

ORIGINAL RESEARCH

Parking space inventory from above: Detection on aerial images and estimation for unobserved regions

Jens Hellekes¹  | Ariane Kehlbacher²  | María López Díaz²  | Nina Merkle¹  |
Corentin Henry¹  | Franz Kurz¹  | Matthias Heinrichs² 

¹German Aerospace Center (DLR), Remote Sensing Technology Institute, Wessling, Germany

²German Aerospace Center (DLR), Institute of Transport Research, Berlin, Germany

Correspondence

Jens Hellekes, Muenchener Str. 20, 82234 Wessling, Germany.

Email: jens.hellekes@dlr.de

Funding information

German Aerospace Center (DLR)

Abstract

Parking is a vital component of today's transportation system and descriptive data are therefore of great importance for urban planning and traffic management. However, data quality is often low: managed parking places may only be partially inventoried, or parking at the curbside and on private ground may be missing. This paper presents a processing chain in which remote sensing data and statistical methods are combined to provide parking area estimates. First, parking spaces and other traffic areas are detected from aerial imagery using a convolutional neural network. Individual image segmentations are fused to increase completeness. Next, a Gamma hurdle model is estimated using the detected parking areas and OpenStreetMap and land use data to predict the parking area adjacent to streets. A systematic relationship is found between the road length and type and the parking area obtained. It is suggested that these results are informative to those needing information on parking in structurally similar regions.

1 | INTRODUCTION

Many applications in urban planning and traffic management require information on the amount of space dedicated to parking at the individual street level. For instance, transport modelling relies on parking supply data to analyse parking choice behaviour and the impact of policy measures. This is particularly relevant in the discussion about re-purposing on- and off-street parking [1]. Parking supply is characterized by a number of variables, including location and number of parking spaces, as well as type of parking and ownership. However, data on parking are scarce or inconsistent. Ref. [2] attributes the lack of parking data to administrative and bureaucratic as opposed to technical barriers. To fill this gap, several approaches have been explored: Ref. [3] uses crowd-sensing from car sensors to map out parking space. The effort required by this approach is considerable, since cars have to drive through every street in the area of interest. Ref. [4] utilizes a LiDAR sensor system in a public facility to obtain parking space conditions. Ref. [5] applies classification methods to images captured by fixed cameras in order to estimate the occupancy status. A prediction of parking lot occupancy is performed by [6] based on data from

smart parking meters. In these approaches, data analysis is limited to those areas covered by the respective sensor technology. On the user side, multiple datasets must be combined to characterize the citywide parking supply, as shown in [7]. They did not, however, include data on the surrounding area and therefore did not account for parking capacity constraints outside of the city. This issue was intended to be addressed at a later date by using additional survey data. Ref. [8] obtains information on parking supply using data from an origin-destination survey by taking the maximum number of parked cars at destination as a proxy for parking capacity. A different approach is taken by [9] for setting up an agent-based parking model: the total number of on-street parking spaces is estimated assuming that parking can occur along entire road segments using a fixed distance between parking lots. Spatial characteristics and restrictions are taken into consideration such as buildings having private garages or parking in driveways not being allowed.

The contribution of this study to the literature on parking space inventory is that parking spaces are automatically mapped and predictions on parking areas can be made for structurally similar regions. We develop a new approach which combines features from aerial imagery with road network information and

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial License](https://creativecommons.org/licenses/by-nc/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *IET Intelligent Transport Systems* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

