Electric vehicle charging point placement optimisation by exploiting spatial statistics and maximal coverage location models

Guanpeng Dong[a], Jing Ma[b,*], Ran Wei[c], Jon Haycox[a]

[a] *Department of Geography and Planning, University of Liverpool. Roxby Building, Chatham St, Liverpool L69 7ZT.*
[b] *Faculty of Geography, Beijing Normal University.*
[c] *School of Public Policy, University of California, Riverside, CA, USA.*

Abstract

Electric vehicles (EVs) are increasingly considered as a promising solution to tackle climate change impacts, improve air quality, and enhance growth sustainability. This paper proposes a two-step approach for optimally deploying charging points (CPs) by bringing together spatial statistics and maximal coverage location models. CP locations are conceptualised as a spatial point pattern, driven by an underlying stochastic process, and are investigated by using a Bayesian spatial log-Gaussian Cox process model. The spatial distribution of charging demand is approximated by the predicted process intensity surface of CP locations, upon which a maximal coverage location model is formulated and solved to identify optimal CP locations. Drawing upon the large-scale urban point of interest (POI) data and other data sources, the developed method is demonstrated by exploring the deployment of CPs in London. The results show that EV charging demand is statistically significantly associated with workplace population density, travel flows, and densities of three POI categories (transport, retail and commercial). The robustness of model estimation results is assessed by running spatial point process models with a series of random subsets of the full data. Results from a policy scenario analysis suggest that with

*Corresponding author
Email address: jing.ma@bnu.edu.cn (Jing Ma)

increasing numbers of charging stations to be planned, optimal CP locations gradually expand to the suburban areas of London and the marginal gains in charging demand covered decrease rapidly.

Keywords: Electric vehicle charging point placement, spatial analysis, the maximal coverage location model, log-Gaussian Cox process

---

1. Introduction

Promotion of the transition from high-emission and fossil-fuelled vehicles to low-emission electric vehicles (EVs) has been increasingly considered as a promising solution to mitigate climate change impacts, improve air quality and reduce health risks, and enhance sustainability (Cai et al., 2014; Ellingsen et al., 2016). Goals for the adoption rate of EVs have been set in many countries along with a set of economic and policy incentives for EV promotion that have been implemented. For instance, the UK government is committed to terminate new sales of petrol and diesel cars by 2040 and achieve a 100% use of zero emission of vans and cars by 2050 (Butcher et al., 2018). New EV registrations are increased from 3,500 in 2013 to more than 47,000 by the end of September 2017 in the UK, with a sharp acceleration since 2014 (Butcher et al., 2018).

A continuing adoption of EVs by residents and business relies heavily on a matched progress on the provision of charging infrastructure network as low spatial coverage and inferior sitting of public charging infrastructure could seriously hamper the uptake of EVs from prospective customers (e.g. Cai et al., 2014; Namdeo et al., 2014). Providing an adequate and spatially optimal public charging points (CP) network could, to a large extent, reduce "range anxiety" – relatively short driving range of EVs, which ranks the top of consumers' concerns with the adoption of EVs (Chen et al., 2013). While various provision plans on publicly accessible CP infrastructure have been put forward in many cities and countries, studies on developing computational and mathematical frameworks for optimal CP sitting over space are also experiencing fast growth (e.g. Wagner et al., 2013; Shahraki et al., 2015; Tu et al., 2016; Wei et al., 2018).

The CP location problem usually involves two steps: estimating the spatial distribution of charging demand in a study area and uncovering optimal CP locations by formulating a mathematical programming model (e.g. Tu et al., 2016). The estimation and prediction of public charging demand in the first step, however, is challenging primarily due to lack of realistic and detailed data on usage of charging stations (Cai et al., 2014). Various proxy variables for charging de-

mand have been proposed in previous studies, for instance road traffic density and vehicle ownership rates (see a review in Cai et al. 2014). Considering the fact that EVs charging is more likely to take place at a trip destination rather than the middle of a trip (Sadeghi-Barzani et al., 2014; Shahraki et al., 2015), travel behaviour models, calibrated based on household travel surveys, were used to predict origin-destination flows and approximate EV charging demand (e.g. Chen et al., 2013; Xi et al., 2013). However, it has been argued that the spatial distribution of public EV charging demand could differ substantively from that of total travel demand due to the required time of EV charging and the possibility of EVs charged at users' residences (Namdeo et al., 2014). In addition, the representativeness of demand estimates based on individual travel surveys or small samples of realistic use of EVs is also unclear (Cai et al., 2014). In the second-stage analysis of a CP location problem, a mathematical programming model was built and solved to identify optimal CP locations that collectively maximise (or minimise) certain objective functions, which will be discussed below. The objective function could vary across various research focuses and nature of data under study (Shahraki et al., 2015; Ko et al., 2017).

Recently, urban big data and fine spatio-temporal resolution public transport trajectory data in particular have been exploited to infer EV charging demand distribution in the CP placement studies. Drawing upon a large-scale daily trajectory data of taxis in Beijing, Cai et al. (2014) found that collective parking hotspots of taxis could serve as good indicators for EV charging demand. Based on the revealed charging demand from taxi travel patterns and existing gas stations (as candidate charging station locations), Shahraki et al. (2015) built a mathematical programming model with an objective of maximising the amount of electrified fleets to find optimal locations of CPs. Considering the difference in daily travel patterns between electric and conventional gasoline taxis, Tu et al. (2016) explicitly derived the spatio-temporal charging demand dynamics by using massive trajectories of electric taxis in Shenzhen, China and solved the optimal placement of charging stations by formulating a mathematical model that takes into account spatio-temporal variabilities in daily charging

demand. Real-world massive taxi trajectories undoubtedly offer great benefits to reveal fine-granular travel patters in space and time, thus improving charging demand estimates. Nonetheless, how and to what extent the charging demand distribution revealed by taxis alone could approximate total charging demand still needs a careful validation. For instance, daily travel patterns and charging behaviours of private EV drivers and taxi drivers could differ substantially. From an implicit charging demand point of view, Wagner et al. (2013) exploited associations between large-scale urban points of interest (POI) data and existing CP locations, upon which the spatial distribution of EV charging demand was calibrated. Point-referenced POIs characterise urban land use differentiation (e.g. work, shopping, and recreation) at the finest spatial resolution and constitute residents' potential daily trip destinations. By associating the usage of existing CPs to densities of various POI categories via a simple linear regression model, specific POI categories exerting statistically significant influences on EV charging demand were selected to construct spatial charging demand indices (Wagner et al., 2013).

