DTU Library

Inhomogeneous Markov Models for Describing Driving Patterns

Iversen, Jan Emil Banning; Møller, Jan Kloppenborg; Morales González, Juan Miguel; Madsen, Henrik

Link back to DTU Orbit

Inhomogeneous Markov Models for Describing Driving Patterns

Emil B. Iversen[a,*], Jan K. Møller[a], Juan M. Morales[a], Henrik Madsen[a]

[a]*DTU Informatics, Technical University of Denmark, Asmussens Alle, building 305, DK-2800 Lyngby, Denmark*

Abstract

It has been predicted that electric vehicles will play a crucial role in incorporating a large renewable component in the energy sector. If electric vehicles are integrated in a naive way, they may exacerbate issues related to peak demand and transmission capacity limits while not reducing polluting emissions. Optimizing the charging of electric vehicles is paramount for their successful integration. This paper presents a model to describe the driving patterns of electric vehicles, in order to provide primary input information to any mathematical programming model for optimal charging. Specifically, an *inhomogeneous Markov model* that captures the diurnal variation in the use of a vehicle is presented. The model is defined by the time-varying probabilities of starting and ending a trip and is justified due to the uncertainty associated with the use of the vehicle. The model is fitted to data collected from the actual utilization of a vehicle. Inhomogeneous Markov models imply a large number of parameters. The number of parameters in the proposed model is reduced using B-splines.

Keywords:
Driving patterns, Inhomogeneous Markov chain, B-splines, Electric vehicles

1. Introduction

Electric vehicles (EVs) have no emissions and are a sustainable alternative to conventional vehicles, provided that the energy used for charging is generated by renewable sources. Production levels from renewable energy sources, such as wind, solar and wave energy depend on weather conditions and consequently there is a high degree of variability in power generation. With the current absence of large-scale energy storage, today electricity has to be produced and

---

*Corresponding author. Tel.: +45 45 25 33 75; Fax: +45 45 88 26 73

Email addresses: `jebi@imm.dtu.dk` (Emil B. Iversen), `jkm@imm.dtu.dk` (Jan K. Møller), `jmmgo@imm.dtu.dk` (Juan M. Morales), `hm@imm.dtu.dk` (Henrik Madsen)

consumed at the same time. With electricity coming from renewable sources it is not possible to produce additional power, if weather conditions do not allow for it. Moreover, in times of high availability from renewable sources, the demand for power may be low and the economic potential of renewables is thus wasted. The battery in an EV is basically a storage device for energy and has the potential to help overcome some of the issues regarding the large-scale introduction of renewable energy. This is done by charging the EVs when energy from renewable sources is abundant and by supplying power into the electrical grid at times of high demand. As long as electric vehicles are charged with electricity from renewable sources, they represent a sustainable zero-emissions alternative to conventional fossil-fuel-based vehicles. However, if EVs are charged in a naive way, they may increase the peak electricity demand. As a consequence, the extra energy needs would be covered by peak-supply units, typically fossil-fuel-based, which would nullify the decrease in emissions gained by switching from conventional to electric vehicles. Furthermore, an increased peak demand could lead to a shortage of transmission capacity, which would force an expansion of the electrical grid to handle the higher peak demand. This is costly and undesirable. To avoid these problems, EVs should be charged in a smart fashion.

As EVs are primarily used for transportation, and not for energy storage, it is essential to charge each vehicle such that there is enough energy to cover any desired trip. Hence a decision-support tool is required to determine whether it is possible to postpone charging the EV or whether it should be charged right away. For such a tool to produce optimal charging decisions, a model capturing the utilization of a specific vehicle is essential. The complexity of human behavior points to a stochastic model to adequately describe the driving needs of EV users.

In the technical literature, however, the usual approach is to define a deterministic driving scenario based on expected values and averages. Needless to say, such an approach fails to capture the dynamics and stochasticity in the use of a specific vehicle. Observed vehicle usage has been considered in several studies (Pearre and Elango, 2011; Golob, 1998). An issue that has received little attention is indeed the stochastic modeling of driving patterns (Green et al., 2010). Rather, the scientific community has been more focused on both the analysis of the potential impact of charging EVs and the design of models to decide when to charge (Rotering and Ilic, 2011). In this vein, the effect of the large-scale integration of EVs into the power grid has been studied in several papers, (Gan et al., 2011; Ma et al., 2010; Lojowska et al., 2011; Acha et al., 2010; Richardson et al., 2012). Issues such as peak load, different charging strategies, network losses, minimizing costs and market equilibrium strategies have been considered.

In this paper we propose a stochastic model for the use of a vehicle. The model can be easily exploited, for example, by decision-making tools for charging an EV. Furthermore, our model provides advances in modeling driving patterns and does not rely on the typical, average or stylized use of a vehicle. The model is fitted to a specific vehicle based on observed data from the utilization of that vehicle. An *inhomogeneous Markov model* is applied to capture the

diurnal variation of the driving pattern. A major disadvantage of these types of models is the high number of parameters to be estimated. B-splines are then applied to substantially reduce this number. An algorithm is proposed to place knots and to find the appropriate number of knots needed for the B-splines. The proposed model does not rely on any assumptions regarding the use of the vehicle, and consequently a versatile model is obtained. Applying the model within a stochastic programming framework will allow for capturing issues related to charging, availability, and costs of using an EV. The proposed approach can be easily extended to a mixed-effects model in order to capture the dynamics of a population of vehicles. A somewhat similar approach was applied in Madsen and Thyregod (1986) to the problem of modeling diurnal variation in cloud cover.

The paper is organized as follows: Section 2 gives a brief introduction to inhomogeneous Markov chains. In Section 3 the number of parameters in the model is reduced by applying B-splines to a generalized linear model. Section 4 provides a numerical example of the model, where the parameters are fitted to observed data from a single vehicle. Section 5 concludes and provides directions for future research within this topic.