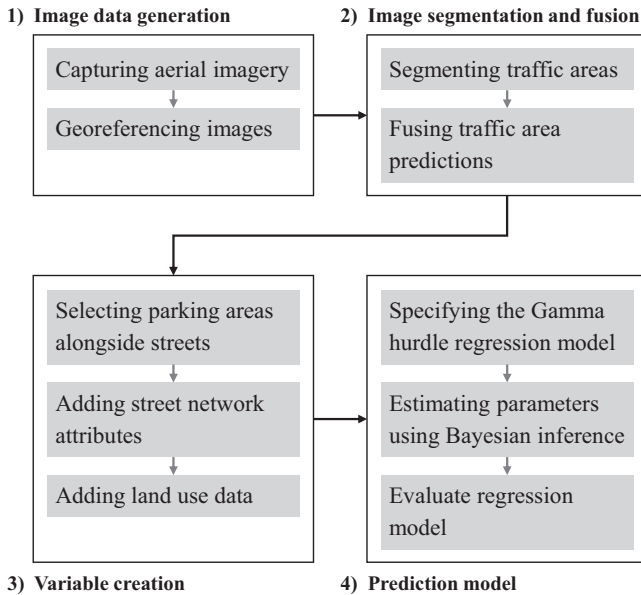


FIGURE 1 Overview of our approach which involves four sequential steps

land use data. Remote sensing has the advantage that it provides a large-scale and complete overview of land use in the area of interest. Moreover, areas inaccessible to ground surveys are captured, and creating training data for subsequent automated image segmentation is less costly than cadastral measurements on the ground. First, a deep neural network is applied to aerial images to detect parking areas. The segmentation results are fused to increase the robustness of detections. Next, the detected parking areas are combined with information on the road network from OpenStreetMap (OSM) and land use data. Variables are extracted from blending the data and then used to estimate a Gamma hurdle model to predict on-street parking area using Bayesian inference. The uncovered relationships between street characteristics and land use and detected parking areas make sense, and predictive accuracy of the model is good. The workflow is depicted in Figure 1.

2 | IMAGE SEGMENTATION AND FUSION FOR MAPPING PARKING AREAS

To detect parking, high-resolution remote sensing data are required. Often, this imagery is captured at periodic intervals through aircraft overflights on behalf of public authorities, and partly released for use free of charge (e.g. 10 cm ground resolution for the state of North Rhine-Westphalia, Germany). This imagery may be sufficient for an initial mapping of parking areas. In this analysis, aerial images from six flight campaigns over Brunswick (Germany), taken by DLR 3K camera system [10] in 2019 and 2020, are used. The acquisition scheme included two flights each in spring, summer, and fall, with a morning and afternoon flight at each time. This ensured that seasonal and daily traffic patterns on working days are accounted for in the data set, as well as different occlusions

by vegetation, varying levels of illumination, and weather conditions. Urban, sub-urban and industrial areas of Brunswick (about 40 km²) were captured twice during each flight with a ground resolution of about 9 cm per pixel. Due to the acquisition frequency of 1 Hz and a resulting overlap of successive images of about 80%, scenes were captured up to 60 times in total. In combination with Global Navigation Satellite System (GNSS) and inertial data, a digital elevation model (DEM) and ground control points, the imagery is precisely georeferenced through bundle adjustment.

2.1 | Segmentation of traffic areas in aerial images

The most successful tools for performing semantic segmentation nowadays are neural networks. Contrary to traditional methods, where features of interest are extracted using hand-crafted detectors, neural networks are capable of automatically learning which are the best suited detectors for the task at hand. From a remote sensing perspective, current research on stationary traffic mainly addresses the detection of marked parking lots [11–13] and their utilization based on surveillance cameras as well as airborne and satellite imagery [14–16]. For a most complete parking map, data is needed not only for dedicated parking spaces but also on areas that are regularly used to park a vehicle (e.g. on-street, in backyards, on factory premises). Since the mapping of such dual-use areas has not been investigated before, we created a new dataset composed of 47 non-overlapping aerial images, with a size of 5616 × 3744 pixels acquired over Brunswick and covering a total area of about 10 km². For each of the image fine-grained manual annotations for the following four classes were generated: access ways (total surface of 0.70 km²), roads (0.70 km²), parking areas (0.42 km², officially dedicated area or used for parking at the time of recording), and background. In contrast to the road network with its connecting function, access ways only serve to link the network to specific destinations, for example, parking lots. This functional distinction is useful for separating on-street and off-street parking areas. The annotations are considered as ground truth because they partition visible areas according to the defined classes.

In this paper, we use a neural network called Dense-U-Net [17] for the task of traffic area segmentation. This network is derived from U-Net [18], a widely successful fully-convolutional architecture especially used in remote sensing. It consists of two consecutive parts: an encoder and a decoder. The aim of the encoder is to extract all the relevant features from the input data. The decoder on the other hand, takes the features as input and interprets them as object classes pixel-wise. In order to extract more spatial and semantic information from the low-level and high-level layers respectively, several layers of the encoder are directly connected to the decoder through so-called skip-connections. The main advantage of Dense-U-Net over similar U-Net architecture is its commonly used densely-connected backbone architecture for which pre-trained weights are easily available, and its replacement of the

