Vepsäläinen, Jari; Otto, Kevin; Lajunen, Antti; Tammi, Kari

**Computationally efficient model for energy demand prediction of electric city bus in varying operating conditions**

*Published in:*  
Energy

*DOI:*  
10.1016/j.energy.2018.12.064

Published: 15/02/2019

*Document Version*  
Publisher's PDF, also known as Version of record

*Please cite the original version:*  
https://doi.org/10.1016/j.energy.2018.12.064
Computationally efficient model for energy demand prediction of electric city bus in varying operating conditions

Jari Vepsäläinen a, *, Kevin Otto a, Antti Lajunen a, b, Kari Tammi a

a School of Engineering, Department of Mechanical Engineering, Aalto University, Espoo, Finland
b Faculty of Agriculture and Forestry, Department of Agricultural Sciences, University of Helsinki, Helsinki, Finland

A R T I C L E   I N F O
Article history:
Received 11 May 2018
Received in revised form 16 November 2018
Accepted 11 December 2018
Available online 11 December 2018

Keywords:
Surrogate modeling
Simulation
Energy demand
Electric bus
Sensitivity analysis
Uncertainty

A B S T R A C T
The uncertainty of operating conditions such as weather and payload cause variations in the energy demand of electric city buses. Uncertain variation in energy demand is a challenge in the design of charging systems and on-board energy storages. To predict the energy demand, a computationally efficient model is required for real-time applications. We present a novel approach to predict energy demand variation with a wide range of uncertain factors. A factor identification is carried out to recognize the range of variation in the operating conditions. A computationally efficient surrogate model is generated based on a previously developed numerical simulation model. The surrogate model is shown to be 10 000 times faster than the numerical model. The surrogate model output corresponds with the numerical model with less than 1% error. The energy demand of the surrogate model varied from 0.43 to 2.30 kWh/km, which is realistic in comparison to previous studies. Successful sensitivity analysis of the surrogate model revealed the most crucial factors. Uncertainty in temperature, rolling resistance and payload contributed most to the variation in energy demand. Variation in these factors should be taken into account when predicting energy consumption and while planning schedules for a bus network.

© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Most city buses are still diesel-powered, even though an electric powertrain would offer superior efficiency, higher peak torque at low speeds, zero tank-to-wheel emissions and lower noise levels [1]. Diesel buses are still preferred because of the higher investment costs of electric buses, which are roughly double compared to diesel counterparts [2]. The costs are dictated by the expensive lithium-ion batteries and charging systems. The investment costs can be lowered with optimal battery sizing and shared charging infrastructure [3]. Electric city buses are often recharged at the route terminuses or at bus stops to allow minimal battery size and charging power [4]. Electric passenger vehicles on the other hand are typically recharged overnight. In addition, passenger cars do not operate on predetermined routes, have no driving schedules nor need to have extra-auxiliary devices, such as pneumatic actuators for doors. For these reasons, the electric vehicle studies can be used as a reference point, yet separate analysis is needed for the energy consumption of electric city buses.

The prediction of the energy consumption of electric vehicles and buses have been investigated previously and can be divided according to the given input factors. These input factors are mission-related kinematic factors (i.e., driving aggressiveness, stops/km and travel time) and vehicle-related factors (i.e., ambient temperature, auxiliary power consumption, payload, and drive efficiency). Mission-related studies of electric passenger and lightweight vehicles report higher energy consumption on highway routes than on city routes [5–7]. The route characteristics and stops per kilometer have also been reported to affect the consumption of electric city buses [8–10]. In addition, the driving style has been shown to have a notable impact on energy demand variation of electric city buses [11].

Previous studies have also investigated the auxiliary and heating device usage, which has a significant impact on energy consumption [5,7,12]. Lajunen [13] further investigated the cost-benefits regarding different electric and hybrid powertrains of buses with varying operating routes. The powertrain design had a significant effect on consumption, but the effect was also dependent on the route and scheduling. De Cauwer, Van Mierlo and Coosemans [14]
implemented a consumption prediction model using a longitudinal dynamics model (LDM) of an electric passenger vehicle. A similar LDM was used by Asamer et al. [15] to study the impact of variation in vehicle-related factors to the energy demand. They reported that the drive efficiency, rolling resistance and auxiliary power demand were the most crucial factors causing variation in energy demand.

In the long term, the bus manufacturer and operator, the public transport authority (PTA) and the public transport users benefit from the quantification of the energy uncertainty. Better prediction of variation in energy demand would improve the reliability of bus schedules and reduce unnecessary overload of the electric grid due to concurrent recharging events [16,17]. Overload peaks can also be alleviated by installing more charging stations to share the charge load amongst multiple chargers for significant savings in charging costs [18]. The characterization of energy demand under uncertainty can also help in sustainable route planning [19]. Thus, the energy demand uncertainty is quantified here to minimize costs and the environmental burden. Unnecessary over-dimensioning of on-board energy storage leads to under-utilization of the battery capacity and increased consumption because of the extra payload. Furthermore, Reiter et al. [20] argued that the electric vehicle battery capacity could be safely utilized more than the typical 70−85% to increase range, especially in rare worst-case events.