2. An Inhomogeneous Markov Chain

A state-space approach is proposed to describe the use of a vehicle. This approach models the vehicle as being in one of several distinct states. In its simplest form the model has two states, which capture whether the vehicle is either *driving* or *not driving*. A more extensive model may include information about where the vehicle is parked, where it is driving, or what type of trip the vehicle is on. In this section we start from a general state-space approach and finish with a detailed description of the two-state model.

2.1. Discrete Time

Let X_t, where $t \in \{0, 1, 2, \ldots\}$, be a sequence of random variables which take on values in the countable set S, called the state space. Denote this sequence X. Without loss of generality, we assume that the state space includes N states. A Markov chain is a random process where future states, conditioned on the present state, do not depend on the past states (Grimmett and Stirzaker, 2001). In discrete time, $\{X\}$, is a Markov chain if

$$\mathbb{P}\left(X_{t+1} = k | X_0 = x_0, \ldots, X_t = x_t\right) = \mathbb{P}\left(X_{t+1} = k | X_t = x_t\right) \quad (1)$$

for all $t \geq 0$ and all $\{k, x_0, \ldots, x_t\} \in S$.

A Markov chain is uniquely characterized by the transition probabilities, i.e.

$$p_{jk}(t) = \mathbb{P}\left(X_{t+1} = k | X_t = j\right). \quad (2)$$

If the transition probabilities do not depend on t, the process is called a homogeneous Markov chain. If the transition probabilities depend on t, the process is known as an inhomogeneous Markov chain.

Considering the use of a vehicle, it is reasonable to expect that the probability of a transition from state j to state k at any specific weekday is the same. Thus the transition probabilities on Tuesday in one week are assumed to be the same as on Tuesdays in other weeks. Furthermore, it is natural to assume that the transition probabilities are the same on all weekdays, that is, from Mondays to Fridays. If the sampling time is in minutes, and as there are 1440 minutes in a day, this leads to the assumption:

$$p_{jk}(t) = p_{jk}(t + 1440). \tag{3}$$

This implies that the transition probabilities, defined by (2) are constrained to be a function of the time s in the diurnal cycle. The matrix containing the transition probabilities is given by

$$\mathbf{P}(s) = \begin{pmatrix} p_{11}(s) & p_{12}(s) & \ldots & p_{1N}(s) \\ p_{21}(s) & p_{22}(s) & \ldots & p_{2N}(s) \\ \vdots & \vdots & \ddots & \vdots \\ p_{N1}(s) & p_{N2}(s) & \ldots & p_{NN}(s) \end{pmatrix}, \tag{4}$$

where $p_{jj}(s) = 1 - \sum_{i=1, i \neq j}^{N} p_{ji}$.

If the model is formulated with time resolution in minutes, $s \in \{1, 2, \ldots, 1440\}$. The assumed periodicity from (3) implies that all the observations from different days are lumped together and a transition is only denoted by its time of day. It follows that the conditional likelihood function, for the model with N states, is given by (Pawitan, 2001):

$$L\left(\mathbf{P}(1), \mathbf{P}(2), \ldots, \mathbf{P}(1440)\right) = \prod_{s=1}^{1440} \prod_{j=1}^{N} \prod_{k=1}^{N} p_{jk}(s)^{n_{jk}(s)}, \tag{5}$$

where $n_{jk}(s)$ is the number of observed transitions from state j at time s to state k at time $s + 1$, where s is the time in minutes of the diurnal cycle.

From the conditional likelihood function the maximum-likelihood estimate of $p_{jk}(s)$ can be found as:

$$\widehat{p}_{jk}(s) = \frac{n_{jk}(s)}{\sum_{k=1}^{N} n_{jk}(s)}. \tag{6}$$

A discrete time Markov model can be formulated based on the estimates of $\mathbf{P}(1), \mathbf{P}(2), \ldots, \mathbf{P}(1440)$. One apparent disadvantage of such a discrete time model is the huge number of parameters, namely $N \cdot (N - 1) \cdot 1440$, where $N \cdot (N - 1)$ parameters have to be estimated for each time step. Needless to say, the number of parameters to be estimated increases as the number of states increases. Another problem is linked to the number of observations, i.e. if $\sum_{k=1}^{N} n_{jk}(s') = 0$ for some s', then $\widehat{p}_{jk}(s)$ is undefined.

A reduction in parameters may be obtained if the diurnal variation is negligible for some transitions, i.e. $p_{jk}(s)$ does not depend on s for some pair $\{j, k\}$.

One way to reduce the parameters is to increase the time between samples. If the sampling time is every 10 minutes, the number of parameters would decrease to $N \cdot (N-1) \cdot 144$. This approach is a bit coarse and the number of parameters is still large. Besides, if another parameter reduction technique is subsequently applied to the data, information is lost compared to directly applying the technique to the data with a sampling time in minutes.

In the model with only two states, namely *driving* and *not driving*, the one-minute transition probability matrix becomes:

$$\mathbf{P}(s) = \left(\begin{array}{cc} p_{11}(s) & p_{12}(s) \\ p_{21}(s) & p_{22}(s) \end{array} \right) = \left(\begin{array}{cc} 1 - p_{12}(s) & p_{12}(s) \\ p_{21}(s) & 1 - p_{21}(s) \end{array} \right). \tag{7}$$

The number of parameters is $2 \cdot 1440$. Assuming that the duration of the trip does not depend on the time of the day, i.e. $p_{21}(s) = p_{21}$, (with 2 being "*driving*" and 1 "*not driving*") the number of parameters is reduced to $1440+1$. Note that, as a result of this reduction, the duration of a trip is captured by a single parameter.

It follows that the conditional likelihood function, for the model with two states, is given by:

$$L\left(\mathbf{P}(1), \mathbf{P}(2), \ldots, \mathbf{P}(1440)\right) = \prod_{s=1}^{1440} \prod_{j=1}^{2} \prod_{k=1}^{2} p_{jk}(s)^{n_{jk}(s)}, \tag{8}$$

and the maximum-likelihood estimate $\hat{p}_{jk}(s)$ is computed from (6).