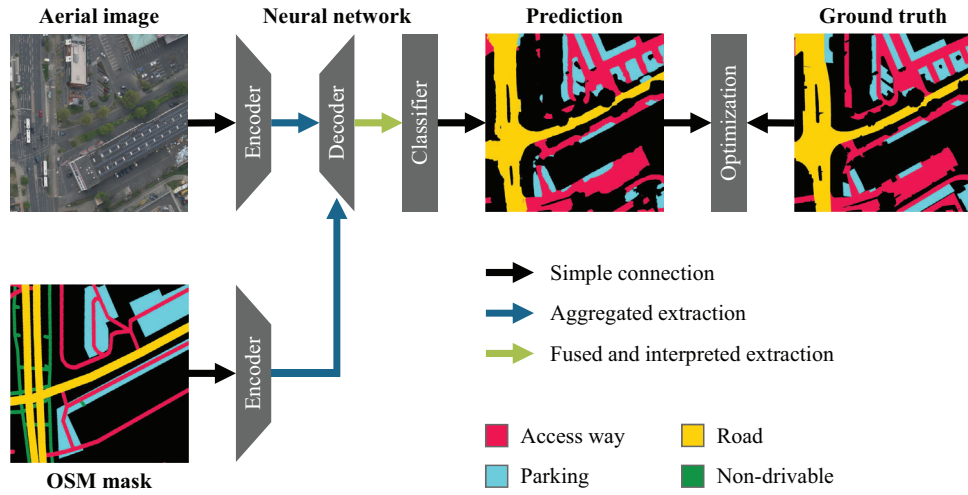


FIGURE 2 Illustration of the training pipeline for the aerial and OSM fusion network

typical shallow decoder architecture with densely-connected layers mirroring those of the encoder. It was shown in [17] that it leads to more accurate segmentation results than competing architectures when fine-grained details recovery is important.

To extract as much information as possible from the image, we additionally exploit a free available data source, namely OSM and integrate these data in our approach. Although OSM data are sometimes spatially and semantically inaccurate, they provide a good baseline for distinguishing roads, access ways and parking areas. Therefore we take inspiration from the fusion strategy of FuseNet [19] and implement it within our Dense-U-Net architecture: the main (here RGB) and auxiliary (here OSM) input images from the two data sources are fed into two separate encoders of the same architecture but not sharing weights, and the feature maps of the auxiliary encoder are merged by addition to the corresponding feature maps of the main encoder, and used by the next layer of the main encoder. As shown in our previous study, the slightly different merging scheme Skip-FuseNet [20] is more effective for RGB and OSM fusion: the added feature maps from the main and auxiliary encoders are not fed into the next main encoder layer, rather they are used as inputs for the skip-connections to the decoder. This separation reduces the risk of the network being confused during training, as the features extracted in both branches are likely to be considerably different. To prepare the OSM data to be used as input to the neural network, we extract pixel-wise masks for seven traffic-related objects categories (drivable and non-drivable ways, access ways, parking spaces, gas stations, bicycle parking and parking vending machines). An overview of the network and of the input and output data is shown in Figure 2.

An essential step when dealing with neural networks is the training phase. In this phase, the network is trained to solve the desired task by running the network over the input images and comparing its outputs (the so-called predictions) with the ground truth using a loss function. Afterwards, the computed loss is “back-propagated” through the network to optimize it so that the next iteration makes more accurate predictions. The data set is split into three subsets: (1) the training set

consisting of 35 images, which is only used during the training of the network, (2) the validation set consisting of five images, which is used to evaluate the network’s performance during the training and (3) the test set consisting of 7 seven images, which is used to evaluate the performance and generalization capacities of the network after the training. This final step ensures that the network generalizes well to images unseen by both the network and the operator monitoring the training, to avoid overfitting the network on training or validation data.

For the training of the network, we crop our training images to smaller patches with a size of 512×512 pixels and perform the training process over these input images over 100 iterations using a cross-entropy loss, an Adam optimizer and a learning rate initialized at 10^{-4} decaying at a rate of 0.937 after each epoch. The weights of the RGB encoder are initialized with weights pre-trained on ImageNet, and the RGB input data is normalized according to the recommendations of the PyTorch library. The other layers’ weights are initialized randomly with a Xavier uniform distribution and their biases are initialized to 0. Afterwards, we perform the final evaluation on the test set. The predictions output by the network are segmentation maps, that is, images where each single pixel is classified. This classification results from a maximum-voting among all four classes, where the class with the highest predicted probability is assigned to each pixel. Upon visual inspection, we establish that the network accurately delineates parking areas, roads and access ways (see Figure 4 for examples). The areas predicted for all classes smoothly and precisely follow the visible object boundaries in the input images, false detections are rare and the extraction complete in most regions. The most common type of confusion appears to be between roads and access ways, and therefore a minor concern since our main objects of interest are parking lots. To provide a comparable baseline, we provide a quantitative performance evaluation in Table 1. Three usual metrics are listed there: (1) the Intersection over Union (IoU), which measures the accuracy of our network by quantifying the percentage of overlapping pixels between the ground truth and our predictions, (2) precision, which is highest if the network extracted