This study presents a new approach to address the optimal CP location problem by using a combination of spatial statistics and mathematical programming models. It innovatively considers the spatial pattern of existing CP locations as a spatial point process, with the *latent charging demand* approximated by a spatially continuous process intensity surface. Intuitively, areas with larger charging demands are expected to be allocated with more CPs. The same reasoning is seen in a spatial point process model – a stronger process intensity of an event implies a larger number of event occurrence (Diggle, 2013). By establishing this link, a spatial statistics model, more specifically, the log-Gaussian Cox process model (Møller et al., 1998; Illian et al., 2012; Diggle, 2013) is employed to explore associations between the spatial pattern of CPs and densities of POIs by types and other factors. Predictions of the spatial distribution of EV charging demand are then estimated from the model. Subsequently, the optimal CP location problem is formulated as a standard maximal coverage location problem (MCLP, Church and ReVelle 1974) – maximising the coverage of EV

charging demand by locating a fixed amount of CPs. MCLP has been proven to be a rigorous locational analysis framework for optimally placing facilities on the landscape with constraints, being widely used in regional sciences, geography and environmental sciences fields (Murray, 2016; Snyder and Haight, 2016; Wei, 2016).

Our contributions to the CP placement literature lie in several aspects. Primarily, the locations of CPs are explicitly modelled as a spatial point process for the first time in the literature, to the best of our knowledge, allowing for an intuitive derivation of the distribution of charging demand over space with fine resolution. The spatial log-Gaussian Cox process model is employed to predict the demand of EV charging and its spatial distribution. With a spatial statistics modelling framework, there is great flexibility to capture potential spatial heterogeneity and dependency in the latent spatial process intensity (or the clustering pattern of CP locations). Moreover, this study provides insights into the potential spatial factors of EV charging demand, drawing upon open urban POI, traffic and socio-demographic data and open source statistical modelling and mathematical programming tools. With ever increasing coverage rate and positional accuracy of POIs in cities, the proposed approach would be beneficial to city planners in terms of placing new CP infrastructure and evaluating efficiency of existing CP configuration and making according adjustments.

The reminder of this paper is structured as follows. Section 2 describes the proposed methodology and data. Section 3 presents findings obtained from the spatial statistical and mathematical models, as well as results from robustness and policy scenario analyses. Section 4 concludes with a brief summary of findings.

2. Methodology

In a nutshell, our method consists of two steps: (1) a spatial point process statistical model of CP locations; (2) a mathematical maximum coverage location model. In Step 1, the inputs of the statistical model are the spatial

pattern of existing CP locations as outcomes and densities of POI, population and traffic as independent variables (detailed below). The outputs are estimates on spatially continuous charging demand intensity surface in the study area and covariate effects. In Step 2, the maximum coverage location model takes estimated charging demand as an input variable and produces the optimal deployment of CP locations as output. The schematic diagram of the methodology is illustrated in Fig. 1.

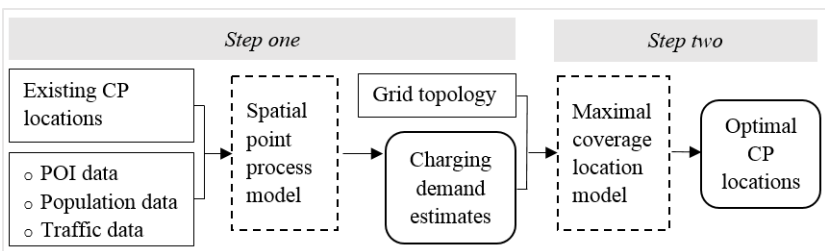

Fig 1. The schematic diagram of the proposed methodology.

2.1. A spatial log-Gaussian Cox process model of CP locations

A Cox process model assumes the realisation or presence of a spatial point pattern (e.g. the pattern of existing CP locations) to be driven by a latent non-negative stochastic process intensity surface, and that given the intensity field, the point pattern follows a Poisson process, i.e. a random number of points (CPs in this study) are located independently and uniformly in an area (Møller et al., 1998; Law et al., 2009; Diggle, 2013). Denote a study area as $S \subset \mathbb{R}^2$ with space indexed by s, and the intensity surface as $\lambda(s)$. The number of CPs in $S' \subset S$, $N_{S'}$, follows a Poisson distribution with an expected value of $\Lambda(S') = \int_{S'} \lambda(s)ds$. Given a spatial pattern P and intensity surface $\lambda(s)$, the data likelihood is expressed as (Simpson et al., 2016),

$$l(P|\boldsymbol{\lambda}) = \exp\left\{|S| - \int_S \lambda(s)ds\right\} \prod_{s_k \in P} \lambda(s_k), \tag{1}$$

where $|S|$ is the area of P. The log-Gaussian process model is a special class of a Cox model, in which the log-intensity surface, $\log\{\lambda(s)\}$, is a Gaussian process or random field, say $\eta(s)$ (Møller and Waagepetersen, 2007; Diggle, 2013). There is great flexibility to include into $\eta(s)$ any relevant covariate effects and spatial effects such as spatial heterogeneity and dependency if desired (Taylor et al., 2015).

For the log-Gaussian Cox process model to be practically implemented, a study area is usually approximated by a fine-resolution grid geography (Illian et al., 2012; Taylor et al., 2015) – a discrete representation of space. In this study, we delineate the study area into a set of spatial grids with a resolution of 1 km×1 km, which will serve as our basic analysis units. The choice of such a spatial resolution is pragmatic. On one hand, adopting spatial grids with much finer resolution (say 100 metres) will substantially increase the sample size to a magnitude ($\sim 170,000$) that would make the computation of a log-Gaussian Cox process model intractable. On the other hand, a spatial resolution of 1 km implies a service radius of 1km for charging stations, which is similar to the 1.6 km range suggested by Cai et al. (2014). As there are only a small proportions of households with off-street parking facilities in London (Namdeo et al., 2014), we adopt a relatively smaller service range. However, we acknowledge that a reliable estimate on the service range of charging stations, ideally learned from data on long-term daily use of charging stations, shall be one priority in the future CP deployment studies. With the grid topology, s indexes each grid and $\lambda(s)$ presents the grid-level intensity process. Let N_s be the number of CPs in grid s, the spatial log-Gaussian Cox process model in this study is defined as follows:

$$\begin{aligned} N_s &\sim \text{Poisson}(\lambda(s)) \\ \log\{\lambda(s)\} &= \eta(s) = X(s)\beta + \psi(s) + \epsilon(s), \end{aligned} \tag{2}$$

where $X(s)$ contains a vector of independent variable values at grid s, including POI densities, population density, traffic flows, and an intercept term. β