We present a novel approach to gain an insight into the energy consumption of electric buses. The approach includes identification of uncertain input factors, creation of fast-computing surrogate model and sensitivity analysis. The uncertainty margins of 14 noise factors affecting the energy demand are identified. The surrogate model is developed from Monte Carlo simulations executed with a dynamic electro-mechanical model of the electric bus. The electro-mechanical model was developed previously [21], specifically to study uncertainties in the energy demand. For the sensitivity analysis, we extend the idea presented by Asamer et al. [15] to electric buses with a wider scope of input factors. We claim that the approach increases the understanding of energy consumption, its prediction, and possible variations under different driving conditions and thus represents an innovation.

The paper consists of three parts. First, the original electric city bus (ECB) model is presented, followed by the development of the surrogate polynomial chaos expansion (PCE) model that is used for the sensitivity analysis. After the sensitivity analysis procedure has been explained, the factor uncertainty identification, model fidelity and sensitivity analysis results are presented. Finally, discussion of the research is followed by the conclusion.

2. Methods

2.1. ECB model

The previous developed ECB is openly available in Ref. [22] and it is based on the model presented by Halmeaho et al. [23]. The model is presented in Fig. 1 and it consists of Newtonian 1-DOF mechanics, simplified powertrain, and a lithium-ion battery model, described in Appendix A. The model is used to generate a surrogate PCE model, to avoid extensive computation of the heavier ECB and for future light-computation use in energy consumption prediction. The PCE model is generated from uniformly distributed Monte Carlo simulations (MCSs). The surrogate model fit and sensitivity analysis are calculated with the UQLab toolbox for MATLAB [24]. The uncertain input factors are called noise factors (NFs). The 14 considered NFs are showed in Fig. 1, according to the component they affect. The names of specific NFs and their uncertainty margins are listed in Table 1.

2.2. Surrogate PCE model

The primary tool for input factor prioritization is sensitivity analysis (SA). A sensitivity analysis can be performed locally (LSA) by studying one factor at a time or globally (GSA), when all the factors are studied at the same time. GSA approach was chosen, because LSA is limited to linear regression models and lacks the factor interaction and high-order effects provided by the GSA [25]. The GSA process phases are shown in Fig. 2, starting from the numerical simulation model development and resulting with sensitivity indices.

The previously developed simulation model, ECB, requires a runtime of 70 s to simulate a 25-min driving cycle (Espoo 11 route) on a powerful desktop computer. The Espoo 11 route driving cycle variations are not considered here, yet have been previously considered in the works of [26−28]. MCSs with the ECB model require excessive computing, which is why a surrogate model is developed. A surrogate model is a simplified analytical description that approximates the original model [29]. There are many surrogate modeling techniques, yet PCE is selected for its simplicity and accuracy [30]. The ECB model M(X) is a function of stochastic input noise factors x_i, $Y_{DR} = M(X)$, (1)

where $Y_{DR}$ are the model responses (energy demand and regeneration). The margins of the 14 input noise factors $X = x_1, x_2, \ldots, x_{14}$ need to be identified to set a realistic search space, before executing the MCSs with the ECB model. The objective is to run as few MCSs as possible, yet the multidimensional search space must be evenly and thoroughly explored. A quasi-randomized Sobol’ sequence is used as the space filling method to generate a joint, uniform distribution of the input factors [31]. Such low-discrepancy set of points is required to achieve a uniform response surface to accurately fit a PCE surrogate model.

The ECB model $M(X)$ is considered as a black-box and the simulation data is used for the surrogate modeling instead of analytical representation, which would be unnecessarily complex. The PCE approximation is represented as

$$M(X) = M^{PCE}(X) = \sum_{a \in A} y_a \phi_a(X),$$ (2)

where the surrogate PCE model $M^{PCE}(X)$ is the sum of orthogonal multivariate polynomials $\phi_a(X)$ with the respected coefficients $y_a$, for the given multi-indices $a$ [32]. A multiple linear regression model consist of 1-degree multivariate polynomials for each variable. Nonlinear behavior can also be represented by adding more degrees of freedom to describe the effect of each variable. To create a finite model, polynomial degree-level of $A$ is selected. This selection is referred as truncation and the finite model as full-basis model. The standard truncation scheme was used, which considers all the variables of degrees less than $A = 10$.