TABLE 1 Quantitative performance of the neural network for the segmentation of parking areas, roads, and access ways for the city of Brunswick, Germany

	IoU (%)	Recall (%)	Precision (%)
Parking	68.16	78.22	84.12
Road	77.44	84.68	90.05
Access way	64.40	75.69	81.19

only correct objects (i.e. no false positive), and (3) recall, which is highest if the network missed no object (i.e. no false negative). The high precision scores for all three classes ($\approx 84\%$ for the parking areas) prove that our network is conservative and therefore is not likely to wrongly detect objects. We can see, however, that despite high recall scores, the completeness of our predictions could be further improved, especially for the parking class ($\approx 78\%$). This is mostly visible under tree shadows and on factory premises. For more details about the training setup, the analysis and discussion of the results as well as a comparison of training configurations (e.g. with and without OSM masks and with and without the fusion scheme), we refer to our paper [20].

2.2 | Fusion of traffic area predictions

We expect an improved completeness of segmented traffic areas when the predictions on overlapping aerial images are combined: changing acquisition conditions (e.g. illumination, occlusion, viewing angle, weather) result in altered network predictions and lead to more reliable classifications through mutual comparison. Ref. [21] proposes using another neural network for this task, which utilizes a baseline prediction as a prior mask and generates a refined prediction for the same area with a slightly different configuration. To enhance classification accuracy, ref. [22] combines the results of three neural networks with a naive Bayesian fusion layer.

To detect traffic and especially parking areas on a citywide basis, two main challenges arise for the fusion of predictions: first, the requirement to detect areas that are regularly used for on-street parking implies that these areas might also be temporarily available for moving traffic. To classify an area in this regard, intertemporal comparisons are necessary and the correct assignment to a class may change based on the observed traffic situation during a flyover. The main purpose of a traffic area is approximated by using the segmentation results as a proxy for non-observed periods. Second, the regression model estimation conducted below requires a parking area inventory that offers a high level of heterogeneity in terms of spatial and infrastructural characteristics in order to facilitate transferability to other regions in the future. For the study area of Brunswick, these requirements translate into several thousand aerial images that must be fused with a high degree of interdependency due to the overlaps and informative content inherent in each prediction.

Classical remote sensing methods for mosaicking, such as a classification by sum of probabilities and/or passing thresholds, do not adequately account for known pixel-wise

power compared to other image-wise predictions of the same scene. Therefore, we propose a novel approach that estimates per-pixel confidence intervals of the class probabilities via the statistical method of bootstrapping and selects the resulting class with a classifier. For this, the probability predictions per class for all aerial images are retrieved by the previously described aerial and OSM fusion network. By projecting the individual predictions on the DEM, a pixel-wise stack of probabilities for the same scene can be obtained. An empirical/basic bootstrapping as described in [23] is used to produce a confidence interval for the class probability's true value. The classifier then returns the class with the highest upper limit of the corresponding confidence interval. Figure 3 visualizes the approach, simplified with six instead of about 60 prediction layers used for generating the bootstrapped segmentation map of traffic areas in Brunswick. The method serves as a framework within which the construction of the bootstrapping procedure and the classifier can be chosen. Subsequent work could examine whether the assumption of independently and identically distributed predicted class probabilities is warranted, and the bootstrapping technique could be adjusted accordingly, depending on the data used.

Comparing the single with the fused predictions for two exemplary aerial images of the test set in Figure 4 shows an overall tendency for a better matching between the fused predictions and the ground truth. It should be noted that the ground truth was created on the corresponding aerial image of the individual prediction depicted, so the heterogeneous combination of all acquisition conditions in the fused prediction is only partially reflected by it. This insufficient representation is most evident in Panel 4(II), where in the majority of aerial images the parking area is obscured and therefore does not appear in the fused prediction. The strength of the approach is apparent in Panel 4(III): here, the single prediction exhibits high uncertainty (indicated by the granular segmentation), while the fusion process correctly classifies the pedestrian walkway as background. An incorrect classification as parking area due to a vehicle in a loading zone (Panel 4(IV)) is also corrected in the fused prediction. If — as seen in Panel 4(VI) — a residential road is frequently used for on-street parking, the fusion process reveals this. By combining the numerous observations, objects that show great similarity to a parking lot on aerial imagery can be ruled out by fusing the predictions (see temporarily placed containers in Panel 4(VII)). It can also be seen in this image cutout that bootstrapping causes only regularly used areas to be classified as parking spaces. Since all flights took place during the vegetation period, the street in Panel 4(VIII) can be better but not completely identified by fusion.