is the associated covariate effect vector to estimate. The sum of $\psi(s)$ and $\epsilon(s)$ describes the intensity process unexplained by the included covariate effects. More specifically, $\psi(s)$ is a spatially structured random effect, which captures potential spatial dependency or smoothness in the intensity process or clustering pattern of CP locations. In this study, spatial dependence or correlation refers to that process intensity values of nearby grids tend to be more similar than that of distant grids. Ignoring such correlation effects would lead to unreliable statistical inferences on covariate effects (e.g. Congdon, 2014). An intrinsic conditional autoregressive (CAR) model (e.g. Lee, 2011; Congdon, 2014; Dong et al., 2016) is specified for $\psi(s)$, resulting in the distribution of $\boldsymbol{\psi}$ (over grids) being a multivariate Normal, $p(\boldsymbol{\psi}) \sim MVN(0, (D_W - W)^{-1}\sigma_\psi^2)$ where W is the neighbourhood structure matrix with element $w(s_i, s_j) = 1$ if grids s_j and s_i share an edge, and $w(s_i, s_j) = 0$ otherwise. D_W is a diagonal matrix with elements being the number of neighbouring grids for each grid or the row-wise sums of W. σ_ψ^2 is a variance parameter measuring the variation of ψ. An independent random effect $\epsilon(s)$ is specified to capture potential spatial heterogeneity in the process intensity surface and the well-known over-dispersion effect often found when applying Poisson models to spatial count data (e.g. Congdon, 2014). It is assumed to follow a Normal distribution, $\psi \sim N(0, \sigma_\psi^2)$. The spatial heterogeneity effect could be understood as an unobserved place-specific (i.e. grid-specific) effect on point process intensities, leading to potential local discontinuities on the intensity surface. The specification of two additive sets of random effect ($\psi(s)$ and $\epsilon(s)$) is equivalent to the well-studied BYM model in the spatial statistics literature (Besag et al., 1991).

The stochastic intensity surface driving the spatial pattern of CP locations is intuitively linked to the spatial distribution of charging demand of EVs in the study area. A higher intensity value at a grid implies a larger demand of EV charging, and thus a greater amount of CPs expected to be placed there. After implementing the log-Gaussian Cox process model displayed in Eq 2, the fitted or predicted intensity surface is conceptualised as the *latent* EV charging demand and serves as a key input in the maximal coverage location model

(detailed below).

Fitting log-Gaussian Cox process models is typically based on the Bayesian Markov chain Monte Carlo (MCMC) simulation approach due to the intractable data likelihoods of Eqs 1 and 2 (Taylor et al., 2015; Simpson et al., 2016). Instead of the computationally expensive MCMC method, we resort to a recently developed approximate Bayesian inference approach, the integrated nested Laplace approximations (INLA, Rue et al. 2009), for model estimation and comparison. The INLA approach has been made conveniently accessible via an open source INLA software package (Rue et al., 2009; Martins et al., 2013) in R (R Core Team, 2017). Fitting specifically the log-Gaussian Cox process models with INLA is discussed in Illian et al. (2012, 2013), among others.

2.2. The maximal coverage location model

Before formulating the MCLP model we introduce relevant notations following Wagner et al. (2013) and He et al. (2016):

$\mathbb{S}$ = the set of grids, indexed by $i, i = 1, \ldots, G$

$\mathbb{S}'$ = the set of potential CP locations (grids), indexed by $j, j = 1, \ldots, G$

$N_i = \{j | w_{ij} = 1\}$, the set of CP locations that covers the charging demand of grid i

P = the number of CP to be deployed

$$y_j = \begin{cases} 1, & \text{if grid } j \text{ is selected to place CP} \\ 0, & \text{otherwise} \end{cases}$$

$$z_i = \begin{cases} 1, & \text{if the demand at grid } i \text{ is met} \\ 0, & \text{otherwise} \end{cases}$$

We note that N_i entails the i-th grid itself and the neighbouring grids as defined by the neighbourhood structure matrix W in the log-Gaussian Cox process model. It is equivalent to a 1 km service radius for each CP. y_j is the key decision variable, indicating whether grid j is selected to locate a CP or not. z_i is another decision variable implying whether the charging demand at grid i is

met, which depends on the values of y at grid i and neighbouring grids. The MCLP is formulated as follows:

$$\text{Maximise} \quad \sum_i \hat{\lambda}_i z_i \tag{3}$$

$$\text{Subject to} \quad \sum_{j \in N_i} y_j - z_i \geq 0, \quad \forall i \in \mathbb{S} \tag{4}$$

$$\sum_{j \in \mathbb{S}'} y_j = P \tag{5}$$

$$\sum_{j \in N_i} y_j \leq \bar{k}, \quad \forall i \in \mathbb{S} \tag{6}$$

$$y_j \in \{0,1\}, z_i \in \{0,1\}, \quad \forall i \in \mathbb{S}, \forall j \in \mathbb{S}' \tag{7}$$

Eq 3 is the objective function maximising the total EV charging demand coverage as in Wagner et al. (2013) and He et al. (2016). Constraint 4 determines whether the demand at grid i is covered. If the demand of grid i is served, i.e., $z_i = 1$, then at least one CP has to be placed in the neighbourhood N_i and thus $\sum_{j \in N_i} y_j \geq 1$. In contrast, if the demand of grid i is unmet, i.e., $z_i = 0$, there should be no CPs deployed in N_i, and thus $\sum_{j \in N_i} y_j = z_i = 0$. Constraint 5 defines that P CPs are to be deployed. Constraint 6 is to avoid spatial clustering of CP location by limiting the number of CPs capable of serving the demand of grid i ($\forall i \in \mathbb{S}$) to $\bar{k}$, which is set to three in the study. Substantively, this constrain is to reduce excess supply and extend the geographical coverage of CP locations. Similar idea was discussed in Wagner et al. (2013) but addressed by using a computationally expensive iterative MCLP approach. Lastly, constraint 7 ensures both decision variables to be binary integers. It is useful to note that the estimated charging demand $\hat{\boldsymbol{\lambda}}$ represents an approximation for current charging demand but not necessarily the potential demand because an implicit assumption here is that all CPs are equally used. Such a restrictive assumption is imposed due to lack of data on actual usage of individual charging stations. The MCLP formulated above is a standard mixed integer linear programming model. It is solved by using an open source R library, OMPR (Optimization

Modelling Package, Schumacher 2017).

2.3. Data and Variables

This study focuses on London, one of the nine regions in England, with a population of more than eight million according to the 2011 census data. The number of EVs in London has grown significantly to over 10,000 as of 2017, experiencing much faster growth than other regions or the nation as a whole. [1] Associated with the tremendous increase in the EVs is a vast amount of investment to improving the CP infrastructure. However, how and where new CPs are to deploy remain unclear.