The accuracy of the surrogate model is measured with a $k$-fold cross-validation. The $k$ is the number of validation iterations. For
A special case of k-fold cross-validation is the leave-one-out error (LOOCV), which consists of as many TSMs as there are data points [34]. The independent TSMs (each iteration, a new temporary surrogate model (TSM)) is generated from a randomly selected subset of the simulated data. In the case of k = 5, independent TSMs are generated using 80% of the simulated data and the remaining 20% is used for validating the TSM in question. The higher the k, the more accurate the error estimation [33]. A special case of a k-fold cross-validation is the leave-one-out cross-validation (LOOCV), which consists of as many TSMs as there are data points [34]. The independent TSMs (\( \mathbf{M}_{\text{PCE}} \)) each leave one observation out of the model. The predictions of the TSMs are compared with the simulation data of the one excluded observation. The total error of the actual surrogate model is computed with the predicted residual sum of squares (PRESS) [35]. The resulting leave-one-out error \( \varepsilon_{\text{LOO}} \) of the model is

\[
\varepsilon_{\text{LOO}} = \frac{1}{N} \sum_{i=1}^{N} (M(x_i') - \mathbf{M}_{\text{PCE}}(x_i'))^2 = \frac{1}{N} \sum_{i=1}^{N} (M(x_i') - \bar{\mu}_Y)^2, \tag{3}
\]

where \( \bar{\mu}_Y = \frac{1}{N} \sum_{i=1}^{N} M(x_i') \) is the sample mean of the simulation response.

Since the PCE is based on the MCSs, the PCE can never be more accurate than the original model. However, in practical applications, perfect theoretic accuracy is not pursued. Typically, we are more interested in achieving a “close enough” useful estimate with as little training as possible, defined with an error estimate. The error estimation is used to further truncate the surrogate model. To limit the degree of polynomials, the least angle regression (LAR) algorithm is employed to iteratively search for a more compact solution. Typically, most of the higher-level interactions terms do not affect the model response significantly and can be discarded. The full-basis PCE is compared to an iteratively generated PCE model with increasing order of degree. The algorithm is repeated until a target accuracy is achieved or a maximum number of iterations is reached [35]. In Fig. 3, the leave-one-error is shown as a function of PCE models of increasing degree.

2.3. Sensitivity analysis

Once the surrogate model fit is satisfactory, Sobol’s sensitivity indices (SIs) are computed. The SIs were originally introduced by Sobol’ [36] and later enhanced by Saltelli, Andres and Homma [37] and Saltelli et al. [38]. The objective of Sobol’s SIs is to minimize the number of tests required for first-order effect \( S_1 \) and total effect \( S_T \).

\[
S_1 = \frac{V(E(Y/X_i))/V(Y)}{2},
\tag{4}
\]

\[
S_T = 1 - V(E(Y/X_{-i}))/V(Y),
\tag{5}
\]

where \( E \) is the expected value of output \( Y \) with different input matrices. The brute-force calculation of the SIs would require excessive computing, which is why Saltelli et al. [31] present shortcuts to the computation. Their method only requires \( N(p + 2) \) computations whereas the brute-force method requires \( N^2 \). Here \( p \) is the number of factors and \( N \) the number of samples. To reduce computation effort even more, the coefficients of the multivariate polynomials of the PCE surrogate model can also be used to calculate the SIs [25].

3. Results

3.1. Uncertainty identification

The identification of the input factor uncertainty is based on measurements or literature data [5,15,39]. A worst- and best-case scenario values are set for each factor by implementing one-factor-at-a-time tests. The resulting margins for the 14 factors are presented in Table 1, which are divided into tolerance noise and extensive noise. Five of the factors were classified as tolerance factors, which were the battery capacitance and capacity, the motor resistance, permanent flux and inductance. The tolerance bounds were set to ±10% except for the battery capacity which is unlikely to be more than informed by the manufacturer but is possibly worse due to production errors (e.g. contaminated materials, assembly tolerances). Thus, the tolerance range for the battery capacity is set from −10% to 0%. The rest of the factors are classified as extensive noise, which are each examined individually.

**Mass:** the studied city bus is a light-weight aluminum frame bus, with a curb weight of 8500 kg. The maximum load for the vehicle is 6500 kg, resulting in a maximum total mass of 15 000 kg.

**Headwind:** the wind speed and direction in Espoo is analyzed with open source data provided by the Finnish Meteorological Institute (FMI) for each day of the year 2016 [40]. In Fig. 4a, the wind
speed is shown as a function of direction, 0° representing north. The measurement sample time was 10 min and the samples are presented with purple circles. The maximum wind speed was 12.2 m/s and the average standard deviation between every three samples was 0.4 m/s and 14°. However, the nature of the wind is ever-changing and in the worst-case scenario the wind could be...

Fig. 3. The evolution of leave-one-out error $\varepsilon_{\text{LOO}}$ with increasing LAR iterations. The PCE model generated by a LAR iteration with the smallest $\varepsilon_{\text{LOO}}$ will be selected. In this example case, the error saturates after LAR iteration 50.