3 | ADDITION OF FURTHER VARIABLES

From the aerial image segmentation and fusion process all visible parking areas are vectorized. However, for our approach only on-street parking spaces are taken into consideration. So, surface parking areas are left out for further analysis and only

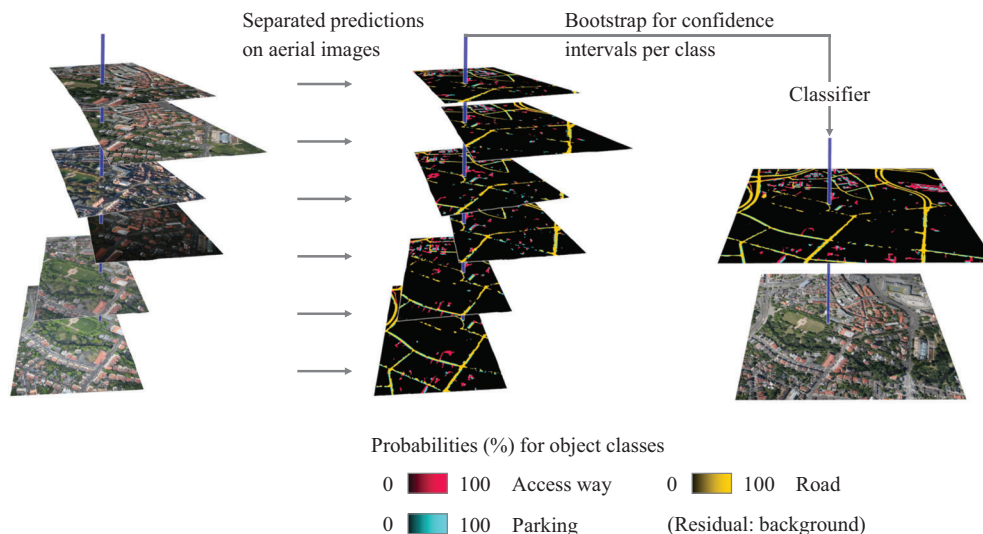


FIGURE 3 Illustration of the scheme for fusing individual predictions on overlapping aerial images into a single segmentation map

parking spaces alongside roads, either on the roadside or on parking lanes (see Figure 5), are taken into consideration. To create a model that makes possible to estimate the parking area of regions that have a comparable infrastructure, we need additional information. This information can be spatially joined to the detected parking areas using additional data sources. Variables required for our model are detected parking area (in square metres), road length (in metres) and road type, extracted from OSM, and land use, sourced from the German digital landscape model by [24].

3.1 | Addition of street data

The OSM road network is mainly represented by line features and identified by the key highway and a value designating different types of roads, ways and paths, such as service, residential or primary. Roads that belong together are sometimes divided into segments in OSM. In order to have more realistic road lengths, adjacent segments of the same type are linked together. Only drivable roads (according to the OSM definition) are considered for analysis, thus leaving out some OSM path types such as footways or steps. Road line features are buffered and converted to polygons. This makes possible to map parking areas onto the buffered roads (Figure 5). Subsequently parking areas are aggregated by road, for which total length and length within the study area are calculated.

3.2 | Addition of land use data

Further, land uses are represented by polygon features and cover the study area continuously. Examples of land uses are industrial, residential or forests. To combine road and land use data firstly road width by road type is estimated visually and used as buffer distances. The buffered roads are subsequently intersected with land use data and percentages of land

uses are calculated for each road. This results in each road having a percentage for each of the land uses. Road width estimates are not strictly necessary, being possible to use an average width for all road types. However, in case of fine-grained land use data, accuracy may be increased. The land uses are condensed to two categories. On the one hand, land uses where parking is more common and on the other hand, land uses for which the occurrence of parking areas is expected to be less likely (what we later call land use with less parking) such as wooded areas, wild vegetation areas, water surfaces, railway related areas, bridges or agricultural areas. Since the first category shows little variation, for the model only the sum of the percentages of the land uses within the second group is used as a predictor.

4 | SPECIFICATION, ESTIMATION AND EVALUATION OF THE MODEL THAT PREDICTS PARKING AREA PER STREET METRE

We use Bayesian inference to estimate our model which requires the specification of the full probability model for the data, via the likelihood function, and parameters, via prior distributions [25]. The advantage of this approach is that uncertainty is explicitly accounted for. This is important in the case of hierarchical models, where parameter uncertainty needs to be transmitted between hierarchies and it is useful when results are fed into further applications such as agent-based travel demand models.