The existing CP data is from the National Charge Point Registry (NCPR), which is a government initiative set up in 2011 to provide a credible database of publicly-funded charge points, in support of government's objectives to promote EV adoption in the UK (NCPR, 2017). Based on the address information of each CP, those located in London were selected and geo-referenced, giving about 400 public CPs used in the following analyses. POIs are recognised as specific point locations of places that residents may find useful or interesting, e.g. public transport facilities, schools, restaurants and offices. POI data used in this study is from the Ordnance Survey Points of Interest product, which contains unique locations of four million different geographic features in a 3-tier hierarchy (9 groups, 52 categories and 616 classes) (Ordanace Survey, 2015). For simplicity and practical computational concerns, we categorised POIs at the group level, which includes accommodation, eating and drinking; commercial services; attractions; sport and entertainment; education and health; public infrastructure; manufacturing and production; retail; and transport. A detailed description of the classification system is provided in the user manual of the product.

Except for the POI data, we also include in our model population character-

[1] CV figures are from the vehicle licensing statistics from the Department of Transport of the UK

istics and traffic flow variables that are potentially related to the EV charging demand (Chen et al., 2013; He et al., 2016). More specifically, night-time and workplace population counts, obtained from the 2011 UK census, are associated to the spatial pattern of CPs. Both workplace and night-time population are measured at the LSOA (Lower Layer Super Output Areas) scale, a fine-resolution census geography usually accommodating 1,000 to 3,000 residents according to the Office for National Statistics of the UK. There are 4,835 LSOAs in London in 2011. In addition, traffic flows on the major road network in the UK, measured by the annual average daily flows for each junction-to-junction link from 2015 to 2016, have been made publicly available by the Department for Transport of the UK and included in our analyses. As LSOAs and grids are not compatible geographies, a standard GIS areal weighting approach was used to transfer population counts available for LSOAs onto the 1 km $\times$ 1 km grids. We refer to Lloyd et al. (2017) and references therein for a detailed description of areal weighting and potential issues associated with it. The traffic flow of a grid is obtained by finding the highest traffic value of a major road segment that intersects the grid, or if no lines intersect the grid, by using the traffic value of the nearest road segment.

The counts or densities of POIs by category over grids are obtained by overlaying POI point layers with the grid layer. Fig. 2 shows spatial distributions of workplace population and POI density at the spatial grid scale. All the independent variables are log-transformed to reduce the potential influence of heteroskedasticity on model estimation. To evaluate possible impacts of multicollinearity on model estimation results, we conduct a robustness analysis by implementing a series of spatial point process models with randomly drawn subsets of the full data. Moreover, only predictor variables with statistically significant impacts on the spatial pattern of CPs are used for charging demand prediction in our preferred model (detailed below). The descriptions and summary statistics for the variables used in the study are provided in Table 1.

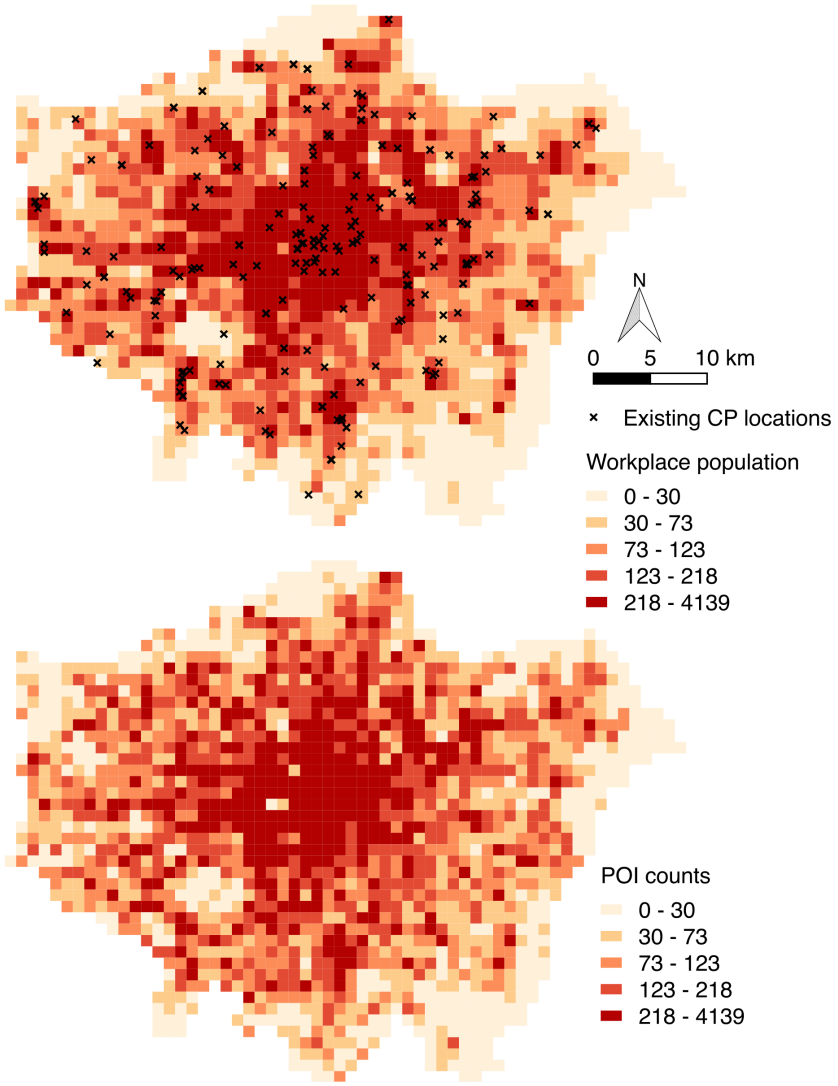

Fig 2. The spatial distributions of existing charging points and densities of workplace population and POIs in London.

Table 1. Descriptive statistics for variables used in the study.

Name	Description	Mean/ Std.Dev
CP counts	The number of charging points per grid cell	0.16 (0.65)
Independent variables		
Night-time population	Log of night-time population per grid in 2011	7.81 (1.50)
Workplace population	Log of workplace population per grid in 2011	6.98 (1.31)
Traffic	Log of annual average daily follows per grid in 2015	10.04 (0.68)
Point of interest (POI) data as of 2017		
Accommodation	Log of the number of POIs of accommodation, eating and drinking	0.20 (0.34)
Attractions	Log of the number of POIs of attractions	0.81 (0.39)
Commercial	Log of the number of POIs of commercial services	1.34 (0.50)
Education	Log of the number of POIs of education and health	0.98 (0.51)
Manufacturing	Log of the number of POIs of manufacturing and production	0.89 (0.48)
Public	Log of the number of POIs of public infrastructure	1.26 (0.37)
Retail	Log of the number of POIs of retail	0.97 (0.57)
Sports	Log of the number of POIs of sports and entertainment	0.82 (0.44)
Transport	Log of the number of POIs of transport	1.15 (0.39)

3. Results

3.1. Spatial point process model estimation results

Four specific log-Gaussian Cox process models were fitted for the spatial pattern of existing CP locations to understand the potential drivers of EV charging demand. Models 1 and 2 are standard Poisson regression models without considerations to spatial heterogeneity and dependency effects. Model 3 takes into account spatial heterogeneity by the inclusion of a set of independent random effect, $\epsilon(s)$. Model 4 further accounts for possible spatial dependency effect. Deviance Information Criterion (DIC, Spiegelhalter et al. 2002), a commonly used model fit index in Bayesian inference, has been used for model comparison. A better model specification is indicated by a smaller DIC value.