2.2. Continuous Time

The continuous time analog to the discrete time inhomogeneous Markov chain is presented below. The continuous time version provides a parameter reduction over the discrete time version, if certain structures are present and can be identified. Specifically, if the number of states is larger than two and it is impossible to switch directly between certain pairs of states, the continuous time variant will lead to a parameter reduction. Hence, if such structures are present, the continuous time variant is preferred over the discrete time model. To introduce the continuous time inhomogeneous Markov chain, we define (Grimmett and Stirzaker, 2001):

$$p_{jk}(t, u) = \mathbb{P}(X(u) = k | X(t) = j), \tag{9}$$

where $t < u$. The model is based on the following assumptions when $\Delta u \to 0$:

$$p_{jj}(u, u + \Delta u) = 1 - q_{jj}(u)\Delta u + o(\Delta u) \tag{10}$$

$$p_{jk}(u, u + \Delta u) = q_{jk}(u)\Delta u + o(\Delta u) \qquad \forall \quad j \neq k, \tag{11}$$

also $0 \leq q_{jj}(u) < \infty$ and $0 \leq q_{jk}(u) < \infty$. The $q_{jk}(u)$'s are known as the transition intensities. These assumptions lead to Kolmogorov's forward differential equation for inhomogeneous Markov processes, expressed in matrix notation as:

$$\frac{\partial \mathbf{P}(t, u)}{\partial u} = \mathbf{P}(t, u)\mathbf{Q}(u) \tag{12}$$

where $\mathbf{P}(t, u) = \{p_{jk}(t, u)\}$, i.e. $\mathbf{P}(t, u)$ is the matrix containing the $p_{jk}(t, u)$'s. The matrix of transition intensities then becomes:

$$\mathbf{Q}(u) = \begin{pmatrix} -q_{11}(u) & q_{12}(u) & \ldots & q_{1N}(u) \\ q_{21}(u) & -q_{22}(u) & \ldots & q_{2N}(u) \\ \vdots & \vdots & \ddots & \vdots \\ q_{N1}(u) & q_{N2}(u) & \ldots & -q_{NN}(u) \end{pmatrix}. \tag{13}$$

Since $\sum_{k=1}^{N} p_{jk}(u, u + \Delta u) = 1$, it follows from (10)-(11) that $\sum_{k=1}^{N} q_{jk}(u) = 0$, i.e. $q_{jj}(u) = \sum_{k=1,k\neq j}^{N} q_{jk}(u)$.

A simple Kolmogorov's differential equation is obtained if $\mathbf{Q}(t)$ is constant in the period $[t, t + T]$:

$$\mathbf{P}(t, t + T) = e^{\mathbf{Q}(t)T}\mathbf{P}(t, t) = e^{\mathbf{Q}(t)T}. \tag{14}$$

Suppose that $T = 1$. Then the one minute transition probabilities are given by:

$$\mathbf{P}(t, t + 1) = \mathbf{P}(t) = e^{\mathbf{Q}(t)}. \tag{15}$$

If the model has two states, the matrix of transition intensities becomes:

$$\mathbf{Q}(u) = \begin{pmatrix} -q_{11}(u) & q_{12}(u) \\ q_{21}(u) & -q_{22}(u) \end{pmatrix} = \begin{pmatrix} -q_{12}(u) & q_{12}(u) \\ q_{21}(u) & -q_{21}(u) \end{pmatrix} \tag{16}$$

As mentioned previously, a continuous time Markov chain will allow for a parameter reduction if certain structures are present. Furthermore, identifying such structures will make the model more theoretically tractable.

As a simple illustration of such a model, consider the case where there are four states, i.e. $N = 4$. State 1 corresponds to the vehicle being parked at home. State 2 corresponds to the vehicle being on a trip that started from home. State 3 corresponds to the vehicle being parked somewhere else. State 4 corresponds to the vehicle starting a trip from somewhere else than at home. The parameter reduction is thus obtained if it is assumed that the vehicle cannot switch directly from being parked at home to being parked somewhere else, that is from states 1 to 3. Also it would be reasonable to assume that the vehicle does not drive from home to return to home, without an intermediate stop. Under these assumptions, the matrix of transition intensities becomes:

$$\mathbf{Q}(u) = \begin{pmatrix} -q_{12}(u) & q_{12}(u) & 0 & 0 \\ 0 & -q_{23}(u) & q_{23}(u) & 0 \\ 0 & 0 & -q_{34}(u) & q_{34}(u) \\ q_{41}(u) & 0 & q_{43}(u) & -(q_{43}(u) + q_{41}(u)) \end{pmatrix}. \tag{17}$$

The discrete time transition probability matrix can then be found by (15). In this case the number of parameters to be estimated for each time step is reduced from $N \cdot (N - 1) = 12$ to 5, by formulating the model in continuous time as

6

opposed to discrete time. The idea behind this specific model is that it can capture whether the vehicle is parked for different lengths of time, depending on the location. Also it can capture whether the vehicle is usually parked at home at night. As the number of states in the model increases, and supposing that certain structures can be identified, the parameter reduction gained by formulating the model in continuous time is increased.

3. Parameter Reduction via B-Splines

As the number of parameters to be estimated is huge, techniques to reduce this number are needed. One such a technique is applying B-splines to approximate the diurnal variation. Other techniques include smoothing splines and kernels, and these could also be applied. The choice is not straightforward. B-splines are preferred here as it is simple to work with basis functions, while the B-spline still provides a spline of the desired order. For a thorough introduction to B-splines as well as smoothing splines and kernels, see Hastie et al. (2008).

3.1. B-Splines

To construct a B-spline, first define the knot sequence τ such that

$$\tau_1 \leq \tau_2 \leq \cdots \leq \tau_M. \tag{18}$$

Let this sequence of knots be defined on the interval where we wish to evaluate our spline. In this particular case the knots should be placed somewhere in the interval $[0, 1440]$, that is over the day.