Fig. 4. Temperature related effects: a) The input wind speed as a function of direction (Espoo 2016), where the zero angle represents north; b) Estimated input HVAC power consumption as a function of ambient temperature (modified from Ref. [41]); c) The input distribution of daily temperatures in Espoo 2012–2016.
facing the bus throughout the route. Given the gusts of wind and the varying course of the route, the headwind margins are set to $-10 \text{ m/s}$ and $10 \text{ m/s}$.

**Rolling resistance coefficient**: the rolling resistance coefficient is determined by tire wear and pressure, terrain and ambient temperature [42]. In the best-case the rolling resistance is 0.006 with a smooth asphalt and new tires during summer time. In the worst-case there is shallow snow on the road and the tires are worn, represented with a rolling resistance coefficient of 0.02.

**Ambient Temperature**: the temperature fluctuation in Espoo was analyzed with open source data acquired from the Finnish Meteorological Institute [40]. The upper and lower temperature bounds were estimated based on data from 2012 to 2016, as displayed in Fig. 4c. The minimum and maximum temperatures were $-26.2$ and $31.6 \degree C$, respectively. The range of temperature was rounded up to $-30 \degree C$ as minimum and to $35 \degree C$ as maximum. Furthermore, the power consumed by the heating, ventilation and air conditioning (HVAC) system is dependent on ambient temperature [43] and humidity [44]. Since the HVAC system is not separately modeled, the temperature dependent HVAC power is estimated and interpolated according to Lajunen and Tammi [41]; as shown in Fig. 4b. The effect of humidity is included in this approximation. Liu et al. [45] confirm a similar consumption-temperature dependency behavior of passenger EVs. Suh et al. [46] also report near 30 kW HVAC power consumption in real-world electric bus tests. In addition, the air density is increased at low temperatures, which increases aerodynamic drag. The air density fluctuates 25% because of the ambient temperature variation.

**Battery Temperature**: the LiFePO$_4$ battery pack of the bus can heat up to 60 \degree C without cooling [47]. The liquid cooling system reduces the effect of ambient temperature changes and high load situations in the battery pack [48]. Nevertheless, the battery temperature fluctuates between 15 and 30 \degree C, according to measurements on a real-world electric city bus.

**Battery State-of-Charge (SoC)**: In the best-case, the battery is initially full and in the worst-case only 50% of the battery SoC is remaining. The worst-case battery SoC was determined as the minimum charge needed to complete the mission when other input factors were also set to their worst-case value.

**Battery Resistance**: Ecker et al. [49] have estimated that at the end of resistive life the internal resistance of a lithium-ion battery has doubled, which was also noted by Rogge, Wollny and Sauer [50]. The internal resistance variation margins are thus 100% and 200% of the original value. In addition, the initial value of internal resistance also grows exponentially as temperature decreases.

**Battery Cycle Age**: The battery ages with charge-discharge cycles, excessive currents, elevated temperatures [51] and storage time [52,53]. Lithium-ion batteries degenerate after each charge-discharge cycle and with time [48]. After 20% decrease in the capacity the solid electrolyte interphase (SEI) has advanced so far that the battery might short-circuit internally and should no longer be used in automotive applications [53–55]. According to Rogge, Wollny and Sauer [50], if either the capacity fade of 20% or the internal resistance is doubling is reached first, the battery is not safe to use anymore.

**Auxiliary Power**: The doors of the bus are operated with a 6 kW air compressor. In addition, hydraulic power is needed for power steering and braking systems which consume approximately 1.5 kW continuously. Other auxiliary devices have an estimated average power of 1 kW. Given these characteristics, the best possible situation is estimated as average continuous auxiliary power of 2 kW. In the worst-case the doors are opened frequently and the auxiliary device power is maximum, leading to an estimate of 7 kW power consumption in the worst-case scenario.

### 3.2. Efficiency of surrogate PCE model

The PCE model was mainly developed to increase computational efficiency. High computational efficiency is imperative for sensitivity analysis with multiple factors and in real-time applications. In the context of electric vehicle modeling, Genikomsakis and Mitrentsis [56] introduced a computation speed benchmark method. They performed 60 model evaluations to determine the average computation speed of a simulation model. The fastest computation speed they achieved on an electric vehicle model that

![Fig. 5](image-url). The computed output probability distributions of the consumption (left column) and regeneration (right column). The top figures are distributions of 1500 evaluations with the original ECB model and the bottom distributions were achieved with 100 000 PCE surrogate model runs. Similarity is obvious, yet the surrogate model results are more consistent.
estimates energy consumption was over 50 000 times faster than real-time.