Table 2 reports estimation results from Models 1 and 2. When interpreting estimates on covariate effects, we use the intensity of the distribution of CPs and *latent* EV charging demand interchangeably for simplicity. From Model 1, we see that workplace population and traffic flows are statistically significantly and positively related to the sitting of CPs, everything else equal. Of the nine POI categories, three exert statistically significant influences on the location of CPs. Transport and retail POI densities influence the location of CPs positively. Putting the magnitude of these effects in perspective, a one unit increase in transport and retail POI densities on the log scale elevates the underlying charging demand by 182% and 61%, respectively. These results are in line with the observation that current CPs are often located in public car parks and service stations (e.g. Chen et al., 2013; Cai et al., 2014). Public car parks, large transport hubs and retail centres tend to offer required spaces for sitting charging stations. Meanwhile they attract residential and business activities taking place there and nearby places. Commercial POI density is negatively related to locations of CPs, everything else equal. A plausible explanation is that although commercial POIs potentially attract population flows, their physical space conditions for charging station deployment might not be ideal. Urban POIs as a whole serve as places where residential and business routine activities take

place, generating traffic flows, but they seem to have differentiating impacts on charging demand.

In Model 2, only variables with statistically significant associations with EV charging demand are included. This leads to a concise model specification and alleviates the issue of multi-collinearity. The DIC values also have a slight decrease from Model 1 to Model 2, implying there is not any loss in model fit by adopting a simpler model specification. The estimates on covariate effects are similar in both models. Model 3 considers potential spatial heterogeneity in the underlying intensity surface of CP locations. A substantial improvement in model fit is achieved, as indicated by a large decrease of 386 in DIC values from Model 2 to Model 3. Estimates on variable coefficients in Model 3 also differ from that in Model 2, especially for workplace population and traffic flow densities. A significant spatial heterogeneity effect signifies place-specific influences, often difficult to be quantified and thus impossible to be included in the model, on charging demand (e.g. spatially varied physical conditions for placing charging stations).

Model 4 further incorporates a spatially structured random effect ($\psi(s)$ in Equation 2 to deal with potential spatial dependencies in the distribution of charging demand. However, the decrease in DIC is only marginal (about 3.7) from Model 3 to Model 4, which implies that the inclusion of spatial dependency effect does not improve model fit significantly. The reason is that the marginal variance of $\boldsymbol{\psi}$, calculated as the empirical variance of posterior estimates on $\psi(s)$ (Blangiardo and Cameletti, 2015), only accounts for a small proportion of the total variance of the latent intensity surface, net of the included covariate effects. In this study, spatial heterogeneity (i.e. independent spatial variations) dominates the spatial distributions of charging demand surface controlling for covariate effects. Thus, Model 3 was used to predicted the latent EV charging demand for each grid, i.e. $\lambda(s)$. We note that fitted values of $\lambda(s)$ from Model 3 and Model 4 are highly correlated with a Pearson coefficient of about 0.98.

Table 2. Estimation results from two Poisson regression models.

	Model 1			*Model 2*		
Variables	Median	2.5%	97.5%	Median	2.5%	97.5%
Intercept	-9.606*	-12.26	-6.992	-9.037*	-10.98	-7.084
Nighttime population	-0.076	-0.257	0.115			
Workplace population	0.381*	0.119	0.635	0.32*	0.126	0.507
Traffic	0.254*	0.061	0.444	0.231*	0.043	0.415
Point of interest variables						
Accommodation	-0.104	-0.307	0.09			
Attractions	-0.145	-0.307	0.017			
Commercial	-0.617*	-0.958	-0.265	-0.529*	-0.797	-0.249
Education	0.103	-0.192	0.407			
Manufacturing	-0.155	-0.341	0.031			
Public	0.398	-0.059	0.860			
Retail	0.475*	0.267	0.682	0.5*	0.298	0.701
Sports	0.027	-0.205	0.260			
Transport	1.04*	0.731	1.355	1.08*	0.781	1.385
DIC	1475.9			1472.3		
PD	12.87			5.986		

Note: * represents statistical significance at the 95% credible interval; DIC represents deviance information criterion; PD represents the effective number of parameters.

Table 3. Estimation results from Poisson regression models accounting for spatial heterogeneity and dependency effects.

	Model 3			*Model 4*		
Variables	Median	2.5%	97.5%	Median	2.5%	97.5%
Intercept	-13.02*	-16.27	-9.982	-13.04*	-16.29	-9.993
Workplace population	0.55*	0.256	0.848	0.552*	0.258	0.852
Traffic	0.351*	0.075	0.628	0.351*	0.075	0.628
Point of interest variables						
Commercial	-0.511*	-0.920	-0.087	-0.511*	-0.920	-0.087
Retail	0.49*	0.187	0.801	0.489*	0.186	0.800
Transport	0.999*	0.576	1.441	1.0*	0.576	1.441
σ^2_ϵ	2.186	1.629	2.968	2.196	1.603	3.815
σ^2_ψ				0.006	0.005	0.008
DIC	1086.4			1082.7		
PD	223.3			221.0		

Note: * represents statistical significance at the 95% credible interval.

3.2. Optimal CP locations

The estimated EV charging demand intensity is shown in the top panel of Fig. 3 with breaking points being quintiles of the distribution. EV charging demand is highly concentrated around the city centre and becomes dispersed when moving towards suburban areas. As there are 161 grids containing CPs as of 2017 in London, the number of grids to deploy CPs (P in Eq 5) is set to 161 in our MCLP model. The optimal CP locations are shown in the bottom panel of Fig. 3. The spatial distribution of optimal CP locations is not highly clustered than would be predicted if the sitting of CPs was purely based on ranks of charging demand, a finding similar to Wagner et al. (2013).