Denote by $B_{i,m}(x)$ the ith B-spline basis function of order m for the knot sequence τ, where $m < M$. The basis functions are defined recursively as follows:

$$B_{i,1}(x) = \begin{cases} 1 & \text{if } \tau_i \leq x < \tau_{i+1} \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

$$B_{i,m}(x) = \frac{x - \tau_i}{\tau_{i+m-1} - \tau_i} B_{i,m-1}(x) + \frac{\tau_i - x}{\tau_{i+m} - \tau_{i+1}} B_{i,m-1}(x) \tag{20}$$

for $i = 1, \ldots, M - m$. These basis functions are polynomials of order $m - 1$ taking values on the interval $[\tau_1, \tau_M]$.

A B-spline curve of degree m is a piecewise polynomial curve defined as follows:

$$S_m(x) = \sum_{i=1}^{M-m} C_i B_{i,m}(x), \tag{21}$$

where C_i, $i = \{1, \ldots, M - m\}$, form the *control polygon*. The $B_{i,m}(x)$ are the B-spline basis functions of order m defined over the knot vector.

As we aim at modeling the diurnal variation in the driving pattern, it is reasonable that the basis splines are periodic. This can be achieved by introducing $2m$ new knots to the existing knots. The new knots are defined as follows:

$$\tau_{1-h} = \tau_{M-h} - (\tau_M - \tau_1) \quad \text{for } h \in \{1, \ldots, m\} \tag{22}$$

$$\tau_{M+h} = \tau_h + (\tau_M - \tau_1) \quad \text{for } h \in \{1, \ldots, m\}. \tag{23}$$

7

More specifically, let the vector containing the new knots be represented by $\tau' = \{\tau_{1-m}, \ldots, \tau_{M+m}\}$. For each B-spline basis function $m + 1$ knots are required, though they may be overlapping. The B-spline basis functions are uniquely defined by the position of the knots. In particular, if the knots are shifted by some constant α, the basis functions will be the same as the original, except that they are shifted by α. If the new knot vector is defined as τ', the basis function defined by the knots $\{\tau_M, \ldots, \tau_{M+m}\}$ will be the same as that defined for the knots $\{\tau_{1-m}, \ldots, \tau_1\}$, except that it is shifted by the interval length $\tau_M - \tau_1$. In this way we can define a basis function that is harmonic in the sense that it is recurrent over different days.

All piecewise polynomial splines of order m defined over the knot vector τ can be constructed from the basis functions defined in (19)-(20). Hence using B-splines does not limit the choice of polynomial splines in any way. Nonetheless an advantage of using B-splines is that the desired spline can be written as a linear combination of predefined basis functions. This proves useful as a generalized linear model is applied to estimate the transition probabilities. Traditionally cubic B-splines are used, i.e. $m = 4$, which is also the case here. A motivation for using cubic B-splines is that the spline produced will be of order 4 and furthermore, if $\tau_i \neq \tau_j$ for all $i \neq j$, it will be $\mathcal{C}^2$ everywhere. A function which is $\mathcal{C}^2$ is indistinguishable from a $\mathcal{C}^\infty$ to the human eye. For a further discussion on why to choose cubic splines, see Hastie et al. (2008).

3.2. A Generalized Linear Model

To reduce the number of parameters in the model, a B-spline can be fitted to the time-varying transition probabilities $p_{jk}(s)$. There are, however, some issues with this approach. Firstly, there is no guarantee that the fitted B-spline is always in the interval $[0, 1]$, which is a problem as we are modeling probabilities. Secondly, if $\sum_{k=1}^{N} n_{jk}(s) = 0$ for some s, the estimate for $p_{jk}(s)$ given by (6) is undefined. A more refined approach is to use a generalized linear model instead. In the following, such an approach is outlined.

Each day, at a specific minute, a transition from state j to state k either occurs or does not occur. Thus for every s on the diurnal cycle we can consider the number of transitions to be binomially distributed, i.e. $n_{jk}(s) \sim B(z_j(s), p_{jk}(s))$, where the number of Bernoulli trials at s, given by $z_j(s) = \sum_{k=1}^{N} n_{jk}(s)$, are known and the probabilities of success, $p_{jk}(s)$, are unknown. The data can now be analyzed using a *logistic regression*, which is a generalized linear model. The explanatory variables in this model are taken to be the basis functions for the B-spline. The logit transformation of the odds of the unknown binomial probabilities are modeled as linear functions of the basis functions $B_{i,m}(s)$. We model $Y_{jk}(s) = n_{jk}(s)/z_j(s)$ and in particular we are interested in $\mathrm{E}[Y_{jk}(s)] = p_{jk}(s)$.

Next we elaborate on how the logistic regression works in this particular case. For a general treatment of this problem see (Madsen and Thyregod, 2010).

We shall use a linear model for a function of p, the *link function*. The

8

canonical link for the binomial distribution is the *logit transformation*,

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right), \tag{24}$$

which is used as the link function. The resulting transformed means are given by $\eta_{jk}(s)$, which is modeled using a *linear model* with the B-spline basis functions as explanatory variables:

$$\eta_{jk}(s) = \log\left(\frac{p_{jk}(s)}{1-p_{jk}(s)}\right) = C_{jk,1} \cdot B_{1,4}(s) + \ldots + C_{jk,M} \cdot B_{M,4}(s). \tag{25}$$

The linear prediction of $\eta_{jk}(s)$ is therefore given by

$$\widehat{\eta}_{jk}(s) = \widehat{C}_{jk,1} \cdot B_{1,4}(s) + \ldots + \widehat{C}_{jk,M} \cdot B_{M,4}(s), \tag{26}$$

where the estimates, $\widehat{C}_{jk,1}, \ldots, \widehat{C}_{jk,M}$, are found by the *iteratively reweighted least squares method*. The inverse transformation of the link function in (25), which provides the probabilities of a transition from state j to state k at time s, is the *logistic function*

$$p_{jk}(s) = \frac{\exp(\eta_{jk}(s))}{1 + \exp(\eta_{jk}(s))}. \tag{27}$$

The estimates of the transition probabilities are thus given by

$$\widehat{p}_{jk}(s) = \frac{\exp\left(\widehat{\eta}_{jk}(s)\right)}{1 + \exp\left(\widehat{\eta}_{jk}(s)\right)} \tag{28}$$

The procedure of applying a *generalized linear model* is implemented in the statistical software package R as the function $\texttt{glm}(\cdot)$.

