Sloan, Energy Consumption in Campus Buildings when No One is around, M.S., University of

517   California, Davis, 2019.

518   https://www.proquest.com/pqdt/docview/2312585978/abstract/41E7485ABECF451DPQ/1 (accessed June 7,

519   2021).

520   [20]   J. St. John, Why Empty Office Buildings Still Consume Lots of Power During a Global Pandemic, Greentech

521   Media. (2020). https://www.greentechmedia.com/articles/read/how-office-buildings-power-down-during-

522   coronavirus-lockdown (accessed July 2, 2021).

523   [21]   S. Kaplan, An Empire State of Green, Wash. Post. (2021).

524   https://www.washingtonpost.com/graphics/2020/climate-solutions/empire-state-building-emissions/ (accessed

525   July 2, 2021).

526   [22]   P. Anand, D. Cheong, C. Sekhar, A review of occupancy-based building energy and IEQ controls and its

527   future post-COVID, Sci. Total Environ. 804 (2022) 150249. https://doi.org/10.1016/j.scitotenv.2021.150249.

528   [23]   A. Franco, F. Leccese, Measurement of CO_2 concentration for occupancy estimation in educational buildings

529   with energy efficiency purposes, J. Build. Eng. 32 (2020) 101714. https://doi.org/10.1016/j.jobe.2020.101714.

530   [24]   B. Hobson, B. Gunay, A. Ashouri, G. Newsham, Wi-Fi based occupancy clustering and motif identification:

531   A case study, in: 2020.

532   [25]   C.M. Stoppel, F. Leite, Integrating probabilistic methods for describing occupant presence with building

533   energy simulation models, Energy Build. 68 (2014) 99–107. https://doi.org/10.1016/j.enbuild.2013.08.042.

534   [26]   J. Page, D. Robinson, N. Morel, J.-L. Scartezzini, A generalised stochastic model for the simulation of

535   occupant presence, Energy Build. 40 (2008) 83–98. https://doi.org/10.1016/j.enbuild.2007.01.018.

536   [27]   W.-K. Chang, T. Hong, Statistical analysis and modeling of occupancy patterns in open-plan offices using

537   measured lighting-switch data, Build. Simul. 6 (2013) 23–32. https://doi.org/10.1007/s12273-013-0106-y.

538   [28]   D. Wang, C.C. Federspiel, F. Rubinstein, Modeling occupancy in single person offices, Energy Build. 37

539   (2005) 121–126. https://doi.org/10.1016/j.enbuild.2004.06.015.

540   [29]   C. Martani, D. Lee, P. Robinson, R. Britter, C. Ratti, ENERNET: Studying the dynamic relationship between

541   building occupancy and energy consumption, Energy Build. 47 (2012) 584–591.

542   https://doi.org/10.1016/j.enbuild.2011.12.037.

543   [30]   B. Howard, S. Acha, N. Shah, J. Polak, Implicit Sensing of Building Occupancy Count with Information and

544   Communication Technology Data Sets, Build. Environ. 157 (2019) 297–308.

545   https://doi.org/10.1016/j.buildenv.2019.04.015.

546   [31]   H.N. Rafsanjani, C.R. Ahn, K.M. Eskridge, Understanding the recurring patterns of occupants' energy-use

547   behaviors at entry and departure events in office buildings, Build. Environ. 136 (2018) 77–87.

548   https://doi.org/10.1016/j.buildenv.2018.03.037.

549   [32]   Y.-S. Kim, J. Srebric, Impact of occupancy rates on the building electricity consumption in commercial

550   buildings, Energy Build. 138 (2017) 591–600. https://doi.org/10.1016/j.enbuild.2016.12.056.

551   [33]   P. Anand, D. Cheong, C. Sekhar, M. Santamouris, S. Kondepudi, Energy saving estimation for plug and

552   lighting load using occupancy analysis, Renew. Energy. 143 (2019) 1143–1161.

553   https://doi.org/10.1016/j.renene.2019.05.089.

554   [34]   X. Liang, T. Hong, G.Q. Shen, Occupancy data analytics and prediction: A case study, Build. Environ. 102

555   (2016) 179–192. https://doi.org/10.1016/j.buildenv.2016.03.027.

556   [35]   N. Alishahi, M. Nik-Bakht, M.M. Ouf, A framework to identify key occupancy indicators for optimizing

557   building operation using Wi-Fi connection count data, Build. Environ. 200 (2021) 107936.

558   https://doi.org/10.1016/j.buildenv.2021.107936.

559   [36]   P. Gandhi, G.S. Brager, Commercial office plug load energy consumption trends and the role of occupant

560   behavior, Energy Build. 125 (2016) 1–8. https://doi.org/10.1016/j.enbuild.2016.04.057.