4.1 | Formulation of the used model

The dependent variable y in our model is the detected parking area (in square metres) per metre of road length. There are $N = 3443$ roads, as defined in Section 3.1, with mean road

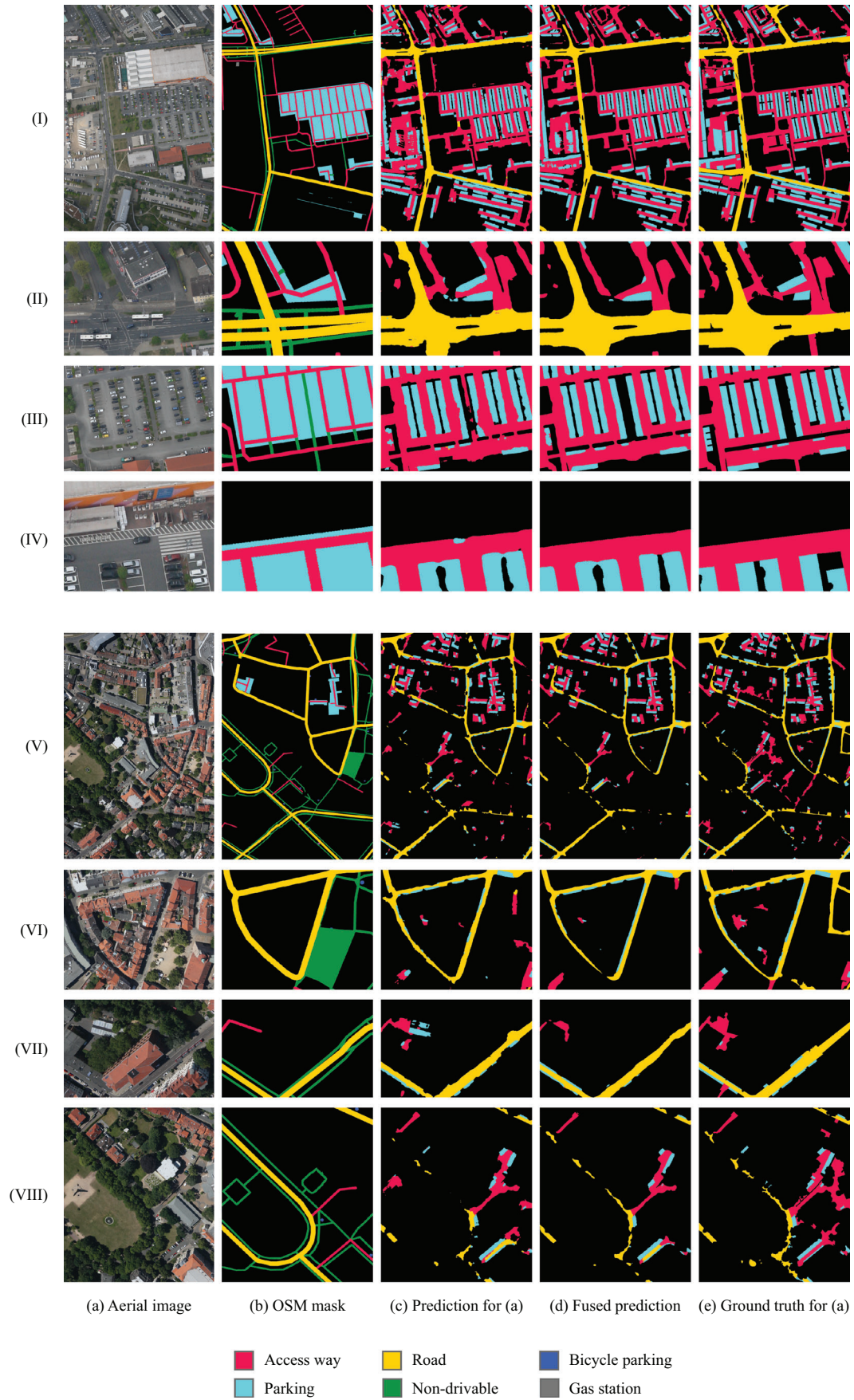


FIGURE 4 Qualitative comparison between predictions on single aerial images, the fused predictions, and the ground truth