3.3. Choosing the Knots

Choosing the amount and position of the knots in the knot vector τ is important to obtain a good fit for the model. A naive method for placing the knots is to distribute them uniformly over the day. A uniform positioning, however, does not take into account the peakedness of the estimate of $p_{jk}(s)$. An algorithm for placing the knots is given in Mao and Zhao (2003). This method requires optimization; it is computationally intensive, does not guarantee a global optimum, and is based on normally distributed errors. Instead an algorithm is proposed here according to which the knots are placed where there are many transitions. This method does not require optimization, but does manage to make the model more sensitive where there are many transitions. As can be observed from the example in Section 4, this approach works well in practice.

Note that the problem of choosing knots in our B-spline-based approach is equivalent to that of choosing a (time-varying) smoothing parameter or bandwidth in the smoothing spline or kernel approaches.

The proposed algorithm for placing the knots runs as follows:

1. Decide first on the total number of knots, M.
2. Decide next on an initial number of knots, $M_{init} < M$, to be dispersed uniformly in the interval, with one at each endpoint. Denote these knots by τ_{init}.
3. Find the two adjacent knots with the highest amount of transitions in the interval between them. Denote these knots τ_j, τ_{j+1}.
4. Place a new knot, $\tau^\star$, in the middle of the interval (τ_j, τ_{j+1}).
5. Go to step 3 if the new number of knots $M^\star < M$. If $M^\star = M$ then stop.

Once an algorithm for distributing the knots is in place, the number of knots to choose, M, has to be decided. If the number of knots is too low, the model can be improved by placing additional knots. If the amount of knots is too high, the model is overparameterized. The amount of knots is found by testing whether placing an additional knot significantly improves the model. This approach is described in the following.

Consider two different models, with a number of knots, M_1 and M_2, respectively, and the same number of initial knots M_{init}. Let $M_1 < M_2$, then model M_1 is a sub-model of model M_2. In this case, using a log-likelihood ratio test is appropriate for distinguishing between the models. If we furthermore consider a host of different models $M_1, \ldots, M_N$ all with the same M_{init} and suppose that $M_{n-1} = M_n - 1$ for $n \in \{2, \ldots, N\}$, then the testing can be performed recursively. That being so, the test statistic is given by:

$$D_n = -2 \cdot \log\left(\frac{L\left(\hat{p}_{ik,M_{n-1}}(s)\right)}{L\left(\hat{p}_{ik,M_n}(s)\right)}\right), \tag{29}$$

which is drawn from a χ^2-distribution with one degree of freedom.

Let us conclude this section with a side note on recursive testing. Suppose there are three models with a number of knots $M_1 < M_2 < M_3$, where the former are sub-models of the latter. Assume that they have the same initial knots. Suppose that no significant improvement is found between models M_1 and M_2. This does not exclude the possibility of a significant improvement between models M_2 and M_3. We recommend testing different models recursively up to some large M, and choosing the number of knots where there does not seem to be any significant improvement beyond this point.

4. Numerical Example

In this section the model is fitted to a sample of data observed from utilization of a single vehicle. The data only allows for fitting a two-state model, as the data set solely contains information on whether the vehicle is *driving* or *not driving*. We focus specifically on how the proposed procedure works for estimating $\widehat{p}_{12}(s)$, as the estimation of other entries in $\mathbf{P}(\mathbf{s})$ is done similarly. Also $\widehat{p}_{12}(s)$ might be the most interesting parameter, as it is the probability of starting a trip within the next minute, conditional on the vehicle not driving at time s.

4.1. Data

The example is based on data pertaining to a single vehicle in Denmark in the period spanning the six months from 23-10-2002 to 24-04-2003, with a total of 183 days. The data is GPS-based and follows specific cars. One car has been chosen and the model is intended to describe the use of this vehicle accordingly. The data set only contains information on whether the vehicle was *driving* or *not driving* at any given time. No other information was provided in order to protect the privacy of the vehicle owner. The data suggest that a model with two states is appropriate. The data set consists of a total of 749 trips. The time resolution is in minutes.
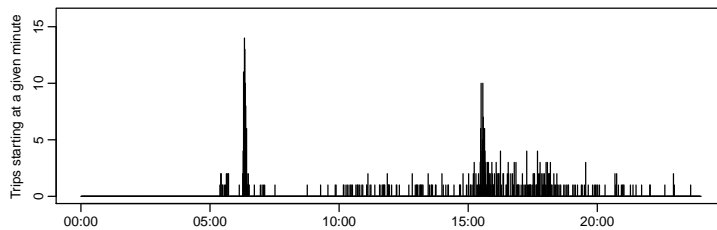
Figure 1: *Trips starting at a certain minute of the day, cumulated for 131 weekdays.*

The data have been divided into two main periods, weekdays and weekends. The observed number of trips starting every minute for the weekdays is displayed in Figure 1. A high degree of diurnal variation is found, with a lot of trips starting around 06:00 and again around 16:00. Also there are no observations of trips starting between 00:00 and 05:00. Other patterns are found for weekends, but as the approach is similar, the focus is on trips starting on weekdays. Annual variations may also be present, however the limited data sample does not allow for capturing such a variation.

4.2. Estimation

Firstly, naive B-splines have been fitted to the data using the logistic regression and the result is shown in Figure 2. These B-splines are described as naive in the sense that the knots defining the basis functions for the B-splines are placed uniformly over the interval. The gray lines are the estimates $\widehat{p}_{12}(s)$ found from (6). As the number of basis-functions increases it is apparent that the fit improves.

The algorithm for placing the knots is implemented considering an initial amount of knots $M_{init} = 10$. The left plot in Figure 3 shows the recursive tests, with the test statistic given by (29), between the subsequent models with the number of knots, M_n.