In order to define the computational gain between our ECB and PCE models, we evaluated both models 60 times. The average computing times for the ECB and PCE model on a 1548 s route were 38.16 s and 0.0038 s, respectively. In comparison, the ECB model was 40 times faster and the PCE model over 400 000 times faster than real-time. The experiment desktop computer had an Intel(R) Xeon(R) CPU @ 3.40 processor with 32 GB of RAM.

Model accuracy and computational efficiency are equally important for real-time prediction of energy consumption. The leave-one-out cross-validation error was less than 1% when comparing the ECB and PCE model outputs. The comparison was carried out with 1500 reference simulations with different combinations of input factor values. In conclusion, the developed surrogate model is both computationally efficient and is able to accurately predict energy consumption in various operating conditions.

3.3. Sensitivity analysis

The sensitivity of the surrogate model was examined with uniformly distributed factors within the uncertainty margins given in Table 1. The studied system responses were energy consumption and energy regeneration. Distributions of the responses are shown in Fig. 5. The energy consumption distribution resembles a Gaussian distribution, with a skew to the decreased consumption. The distribution of regeneration is oppositely skewed, as expected. At worst, none of the energy consumed was recovered through regenerative braking, although up to 54% of the energy spent was regenerated at best. On average, 28.0% of the consumed energy was regenerated with a standard deviation of 8.7%, as shown in Table 2. Björnsson and Karlsson [57] reported an average regeneration of 27% with real-world driving data in Sweden that was analyzed with an electric vehicle model. Soylu [8] also observed a 27% average regeneration of the traction power with a hybrid electric bus, with a maximum of 52% energy regeneration in the most favorable conditions. Electric city buses tend to have higher regeneration rates than other vehicles due to their operation cycle with multiple bus stops. The simulated energy consumption varied between 0.43 kWh/km and 2.30 kWh/km with an average of 1.20 kWh/km. The coefficient of variation (CV) was over 25% for both responses, which is the ratio of the standard deviation to the mean. The CV is a dimensionless value, which allows comparison of variates with different datasets. The CV of both responses, energy demand and regeneration, indicates to notable uncertainty [58].

To achieve low cross-validation errors, only 1000 data points from the MCSs were necessary in the surrogate PCE model generation. The fully descriptive surrogate model would have required 15504 polynomials, but the truncated sparse PCE model had maximum degree of 5. The sparse models are constructed with only 118 and 155 polynomials for consumption and regeneration models, respectively. The truncation usually implies that the interaction between the factors is low and that the effects are mostly linear or quadratic.

The sparse PCE model sensitivity analysis resulted with Sobol' indices for first-order and total effects. Only factors which contributed more than 0.5% to total variation were included in the analysis. In Fig. 6, the six most significant noise factors contributing to the energy consumption variation are shown, with their respective error bars. The ambient temperature, rolling resistance and payload variation contributed 88.2% of the total variation in

<table>
<thead>
<tr>
<th>Metric</th>
<th>Consumption</th>
<th>Regeneration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full model evaluations</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>Full basis size</td>
<td>15504</td>
<td>15504</td>
</tr>
<tr>
<td>Maximal degree</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sparse basis size</td>
<td>118</td>
<td>155</td>
</tr>
<tr>
<td>Leave-one-out error</td>
<td>0.003</td>
<td>0.016</td>
</tr>
<tr>
<td>Mean value</td>
<td>1.20 kWh/km</td>
<td>28.0%</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.32 kWh/km</td>
<td>8.7%</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>26.9%</td>
<td>31.2%</td>
</tr>
</tbody>
</table>

Table 2
The metrics of the PCE surrogate models for consumption and regeneration. The leave-one-out errors indicate a sufficient fit with the results acquired with the original ECB model.

Fig. 6. Sobol’ indices of total (blue) and first-order effects (beige) of the six most significant noise factors contributing to the total energy consumption variation. The error bars of each of the factors are presented in red.
energy demand. The impact of these three factors to the total variation, with ECB and PCE model runs, is displayed in Fig. 7. The more the sample points are spread out at a specific factor value, the less it dominates the consumption variation. However, if the mean consumption of samples with lower factor value is different than with high values, the factor does contribute partially to the response variation.

Three of the most significant contributors to variation in energy regeneration were rolling resistance, battery state-of-charge and headwind, shown in Fig. 8. These factors account for 89.5% of the variation in the regenerated energy. The ECB and the PCE results of each test run for these factors is shown in Fig. 9. Interestingly, the rolling resistance affects the variation in energy demand quite linearly yet the influence on energy regeneration has slightly increasing exponential behavior towards low RRC values. The effect of headwind exhibits exponentially decreasing influence at negative wind speeds. In addition, the available energy regeneration is only depended on the battery state-of-charge when the charge is over 90%.