561   [37]   A. Meier, Putting Vacant Buildings to Sleep, (2020). https://youtu.be/FdEKyGUDPGQ.

562   [38]   O.T. Masoso, L.J. Grobler, The dark side of occupants' behaviour on building energy use, Energy Build. 42

563   (2010) 173–177. https://doi.org/10.1016/j.enbuild.2009.08.009.

564   [39]   Y. Ding, D. Ivanko, G. Cao, H. Brattebø, N. Nord, Analysis of electricity use and economic impacts for

565   buildings with electric heating under lockdown conditions: examples for educational buildings and residential

566   buildings in Norway, Sustain. Cities Soc. 74 (2021) 103253. https://doi.org/10.1016/j.scs.2021.103253.

567   [40]   Y. Kim, Calibration of building energy simulations with occupancy and plug-load schedules derived from

568   metered building electricity consumption, (2014). https://etda.libraries.psu.edu/catalog/23664 (accessed

569   October 27, 2021).

570   [41]   M. Avci, M. Erkoc, A. Rahmani, S. Asfour, Model predictive HVAC load control in buildings using real-time

571   electricity pricing, Energy Build. 60 (2013) 199–209. https://doi.org/10.1016/j.enbuild.2013.01.008.

572   [42]   T. Labeodan, W. Zeiler, G. Boxem, Y. Zhao, Occupancy measurement in commercial office buildings for

573   demand-driven control applications—A survey and detection system evaluation, Energy Build. 93 (2015)

574   303–314. https://doi.org/10.1016/j.enbuild.2015.02.028.

575   [43]   G. Fajilla, M. Chen Austin, D. Mora, M. De Simone, Assessment of probabilistic models to estimate the

576   occupancy state in office buildings using indoor parameters and user-related variables, Energy Build. 246

577   (2021) 111105. https://doi.org/10.1016/j.enbuild.2021.111105.

578   [44]   Z. Yang, B. Becerik-Gerber, Modeling personalized occupancy profiles for representing long term patterns by

579   using ambient context, Build. Environ. 78 (2014) 23–35. https://doi.org/10.1016/j.buildenv.2014.04.003.

580   [45]   D. Trivedi, V. Badarla, Occupancy detection systems for indoor environments: A survey of approaches and

581   methods, Indoor Built Environ. 29 (2020) 1053–1069. https://doi.org/10.1177/1420326X19875621.

582   [46]   B. Khaleghi, A. Khamis, F.O. Karray, S.N. Razavi, Multisensor data fusion: A review of the state-of-the-art,

583   Inf. Fusion. 14 (2013) 28–44. https://doi.org/10.1016/j.inffus.2011.08.001.

584   [47]   S. Naylor, M. Gillott, T. Lau, A review of occupant-centric building control strategies to reduce building

585   energy use, Renew. Sustain. Energy Rev. 96 (2018) 1–10. https://doi.org/10.1016/j.rser.2018.07.019.

586   [48]   B. Chandrasekaran, S. Gangadhar, J. Conrad, A survey of multisensor fusion techniques, architectures and

587   methodologies, SoutheastCon 2017. (2017). https://doi.org/10.1109/SECON.2017.7925311.

588   [49]   H. Saha, A.R. Florita, G.P. Henze, S. Sarkar, Occupancy sensing in buildings: A review of data analytics

589   approaches, Energy Build. 188–189 (2019) 278–285. https://doi.org/10.1016/j.enbuild.2019.02.030.

590   [50]   L. Rueda, K. Agbossou, A. Cardenas, N. Henao, S. Kelouwani, A comprehensive review of approaches to

591   building occupancy detection, Build. Environ. 180 (2020) 106966.

592   https://doi.org/10.1016/j.buildenv.2020.106966.

593   [51]   J. Chaney, E. Hugh Owens, A.D. Peacock, An evidence based approach to determining residential occupancy

594   and its role in demand response management, Energy Build. 125 (2016) 254–266.

595   https://doi.org/10.1016/j.enbuild.2016.04.060.

596   [52]   Z. Li, B. Dong, A new modeling approach for short-term prediction of occupancy in residential buildings,

597   Build. Environ. 121 (2017) 277–290. https://doi.org/10.1016/j.buildenv.2017.05.005.

598   [53]   OpenCV Team, Open Source Computer Vision Library (OpenCV), OpenCV. (2019). https://opencv.org/

599   (accessed June 7, 2021).

600   [54]   M. Olafenwa, J. Olafenwa, ImageAI, an open source python library built to empower developers to build

601   applications and systems with self-contained Computer Vision capabilities, (2018). http://www.imageai.org

602   (accessed June 7, 2021).

603