Current CP locations were superimposed on the optimal CP locations in the bottom panel of Fig. 3. There seems to be a fair overlap between the grids with existing CPs and optimal grids selected to locate CPs, which is not surprising as the distribution of latent charging demand is learnt from the spatial pattern

of existing CPs. Differences in patterns of the current and optimal CP locations are also clear. In the central areas of London, the spatial concentration level of current CP locations is much higher than that of the optimal CP locations. It might be in relation to a lack of strategic spatial planning of charging infrastructure with an aim to maximise charging demand from residents and business (Transport for London, 2018). With respect to the proportion of total charging demand that are accounted for in the study area, it is about 70% for the spatially optimised charging stations and 60% for the current stations

As a demonstration of our approach to be a tool that could be useful for strategic charging infrastructure planning, a simple experiment is conducted. Imagine that the charging demands of grids currently sitting in the lowest decile of the demand distribution are increased to the median demand level (Fig. 4), where should the new CPs be placed? The increase in charging demand could be due to new urban (re)generation programs. Intuitively, new CPs are expected be deployed in a subset of grids whose charging demands have yet to be met, and the selected locations should maximise charging demand coverage.[2] Fig. 4 shows the optimal 50 locations for deploying new CPs in line with the above criteria. Of course, city planners have the flexibility to decide the number of new CPs (other than 50) to be deployed. We note that new CPs do not have to be located in grids where charging demand elevates as the objective in the MCLP model is to maximise the overall demand coverage for grids whose demands have not been met.

[2] An alternative and simple way to explore new CP locations associated with demand shocks would be maximising total demand coverage only conditioning on the existing CP locations. This is achieved by setting the decision variable y_j at the existing CP locations to 1.

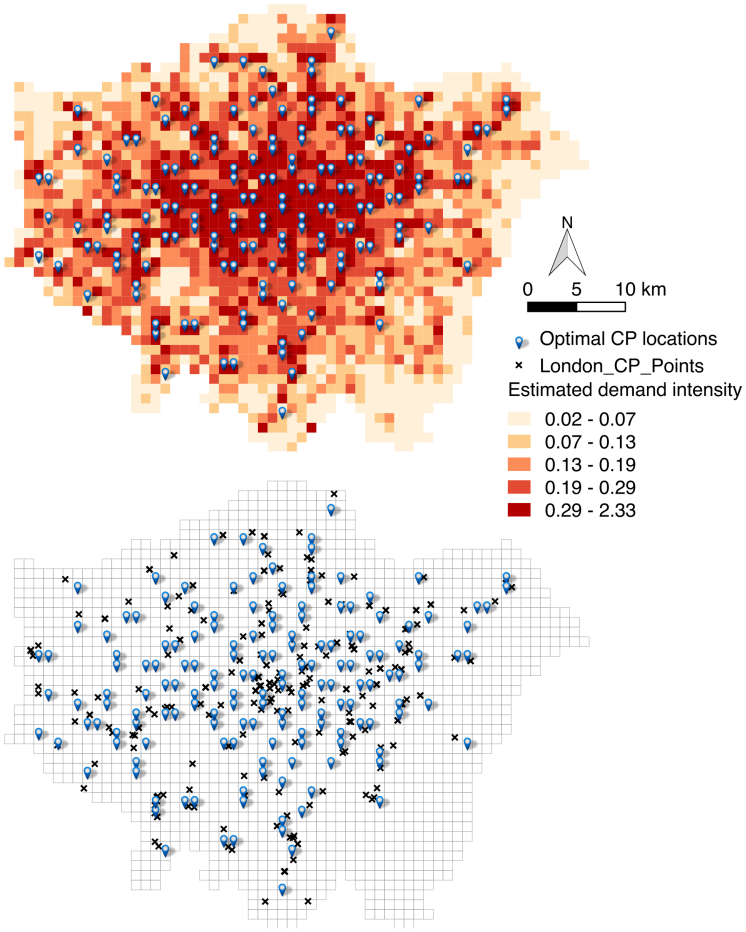

Fig 3. Estimated demand intensity and spatially optimal charging point locations.

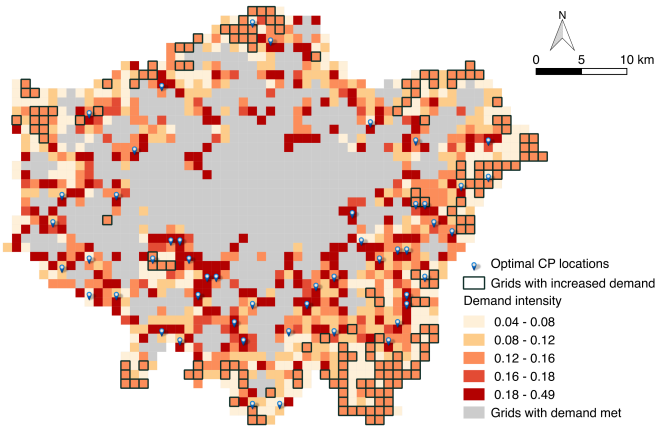

Fig 4. The spatially optimal charging point locations in the presence of sudden changes in charging demand

3.3. A robustness analysis

The optimal CP locations calibrated from the MCLP model rely on the estimate of latent changing demand ($\hat{\boldsymbol{\lambda}}$) from the spatial point process model. To assess the robustness of estimates on the distribution of charging demand and covariate effects, we implemented a series of log-Gaussian Cox process models with random subsets of the full data. Each subset contains observations randomly drawn from the full data with certain probability (ranging from 0.6 to 0.99), leading to 40 subsets in total. Models implemented with subsets of the full data are referred to comparison models for simplicity. Estimates on covariate effects and $\boldsymbol{\lambda}$ from each model were stored and compared to that from the preferred model (Model 3 in Table 3). With respect to covariate effects, three metrics were calculated for each predictor: the variability – squared differences in coefficient estimates between comparison models and Model 3, presented as percentages of estimates from Model 3; the percentage of estimates from comparison models having the same coefficient sign (positive or negative) as the estimates from Model 3 (Column "Coef sign" of Table 4); and the percentage of coefficients that are statistically significant (Column 4 of Table 4). In terms of estimated charging demand, we calculated the Pearson correlation coefficients

Table 4. Robustness check of model estimation results.

Variables	Variability	Coef sign	Statistical significance
Intercept	-0.98	100	100
Workplace population	0.20	100	100
Traffic	0.14	100	87.5
Commercial	-0.01	100	97.5
Retail	0.18	100	100
Transport	0.02	100	100
$\hat{\lambda(s)}$	$0.991[0.987, 0.994]^{a}$		

Note: [a] indicate the mean, lower and upper quantiles of the Pearson correlation coefficients between $\hat{\lambda(s)}$ from comparison models implemented with subsets of the full data and that from the preferred Model 3. All the other numbers are presented as percentages.

between $\hat{\boldsymbol{\lambda}}$ from comparison models and from Model 3. The robustness analysis results are reported in Table 4. It is clearly seen that variabilities in coefficient estimates for all predictors are small and less than 1%, and that coefficient signs of a predictor from comparison models are identical to that from Model 3. In terms of statistical inferences, coefficients of traffic and commercial POI densities are statistically significant in 35 and 39 trials out of 40 while other covariate effects are statistically significant in all models. The distribution of calculated Pearson coefficients has a mean of 0.991 with an interquartile range of $[0.987, 0.994]$, indicating high consistence in estimated charging demand across models. In short, the above results suggest that the estimated covariate effects and latent charging demand are relatively robust and not particular sensitive to sampling variability.