Referring to Figure 3, left, the model with a total number of knots $M_n = 24$ is chosen, as no significant improvement is attained beyond this point. In Figure 3, right, the log-likelihood for models with different numbers of knots is shown. The red dashed line is the log-likelihood of the model with the estimates found
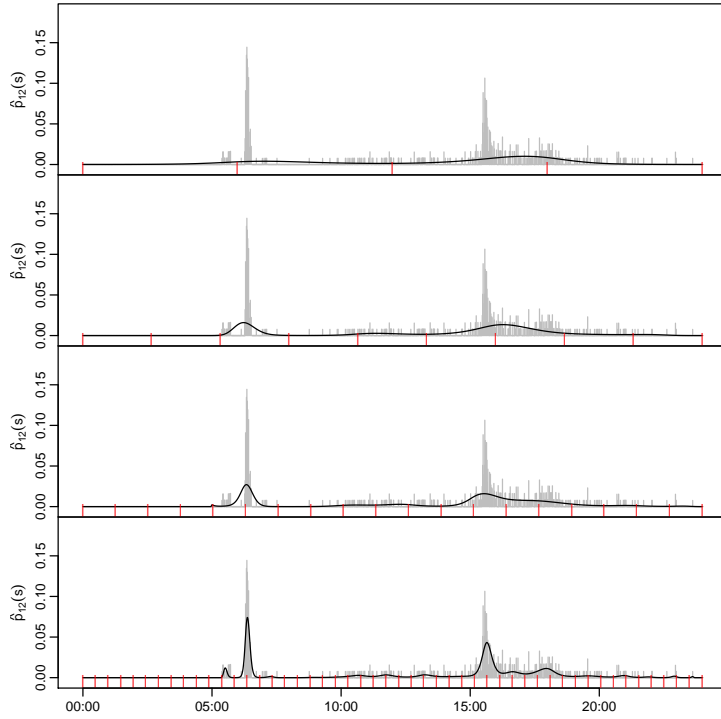
Figure 2: *From top to bottom: Fitting the estimate $\hat{p}_{12}(s)$ where the knots are uniformly distributed in the interval from 00:00 to 23:59 on a weekday, with number of knots $\{5, 10, 20, 50\}$. For reference, the gray bars are the estimates of $\hat{p}_{12}(s)$ from (6) with no parameter reduction. The red bars indicate the knot positions.*
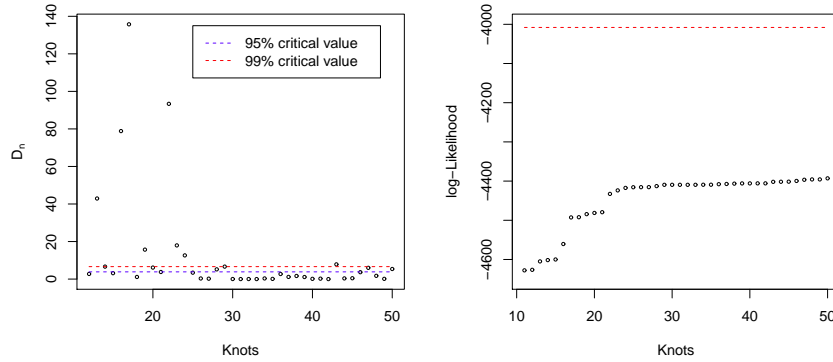

Figure 3: *Left: Log-likelihood ratio test statistic, given by D_n, from the model with n knots vs. the model with $n - 1$ knots. 95% and 99% critical values are shown for a χ^2-distribution with $df = 1$. Right: The log-likelihood of the models with different knots. The red dashed line is the likelihood of the model with estimates based on (6).*

by (6) and corresponds to a perfect data fit. It is in some sense a limit for the

12

fitted models.

The models based on B-splines are sub-models of the model in which a knot is placed at every minute. In this model, the transition probabilities are estimated independently for every minute, and in turn the model corresponds to that with no parameter reduction. The models where the number of parameters is reduced can be tested against the model with no parameter reduction. This leads to a test statistic that will be χ^2-distributed with $1440 - M$ degrees of freedom for each time-varying transition probability. Accordingly the critical value will be very large (> 1475 for estimating one time-varying transition probability and with $M \leq 50$ with 95% significance) and it is difficult to test anything.

The top plot in Figure 4 illustrates the estimate of $\widehat{p}_{12}(s)$ using B-splines with $M = 24$, where the knots are placed by the algorithm introduced in section 3.3. For comparison, the model with the naive knots and $M = 24$ is shown on the bottom plot in Figure 4. By visual inspection, it is observed that the model in which the knots are placed according to the algorithm in section 3.3 better captures the peakedness of $p_{12}(s)$.
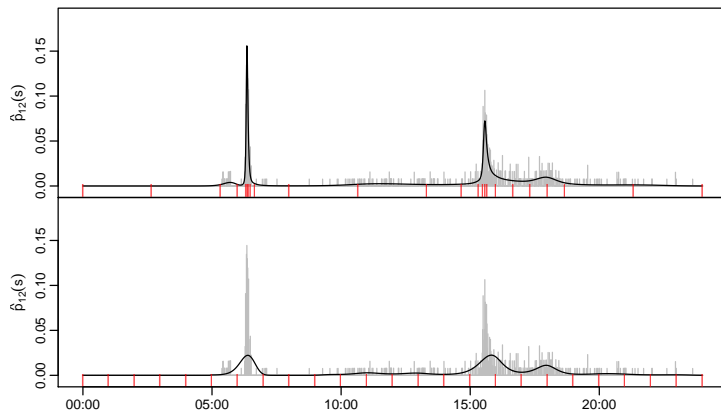

Figure 4: *Top: $\widehat{p}_{12}(s)$ based on the B-splines with $M = 24$ and the knots placed using the algorithm, plotted as the black line over the estimates $\widehat{p}_{12}(s)$ with no parameter reduction. Bottom: $\widehat{p}_{12}(s)$ based on the naive B-splines with $M = 24$, plotted as the black line. The red bars indicate the knot position.*

4.3. Applications

The applications of the proposed stochastic model for driving patterns range from simulating different driving scenarios to calculating the probability of a trip starting within a given interval. In addition, the model is prerequisite to determine the optimal charging scheme for an electric vehicle.