The Sobol’ metrics in Table 3 of both energy consumption and regeneration show that the residual of the analysis is smaller in the consumption analysis. The residual refers to the variation in energy demand that cannot be explained by the factors included in the analysis. Hence, the smaller the residual the more accurate the analysis. The error confidence interval of all factors, for both energy demand and regeneration, was 99%. The confidence interval is a measure of reliability, which is analyzed for each factor individually. The tolerance noise factors influence less than 0.1% to both system responses.

### Table 3

<table>
<thead>
<tr>
<th>Metric</th>
<th>Consumption</th>
<th>Regeneration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual</td>
<td>0.69%</td>
<td>6.05%</td>
</tr>
<tr>
<td>Extensive Factors</td>
<td>99.26%</td>
<td>93.85%</td>
</tr>
<tr>
<td>Tolerance Factors</td>
<td>0.05%</td>
<td>0.1%</td>
</tr>
<tr>
<td>Confidence Interval</td>
<td>99%</td>
<td>99%</td>
</tr>
</tbody>
</table>

4. Discussion

Our simulations with a lightweight bus had a variation in energy demand of 0.43–2.30 kWh/km on a single route. Previous studies report a simulated route-dependent variation in the energy demand of an electric city bus of 0.9–1.42 kWh/km [41] and 1.24–2.48 kWh/km [59] for light- and heavyweight bus structures, respectively. In our recent study, the energy consumption varied from 0.85 to 2.95 kWh/km in 15 routes, where the lowest consumption was observed in average conditions and the highest in worst-case conditions [60]. The weight, aerodynamical drag, rolling resistance and auxiliary power (including HVAC) were considered.

The main factor driving the variation in consumption was the ambient temperature. The form of the temperature-consumption relation (Fig. 7) is caused by HVAC’s temperature dependency presented in Fig. 4b. However, the temperature changes also affect aerodynamic drag, through variation in the air density. Even so, the authors estimate that well over 95% of the impact of the ambient temperature variation is due to the HVAC system, since the aerodynamic drag is one degree more dependent on wind speed than on air density.

Bottiglione et al. [61] report that air conditioning increases the fuel consumption of a hybrid electric bus by 142% in real-world conditions. As demonstrated in Fig. 4b, the power consumed by the heating power at extremely cold temperatures is greater than
that consumed by air conditioning during summertime in Finland, which indicates that our result is reasonable. Especially the heating system power is a challenge in electric vehicle and bus design, because of the lack of excess heat produced by an internal combustion engine.

The second most significant factor was the RRC and it was also the most significant factor influencing the variation in energy regeneration. However, the impact of RRC was linear to consumption yet slightly exponential to energy regeneration; changes in low RRC values increased the regenerated energy more than changes in high RRC values. The effect of a HW decreased slightly more with negative wind speeds. This behavior can be explained by the limitation of motor performance under heavier loads, which resulted in higher utilization of the mechanical brakes. Therefore, more energy is wasted when the motor reaches its performance limit during deceleration.

The RRC and HW also experienced sparse points of decreased regeneration with all values. This “snowflake effect” was caused by the 90% SoC regeneration limit set in the ECB model. The regeneration limit is set to prevent overcharging and overheating of the battery. Thus, if the SoC was 90% or higher at the beginning, at least a portion of the route was driven without regeneration. The battery SoC had no effect when the starting SoC was lower than 90% because the individual measurements align horizontally with respect to regeneration, as shown in Fig. 9. If the SoC limitation were excluded, the minimum regeneration on the route would have been 10%.

Furthermore, it is recommended that the charge-discharge cycles would circle around 50% SoC to increase the lifetime of the batteries [53,62]. However, too shallow depth of discharge (around 10% ΔSoC) might even expedite the degradation of LiFePO4 batteries [63]. Thus, the optimal depth of discharge needs to be further investigated for optimal battery dimensioning.

The small difference between the first-order and total effects of the energy demand (Fig. 6) implies that the interactions of the factors are low. This was expected, since the energy consumption is dominated by a few factors whose physics are more additive than multiplicative. In comparison, the high-order effects had some minor contribution to the variation in energy regeneration (Fig. 8). The minor high-order effects in all of the factors are caused by small correlations that even the most sophisticated quasi-random sequences such as the Sobol’ sequence cause.

More factor interaction is observed in the variation of energy regeneration (Table 3) because of the regenerative braking controller, which disables regeneration if the battery is too full. Furthermore, the amount of mechanical braking depends on different loads, which results in minor unknown variations in the model. It is also important to note that the variation in both the energy demand and regeneration variation was caused mainly by extensive noise factors rather than tolerance noise factors.