3.4. Scenarios with different number of CP locations

To help with city charging infrastructure planning and management, we conducted a set of computational experiments to examine the changes in spatial patterns of charging stations and demand coverage associated with varying total numbers of CP locations (i.e. $P = 40, 60, \ldots, 400$ in Eq 5). The underlying distribution of charging demand is assumed to be $\hat{\boldsymbol{\lambda}}$ from Model 3 for simplicity.

A MCLP model was implemented to find optimal locations for placing charging stations under each scenario. Spatial non-parametric kernel estimation was employed to depict the continuous distributions of process intensities associated with each set of optimal CP locations (Baddeley et al., 2015). A subset of estimated density maps for optimal CP locations under different scenarios are shown in Fig. 5. As a comparison, the kernel density maps for existing CP locations and the optimal CP locations discussed above ($P = 161$) are also included in Fig. 5. Again, it clearly shows a much higher level of concentration in the spatial pattern of existing CPs than in the spatial pattern of optimal CPs. For optimised locations of charging stations in different scenarios, although high density are still observed at the inner London areas, it gradually expands to the suburban areas. The percentages of charging demand covered by optimal CP locations under each scenario and the marginal increases in demand covered by deploying more CPs are calculated and presented in Fig. 6. In general, the marginal gains of placing more numbers of CPs tend to be decreasing. The policy implication of decreasing marginal gains in demand covered on charging infrastructure planning and management is that an optimal number of charging stations could be found if the cost of deploying a new charging station is known.

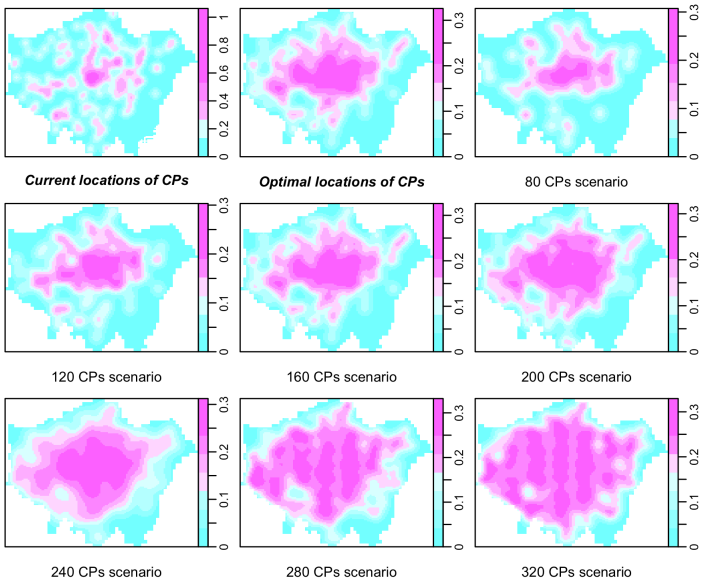

Fig 5. Kernel density maps of the optimal charging stations under each scenario.

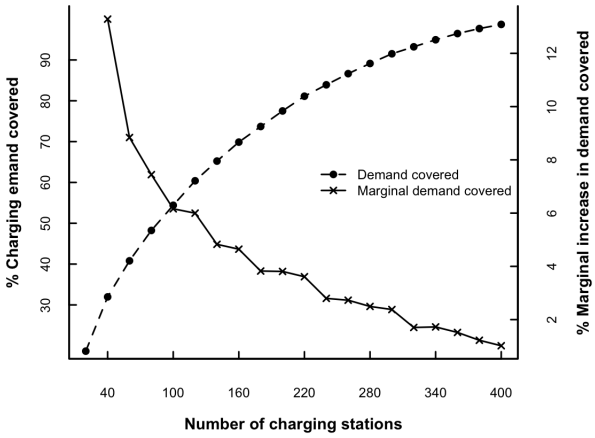

Fig 6. Percentages of total charging demand covered under each scenario.

3.5. Limitations and future work

A spatial point process model is proposed to derive the spatial distribution of charging demand in this study. While it is flexible in terms of allowing for

relevant covariate effects and spatial effects on charging demand to be modelled, a key assumption of equal usage of charging stations is made. Nonetheless, due to charging capacity (e.g. the amount and types of facilities) variations across charging stations, such an assumption may not hold. An important avenue for future work on accurate charging demand estimates is to simultaneously model the location and usage of charging stations with a complex marked spatial point process model (Diggle, 2013; Illian et al., 2013), by collaborating with public and commercial providers of charging stations.

Based on the charging demand derived from a spatial point process model, the second-stage MCLP model identifies optimal locations for deploying charging stations, presented as grid cells with a 1 km $\times$ 1 km resolution rather than precise sites. While a straightforward extension would be adopting a grid topology with a much finer spatial resolution (e.g. 100 metres), it is not feasible due to the huge increase in computational burden of implementing the spatial point process model. An alternative is to identify potential sites for charging stations in a city by selecting the locations of a subset of POIs that can be matched with the current site types of CPs. These potential sites can then serve as the candidate locations for CPs in the MCLP model. This approach shall be explored in our future research. Additionally, the usage of EVs from the perspective of consumers (residents or taxi drivers) and its temporal characteristics (Tu et al., 2016) are not captured in the proposed methodology. Future studies to assess how a further consideration of potential temporal fluctuations in charging demand would affect the optimal deployment of CPs will be needed.

4. Conclusions

This study develops an approach for the EV charging point placement optimisation that brings together spatial statistics and maximal coverage location models. The CP locations are conceptualised as a spatial point pattern driven by an underlying process intensity surface and investigated naturally by using a Bayesian spatial log-Gaussian Cox process model. The model offers great flexi-

bility of capturing potential spatial heterogeneity and dependency effects when identifying potential drivers of the CP locations. Based on the charging demand, estimated as the predicted process intensity surface of CP locations, the optimum placement of CP infrastructure is naturally formulated as a MCLP model with an objective to maximise demand coverage. The methodology is demonstrated by exploring the deployment of CP infrastructure in London. The results show that workplace population and traffic flows are significantly associated with EV charging demand. The densities of transport, retail and commercial POI categories are also statistically significantly linked to the distribution of charging demand in London. Given charging demand estimates from a spatial point process model, optimal locations of new charging stations could be explored timely through a MCLP model, which could be of great benefit to the planning and management of CP infrastructure in a city.

References

Baddeley, A., Rubak, E., Turner, R., 2015. Spatial point patterns: methodology and applications with R. Chapman and Hall/CRC, Boca Raton, FL.

Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics 43, 1–20.

Blangiardo, M., Cameletti, M., 2015. Spatial and spatio-temporal Bayesian models with R-INLA. John Wiley & Sons, Boca Raton, FL.