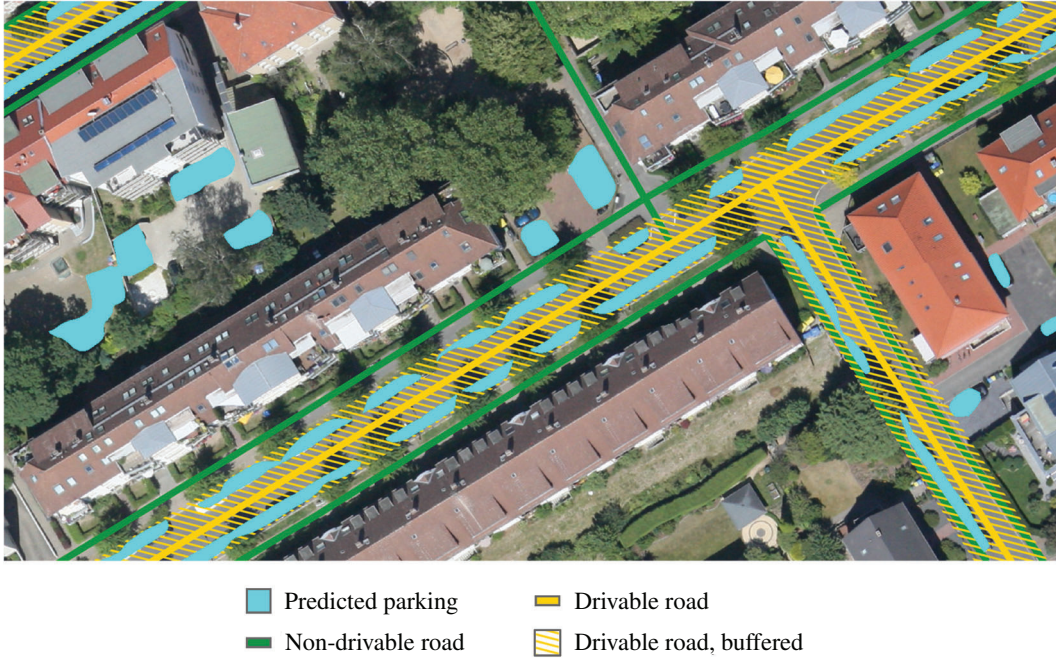


FIGURE 5 Cutout of an aerial image with the OSM road network and detected parking areas

length of 130 m, and median, mean and maximum parking area per metre of $0.2 \text{ m}^2/\text{m}$, $1.3 \text{ m}^2/\text{m}$ and $19.5 \text{ m}^2/\text{m}$, respectively. The distribution of detected parking area per metre is highly skewed and non-normal with a large proportion of zeros (38%). A hurdle specification is chosen to model the data generating processes of the zero observations and the positive observations separately. Various specifications involving different types of probability distributions for the positive observations (log Normal, Gamma) and variables were compared using leave-one-out cross-validation [26]. While the log Normal and Gamma distribution are both continuous probability distributions restricted to the positive domain, the Gamma distribution has slightly fatter right tails, which means it is better suited to modelling extreme values. The most preferred model based on model fit criteria is a hierarchical Gamma hurdle regression. For details on this type of model in general the reader is referred to Section 5.6 in [27]. The probability of whether a road has zero detected parking area is modelled via a Bernoulli distribution. This is a discrete probability distribution for random variables that can take only one of two possible outcomes. Thus, the hurdle equation is specified as follows

$$\pi_i = \alpha_{00} + \beta_0 x_i \quad (1)$$

where $\pi_i = \text{Prob}(y_i = 0)$, and α_{00} and β_0 are parameters to be estimated, and predictors x_i are log road length in metres and the percentage of land use where parking is less likely. The distribution of the positive continuous values of the dependent variable y_i is specified as a Gamma such that

$$\mu_i = \alpha_{10} + \alpha_{1,s[i]} + \beta_1 x_i + \varepsilon_i \quad \varepsilon_i \sim \Gamma(\lambda) \quad (2)$$

where α_{10} is the intercept; $\alpha_{1,s[i]}$ detects systematic departures from α_{10} due differences in detected parking area per street metre in road types, where $s[i]$ denotes the road type $s = 1, \dots, S$ of observation i where $S = 9$ as explained in Section 3.3. The parameter vector β_1 measures differences in the positive continuous detected parking area per street metre due to variation in the predictors, which are the same as in the hurdle equation. The random error ε_i captures unobserved factors influencing detected parking area and follows a Gamma distribution with shape parameter λ . Combining (1) and (2) gives the following hurdle likelihood function

$$p(y_i | \theta_i) = \begin{cases} \frac{1}{1 + \exp(\pi_i)} & \text{if } y_i = 0 \\ \frac{1}{1 + \exp(-\pi_i)} \cdot \frac{\gamma(y_i | \mu_i, \lambda)}{1 - \gamma\text{CDF}(0 | \mu_i, \lambda)} & \text{if } y_i > 0 \end{cases} \quad (3)$$