At the beginning of the study, there were challenges in the uncertainty propagation since it led to failure modes during the MCSs. The problem was resolved by a worst-case scenario analysis. Battery capacity was the main reason for failure modes in worst-case testing and is thus crucial in terms of reliable operation. On the other hand, oversizing the battery causes additional costs, which are around €300–€600 per kWh for passenger cars [64] and even higher for the lithium titanate (LTO) or lithium iron phosphate (LiFePO4) batteries often used in electric buses, which cost around €800 per kWh [16]. Optimal charging strategy can lengthen the lifetime of the batteries [53]. However, small battery and tight operation schedules, which are intuitively desired, increase power demand costs and higher load on the electric grid [65]. Therefore, in future works the optimization of the battery capacity and charging station positioning under uncertain driving conditions when schedule abnormalities occur is imperative.

The battery dimensioning can be carried out with the developed surrogate model that is 10 000 more computationally efficient than the original ECB model (Fig. 1). The PCE surrogate model is of close approximation to the ECB model (Fig. 5), which is confirmed by over 99% confidence interval (Table 3). The minimum number of iterations to achieve the desired accuracy was determined empirically by testing the model fit with after every 100 MCSs. However, in some cases the model error saturates before the desired accuracy is achieved. Model error saturation refers to a situation where the increase of MCS iterations no longer increases accuracy. Furthermore, the relation between the energy demand and regeneration distributions was as expected, since they were skewed in opposite directions: higher regeneration was more probable and so was lower consumption.

The results of the study were in line with the expectations. However, every study has its limitations. As mentioned earlier, the selected input factors do not account for all of the variation in the outputs. Therefore, the analysis will be more complete when various mission-related factors such as traffic and the driving cycle are considered. However, such dynamic factors do not have theoretical worst-case values, and thus cannot be analyzed independently, like the vehicle-related factors analyzed here. Furthermore, the developed PCE model is intended for a specific purpose and it behaves as expected in its defined input space. Extrapolation performance was not studied. In addition, the presented model only considers one route and should be retrained for other routes.

5. Conclusion

The energy demand sensitivity of an electric city bus was successfully investigated with a novel, fast-computing surrogate modeling approach. First, the uncertainty in 14 input factors was identified on the basis of literature and measurements. The margin identification and worst-case analysis were imperative to ensure realistic results and to avoid failure modes in simulations. Then, the input factor uncertainty was propagated using quasi-randomized Monte Carlo simulations with a previously developed electromechanical electric city bus model. The simulated data was used to develop a surrogate model to predict energy demand and regeneration 10 000 times faster than the original model. The surrogate model was compared to the original ECB model with a leave-one-out cross-validation that resulted in less than 1% error.

In addition to factor identification and surrogate model development, the sensitivity of energy demand and regeneration under uncertain conditions was analyzed. The mean energy demand was 1.20 kWh/km with a standard deviation of 0.32 kWh/km and on average 28% of the kinetic energy was recovered through regenerative braking. The ambient temperature, rolling resistance and payload uncertainty contributed most to the variation in energy demand. Extreme ambient temperatures require more power from the HVAC system, which by itself accounted for more than half of the variation in energy demand. The variation in energy regeneration was determined by rolling resistance, battery state-of-charge and headwind. The results indicate that the HVAC power, tire, road and weather variations require further studies to reduce the variation in energy demand.

In future work, the surrogate model presented here will be validated with real-world measurements. The validated surrogate model can be employed to optimize battery size, predict energy demand variation in real-time and to plan charging system capacity utilization. The approach presented here provides an understanding of the factors that drive electric vehicle energy consumption. This understanding of specific cases is crucial, since the
competitiveness of electric vehicles is limited because of range anxiety. Moreover, in the case of electric city buses, range anxiety can lead to battery over-sizing, which unnecessarily increases investment costs and the gross weight of electric buses.

Acknowledgments

The authors thank the Finnish Meteorological Institute for providing open weather data and Sami Ruotsalainen from Linkker for the electric city bus data. The authors also acknowledge the financial support from KAUTE, Honkanen and Henry Ford Foundations for the research reported in this paper.