Butcher, L., Hinson, S., Hirst, D., 2018. Electric vehicles and infrastructure. House of Commons Library, Briefing Paper CBP07480, 1–26. URL: `https://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7480#fullreport`. last accessed: 2018-03-29.

Cai, H., Jia, X., Chiu, A.S., Hu, X., Xu, M., 2014. Siting public electric vehicle charging stations in beijing using big-data informed travel patterns of the taxi fleet. Transportation Research Part D: Transport and Environment 33, 39–46.

Chen, T.D., Kockelman, K.M., Khan, M., et al., 2013. The electric vehicle charging station location problem: a parking-based assignment method for seattle. Transportation Research Record: Journal of the Transportation Research Board 2385, 28–36.

Church, R., ReVelle, C., 1974. The maximal covering location problem. Papers in Regional Science 32, 101–118.

Congdon, P., 2014. Applied bayesian modelling. John Wiley & Sons, Boca Raton, FL.

Diggle, P.J., 2013. Statistical analysis of spatial and spatio-temporal point patterns. CRC Press, Boca Raton, FL.

Dong, G., Ma, J., Harris, R., Pryce, G., 2016. Spatial random slope multilevel modeling using multivariate conditional autoregressive models: A case study of subjective travel satisfaction in beijing. Annals of the American Association of Geographers 106, 19–35.

Ellingsen, L.A.W., Singh, B., Strømman, A.H., 2016. The size and range effect: lifecycle greenhouse gas emissions of electric vehicles. Environmental Research Letters 11, 054010.

He, S.Y., Kuo, Y.H., Wu, D., 2016. Incorporating institutional and spatial factors in the selection of the optimal locations of public electric vehicle charging facilities: A case study of beijing, china. Transportation Research Part C: Emerging Technologies 67, 131–148.

Illian, J.B., Martino, S., Sørbye, S.H., Gallego-Fernández, J.B., Zunzunegui, M., Esquivias, M.P., Travis, J.M., 2013. Fitting complex ecological point process models with integrated nested laplace approximation. Methods in Ecology and Evolution 4, 305–315.

Illian, J.B., Sørbye, S.H., Rue, H., 2012. A toolbox for fitting complex spatial point process models using integrated nested laplace approximation (inla). The Annals of Applied Statistics , 1499–1530.

Ko, J., Gim, T.H.T., Guensler, R., 2017. Locating refuelling stations for alternative fuel vehicles: a review on models and applications. Transport Reviews 37, 551–570.

Law, R., Illian, J., Burslem, D.F., Gratzer, G., Gunatilleke, C., Gunatilleke, I., 2009. Ecological information from spatial patterns of plants: insights from point process theory. Journal of Ecology 97, 616–628.

Lee, D., 2011. A comparison of conditional autoregressive models used in bayesian disease mapping. Spatial and Spatio-temporal Epidemiology 2, 79–89.

Lloyd, C., Catney, G., Williamson, P., Bearman, N., 2017. Exploring the utility of grids for analysing long term population change. Computers, Environment and Urban Systems 66, 1–12.

Martins, T.G., Simpson, D., Lindgren, F., Rue, H., 2013. Bayesian computing with inla: new features. Computational Statistics & Data Analysis 67, 68–83.

Møller, J., Syversveen, A.R., Waagepetersen, R.P., 1998. Log gaussian cox processes. Scandinavian Journal of Statistics 25, 451–482.

Møller, J., Waagepetersen, R.P., 2007. Modern statistics for spatial point processes. Scandinavian Journal of Statistics 34, 643–684.

Murray, A.T., 2016. Maximal coverage location problem: impacts, significance, and evolution. International Regional Science Review 39, 5–27.

Namdeo, A., Tiwary, A., Dziurla, R., 2014. Spatial planning of public charging points using multi-dimensional analysis of early adopters of electric vehicles for a city region. Technological Forecasting and Social Change 89, 188–200.

NCPR, 2017. National charge point registry 2017. URL: `http://www.national-charge-point-registry.uk`. last accessed: 2018-03-29.

Ordanace Survey, 2015. Points of interest: User guide and technical specification. URL: `https://www.ordnancesurvey.co.uk/docs/user-guides/points-of-interest-user-guide.pdf`. last accessed: 2018-03-29.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.R-project.org/`.

Rue, H., Martino, S., Chopin, N., 2009. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. Journal of the royal statistical society: Series B (statistical methodology) 71, 319–392.

Sadeghi-Barzani, P., Rajabi-Ghahnavieh, A., Kazemi-Karegar, H., 2014. Optimal fast charging station placing and sizing. Applied Energy 125, 289–299.

Schumacher, D., 2017. ompr: Model and Solve Mixed Integer Linear Programs. URL: https://CRAN.R-project.org/package=ompr. r package version 0.7.0.

Shahraki, N., Cai, H., Turkay, M., Xu, M., 2015. Optimal locations of electric public charging stations using real world vehicle travel patterns. Transportation Research Part D: Transport and Environment 41, 165–176.

Simpson, D., Illian, J.B., Lindgren, F., Sørbye, S.H., Rue, H., 2016. Going off grid: Computationally efficient inference for log-gaussian cox processes. Biometrika 103, 49–70.

Snyder, S.A., Haight, R.G., 2016. Application of the maximal covering location problem to habitat reserve site selection: a review. International Regional Science Review 39, 28–47.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64, 583–639.

Taylor, B., Davies, T., Rowlingson, B., Diggle, P., 2015. Bayesian inference and data augmentation schemes for spatial, spatiotemporal and multivariate log-gaussian cox processes in r. Journal of Statistical Software 63, 1–48.

Transport for London, 2018. Electric vehicle charging infrastructure: Location guidance for london. URL: http://content.tfl.gov.uk/electric-vehicle-charging-infrastructure-location-guidance-for-london.pdf. last accessed: 2018-03-29.

Tu, W., Li, Q., Fang, Z., Shaw, S.l., Zhou, B., Chang, X., 2016. Optimizing the locations of electric taxi charging stations: A spatial–temporal demand coverage approach. Transportation Research Part C: Emerging Technologies 65, 172–189.

Wagner, S., Götzinger, M., Neumann, D., 2013. Optimal location of charging stations in smart cities: A points of interest based approach. In: 34th Int Conf Inf Syst (ICIS 2013) , 1–18.

Wei, R., 2016. Coverage location models: Alternatives, approximation, and uncertainty. International Regional Science Review 39, 48–76.

Wei, R., Liu, X., Ou, Y., Fayyaz, S.K., 2018. Optimizing the spatio-temporal deployment of battery electric bus system. Journal of Transport Geography 68, 160–168.

Xi, X., Sioshansi, R., Marano, V., 2013. Simulation–optimization model for location of a public electric vehicle charging infrastructure. Transportation Research Part D: Transport and Environment 22, 60–69.