4.3.1. Probabilities and Simulations

Four driving scenarios are simulated and shown in Figure 5. Markov states are indicated in a binary form depending on whether the vehicle is driving "1" or not "0".
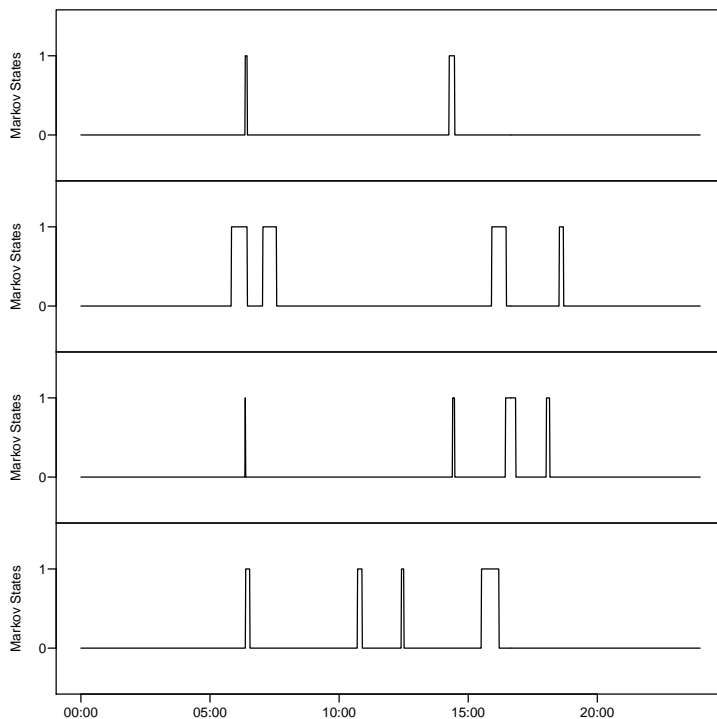
Figure 5: *Four distinct realizations of driving patterns using the proposed stochastic model.*

Next we illustrate how to find the probability of a trip starting within a given interval. Suppose that at time s the vehicle is parked. Denote the waiting time until the next trip starts by Z_s. We have that $Z_s \sim \exp(q_{12}(s))$. The probability of a trip starting within the time interval $[s, s + \tau]$ is thus

$$\mathbb{P}(Z_s \leq \tau) = 1 - e^{-\int_0^\tau q_{12}(s+t)dt}. \tag{30}$$

Using this equation, for example, the probability of a trip starting in the interval from 00:00 to 06:00 is $\mathbb{P}(Z_{00:00} \leq 06:00) = 1 - e^{-\int_0^{6 \cdot 60} q_{12}(t)dt} = 0.131$.

In the top part of Figure 6, the probability of starting a trip within the next hour, conditional on not driving at the beginning of that hour, is depicted. The probability of the vehicle being in use at any time of the day is found using bootstrap (Davison and Hinkley, 1997) and is shown in the bottom part of Figure 6.

4.3.2. Example of Electric Vehicle Charging

A stylized example of charging an EV is considered below. Suppose that the owner desires to have the vehicle fully charged for the next time he/she must use the vehicle. At the same time, he/she wishes to minimize the cost of doing so. As the use of the vehicle is stochastic, it can only be guaranteed that the vehicle is fully charged with some probability. Suppose that the vehicle is
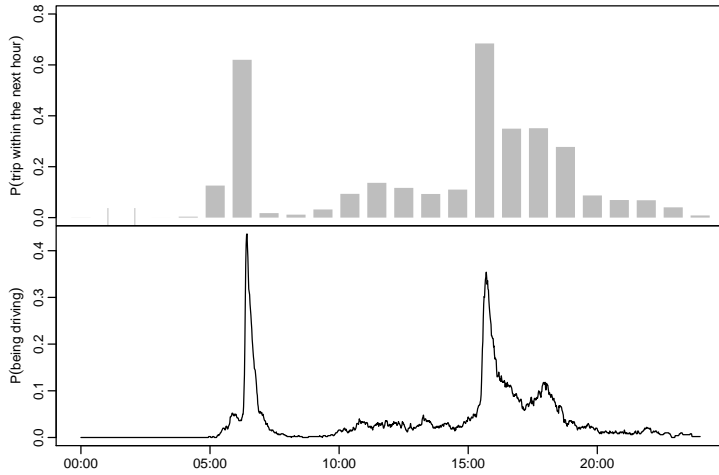
Figure 6: *Top: A graph plotting the probability of starting a trip within the next hour, conditional on not driving at the present time, which is found by applying (30). Bottom: The probability of the vehicle being in use at any time of the day, which is estimated using bootstrap.*

parked at 7:00 am and that the battery is at half capacity. Furthermore, we assume that the battery has a total capacity of 30 kWh and that it takes 6 hours to fully charge the battery, if it is completely empty. In addition we assume a stylized price signal, according to which the electricity price is 60 EUR/MWh until 10:00 am, and then drops to 30 EUR/MWh. We consider that the driving behavior is independent of the electricity price. Moreover, it is assumed that a strategy with early charging is preferred over a strategy which postpones the charging, if both strategies involve identical costs.

In Figure 7 the different charging strategies are presented for varying levels of certainty. It can be observed that lowering the desired level of certainty will delay the charging. As a consequence, the cost of charging is also reduced, which is seen in Table 1.

Certainty	95%	90%	80%	50%
Cost (EUR)	NA*	0.83	0.69	0.45

Table 1: The associated costs for different levels of certainty. *Having the vehicle fully charged with 95% certainty is not feasible. The associated cost for the highest level of certainty that can be achieved (93.9%) is EUR 0.90, which corresponds to the strategy in which the vehicle is immediately plugged in and starts charging, if not driving.