Nomenclature

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUX</td>
<td>Auxiliary power</td>
</tr>
<tr>
<td>BR</td>
<td>Battery internal resistance</td>
</tr>
<tr>
<td>BAge</td>
<td>Battery aging</td>
</tr>
<tr>
<td>BC</td>
<td>Battery transient behavior capacitance</td>
</tr>
<tr>
<td>BQ</td>
<td>Battery capacity</td>
</tr>
<tr>
<td>ECB</td>
<td>Electric City Bus –model</td>
</tr>
<tr>
<td>GSA</td>
<td>Global Sensitivity Analysis</td>
</tr>
<tr>
<td>HVAC</td>
<td>Heating, Ventilation and Air Conditioning</td>
</tr>
<tr>
<td>HW</td>
<td>Headwind</td>
</tr>
<tr>
<td>LAR</td>
<td>Least Angle Regression</td>
</tr>
<tr>
<td>LOOCV</td>
<td>Leave-one-out Cross-Validation</td>
</tr>
<tr>
<td>LSA</td>
<td>Local Sensitivity Analysis</td>
</tr>
<tr>
<td>MASS</td>
<td>Mass of the bus</td>
</tr>
<tr>
<td>MCS</td>
<td>Monte Carlo Simulation</td>
</tr>
<tr>
<td>MF</td>
<td>Flux induced by motor magnets</td>
</tr>
<tr>
<td>MI</td>
<td>Inductance of the motor</td>
</tr>
<tr>
<td>MR</td>
<td>Internal resistance of the motor</td>
</tr>
<tr>
<td>NF</td>
<td>Noise Factor</td>
</tr>
<tr>
<td>PCE</td>
<td>Polynomial Chaos Expansion</td>
</tr>
<tr>
<td>PTA</td>
<td>Public Transport Authority</td>
</tr>
<tr>
<td>RRC</td>
<td>Rolling Resistance Coefficient</td>
</tr>
<tr>
<td>SA</td>
<td>Sensitivity Analysis</td>
</tr>
<tr>
<td>SI</td>
<td>Sensitivity Index</td>
</tr>
<tr>
<td>SoC</td>
<td>Battery state-of-charge</td>
</tr>
<tr>
<td>Ta</td>
<td>Ambient Temperature</td>
</tr>
<tr>
<td>Tb</td>
<td>Temperature of the battery pack</td>
</tr>
<tr>
<td>TSM</td>
<td>Temporary Surrogate Model</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Maximal Polynomial degree</td>
</tr>
<tr>
<td>M</td>
<td>Model</td>
</tr>
<tr>
<td>M</td>
<td>Model approximation</td>
</tr>
<tr>
<td>N</td>
<td>Number of simulated observations</td>
</tr>
<tr>
<td>Y</td>
<td>Response/Output</td>
</tr>
<tr>
<td>X</td>
<td>Input factor matrix</td>
</tr>
<tr>
<td>$\mathbf{v}_a$</td>
<td>Multivariate coefficient</td>
</tr>
<tr>
<td>$\varepsilon_{LOO}$</td>
<td>Leave-one-out error</td>
</tr>
<tr>
<td>$\mathbf{\psi}_a$</td>
<td>Multivariate polynomial</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Sample Mean</td>
</tr>
<tr>
<td>$S_i$</td>
<td>Sensitivity index, first-order effect</td>
</tr>
<tr>
<td>$S_{\eta_i}$</td>
<td>Sensitivity index, total effect</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscripts</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>Demand (of energy)</td>
</tr>
<tr>
<td>R</td>
<td>Regeneration (of energy)</td>
</tr>
<tr>
<td>Y</td>
<td>Response, Output</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Superscripts</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>Single column of the matrix, representing the observations of one factor</td>
</tr>
<tr>
<td>$\sim i$</td>
<td>All other columns of the matrix, except i</td>
</tr>
<tr>
<td>$r$</td>
<td>Individual iteration</td>
</tr>
<tr>
<td>PCE</td>
<td>Polynomial Chaos Expansion</td>
</tr>
<tr>
<td>PCE/$r$</td>
<td>Polynomial Chaos Expansion model iteration</td>
</tr>
</tbody>
</table>

Appendix A

The ECB model has a virtual driver, which is given a measured velocity profile. The speed and current controllers adjust the motor power so that the desired driving behavior is achieved. The motor is coupled to the tires with differential gear and a fixed ratio gear, which produces tractive force $F_x$ according to

$$m_a \cdot \frac{dv_x}{dt} = \sum F_x - F_{res} - m_a g \sin \theta$$

where $m_a$ is the mass of vehicle

$F_{res}$ are the resistive forces

$F_r$ is the rolling resistance force

$F_d$ is the aerodynamic drag force

$\rho_a$ is the air density

$T_a$ is the ambient temperature

$R_a$ is the specific gas constant

The battery model is described with

$$u_b = u_{oc} - R_{int}i_b$$

$$u_{oc} = \sum_{k=0}^{n} (c_k \times (1 - SOC)^k)$$

$$SOC = 1 - \frac{1}{Q_r} \cdot \int_{0}^{t} (i_b) \, dt$$

where

$u_b$ is the battery equilibrium potential

$u_{oc}$ is the open-circuit voltage

$i_b$ is the battery current

$R_{int}$ is the internal resistance

$Q_r$ is the coulombic capacity of the battery

SOC is a nonlinear approximation of the voltage as a function of state-of-charge
Cₖ is the kth polynomial of SOC as it is fitted to a measured LiFePO₄ cell behavior.

References


[29] Li, K et al. Exploring the interactive effects of ambient temperature and vehicle auxiliary loads on electric vehicle energy consumption. Appl Energy 2017;0–1. https://doi.org/10.1016/j.apenergy.2017.08.074. Elsevier Ltd.