5. Conclusion and Future Research

This paper proposes a suitable model that captures the diurnal variation in the use of a vehicle. The number of parameters is significantly reduced by using B-spline basis functions as explanatory variables in a logistic regression. The
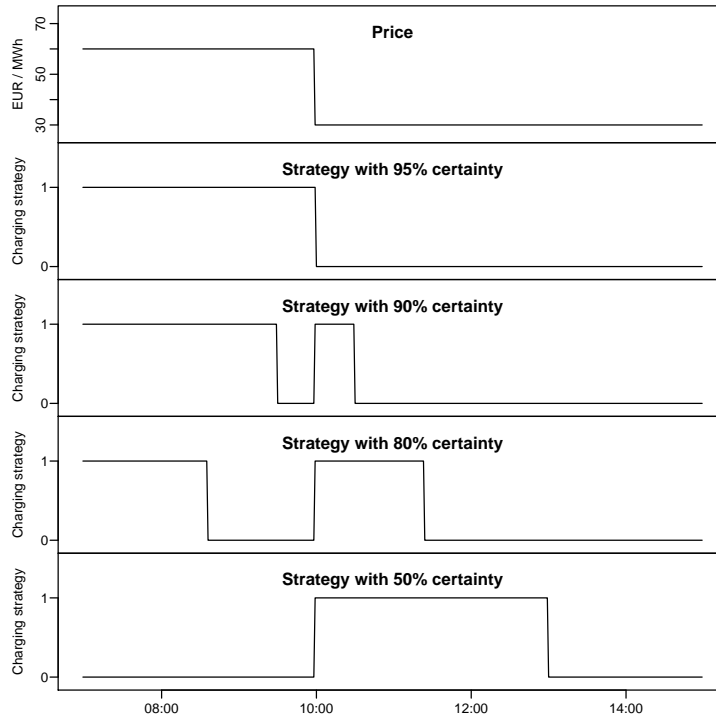
15

Figure 7: *Top: The electricity price. Bottom four: The charging strategies with decreasing certainty of the vehicle being fully charged for the next trip. A "1" indicates that the vehicle is being charged and a "0" indicates that the vehicle is not charging.*

model is versatile and can be applied to describe driving data from any single vehicle, thus providing a reliable model for the use of that vehicle.

A possible extension to the model would be to use hidden inhomogeneous Markov chains. The hidden states could be used to capture where the vehicle is parked, e.g. at home or somewhere else, and to distinguish between different driving environments, e.g. urban and rural. The hidden Markov model does not rely on having observations from each state. The model could be extended to cover a population of vehicles by using a mixed-effect model. Another extension to the model could be to estimate the transition probabilities adaptively in time. An adaptive approach would capture structural changes in the driving behavior, such as variation over the year or a change in use that would follow from the household buying an additional vehicle. Adaptivity will be relevant for applying the model in practice. An obvious next step is to use the model to implement a charging strategy that minimizes the costs of driving considering the underlying uncertainty in the use of the vehicle.

Acknowledgments

References

Acha, S., Green, T.C., Shah, N., 2010. Effects of optimised plug-in hybrid vehicle charging strategies on electric distribution network losses, in: Transmission and Distribution Conference and Exposition, 2010 IEEE PES, pp. 1 –6.

Davison, A., Hinkley, D., 1997. Bootstrap Methods and Their Application. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.

Gan, L., Topcu, U., Low, S., 2011. Optimal decentralized protocol for electric vehicle charging. Technical Report. California Institute of Technology.

Golob, T.F., G.J., 1998. Projecting use of electric vehicles from household vehicle trials. Transportation Research Part B: Methodological 32, 441–454.

Green, R., Wang, L., Alam, M., 2010. The impact of plug-in hybrid electric vehicles on distribution networks: a review and outlook, in: Power and Energy Society General Meeting, 2010 IEEE, pp. 1 –8.

Grimmett, G., Stirzaker, D., 2001. Probability and random processes. Texts from Oxford University Press, Oxford University Press.

Hastie, T., Tibshirani, R., Friedman, J., 2008. The elements of statistical learning: data mining, inference, and prediction. Springer series in statistics, Springer.

Lojowska, A., Kurowicka, D., Papaefthymiou, G., van der Sluis, L., 2011. From transportation patterns to power demand: Stochastic modeling of uncontrolled domestic charging of electric vehicles, in: Power and Energy Society General Meeting, 2011 IEEE, pp. 1 –7.

Ma, Z., Callaway, D., Hiskens, I., 2010. Decentralized charging control for large populations of plug-in electric vehicles: Application of the Nash certainty equivalence principle, in: Control Applications (CCA), 2010 IEEE International Conference on, pp. 191 –195.

Madsen, H., Thyregod, P., 1986. Inhomogenous Markov models for the variation in cloud cover, in: Third International Conference on Statistical Climatology, June 23-27, 1986, The Institute of Mathematical Statistics and Operation Research, The Technical University of Denmark. Vienna, Austria.

Madsen, H., Thyregod, P., 2010. Introduction to general and generalized linear models. Chapman & Hall/CRC Texts in Statistical Science, Chapman & Hall/CRC.

Mao, W., Zhao, L.H., 2003. Free-knot polynomial splines with confidence intervals. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65, 901–919.

Pawitan, Y., 2001. In all likelihood: statistical modelling and inference using likelihood. Oxford science publications, Clarendon Press.

Pearre, N.S., K.W.G.R., Elango, V., 2011. Electric vehicles: How much range is required for a day's driving? Transportation Research Part C: Emerging Technologies 19, 1171–1184.

Richardson, P., Flynn, D., Keane, A., 2012. Optimal charging of electric vehicles in low-voltage distribution systems. Power Systems, IEEE Transactions on 27, 268 –279.

Rotering, N., Ilic, M., 2011. Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets. Power Systems, IEEE Transactions on 26, 1021 –1029.