where θ_i is the set of all parameters to be estimated. To obtain the joint posterior distribution, which is our object of inference, we combine the likelihood in (3) with the following prior distributions for each parameter

$$\begin{aligned} \alpha_{00}, \alpha_{10}, \beta_0, \beta_1 &\sim \mathcal{N}(0, 5) \\ \alpha_s &\sim \mathcal{N}(0, \tau_s) \quad \forall s = 1, \dots, S \\ \tau_s &\sim t_+(3, 0, 2.5) \\ \lambda &\sim \Gamma(0.01, 0.01) \end{aligned} \quad (4)$$

where the standard deviation τ_s denotes the road type level errors. Estimation of the joint posterior is performed in the probabilistic programming language Stan using Markov Chain Monte Carlo sampling via adaptive Hamiltonian Monte Carlo

TABLE 2 Estimates of the posterior means of the model parameters with 95% credible intervals (CI) and Rhat values

	Estimate	1–95% CI	u–95% CI	Rhat	Bulk ESS	Tail ESS
Gamma component						
shape λ	0.73	0.69	0.76	1.00	10,988	7716
intercept α_{10}	1.46	0.84	2.02	1.00	3847	4645
sd(intercept) τ_5	0.48	0.19	1.10	1.00	2769	4888
ln(metres) β_1	−0.26	−0.32	−0.20	1.00	10,415	7305
% land use with less parking β_1	−0.00	−0.01	0.00	1.00	16,931	7091
Hurdle component						
intercept α_{00}	3.51	3.15	3.87	1.00	11,285	7503
ln(metres) β_0	−0.98	−1.06	−0.89	1.00	10,231	7314
% land use with less parking β_0	0.03	0.02	0.04	1.00	12,689	7258

[28], which is an approximate Hamiltonian dynamics simulation based on numerical integration that is then corrected by performing a Metropolis acceptance step. Implementation of our model is done using the R package for Bayesian regression modelling “brms” version 2.15 [29], which is based on Stan. Samples from the marginal posterior distributions of the parameters are drawn using the Hamiltonian Monte Carlo No U-turn sampler [30] with four independent Markov chains with 5000 warm-up iterations and 5000 sampling iterations. Convergence of the chains is monitored visually via traceplots and in terms of the split potential scale reduction factor (Rhat) which compares between-chain parameter estimates and takes a value of 1 at convergence.

4.2 | Parameter estimates

Estimates of the means of the posterior distributions and their 95% credible intervals are presented in Table 2. The potential scale reduction factor Rhat, which measures the ratio of the average variance of samples within each of the four chains to the variance of the pooled samples across chains, is 1. This suggests that all chains are comparable to each other and have converged to the same level. Also, bulk and tail effective sample sizes (ESS) are large, indicating that estimates are reliable as they are based on a large number of independent samples from the posterior distribution. The Gamma distribution is described by its shape parameter of 0.73 and the scale parameter intercept, which decreases as the natural log of metres of street length, ln(metres), increases. This means that as expected the distribution is highly skewed and that as a street gets longer, the amount of detected parking area per street metre decreases. The parameter sd(intercept) is the estimated standard deviation of the varying intercept. It suggests that some of the variation in detected parking area per street metre is associated with differences in road types. Accordingly, the posterior distributions of these effects in Figure 6 show that service and residential roads tend to be associated with large values of detected parking areas per street metre, whereas primary and secondary roads show a slight, albeit uncertain, negative association. The parameters in the hurdle components in Table 2 are the log odds. They

indicate that on average longer streets have a lower probability of having no parking areas as a 1% increase in street length is associated with almost equally sized decrease in the probability of observing zero detected parking area. Also, as % land use with less parking increases for a given street, the average probability of observing zero detected parking area increases, albeit by a very small amount. The conditional effects of the predictor’s parameters are depicted in Table 2 and shown in Figures 7 and 8. Figure 7 suggests the probability of zero detected parking areas per street metre decreases as road length increases and increases as the value of % land use with less parking increases. Figure 8 shows the positive effect of increases in road length on predicted detected parking area, which is strongest at moderate road lengths. It also shows that as the percentage of land use less likely to be associated with parking increases, its effect on detected parking per street metre decreases. This means that as a road is increasingly surrounded by land uses such as wooded areas that are usually not associated with parking, the model predicts less detected parking area, which makes sense.

4.3 | Evaluation of predictive performance

The out-of-sample predictive accuracy of our model is evaluated using Pareto smoothed importance sampling cross-validation [26], an approach that involves fitting a Pareto distribution to the tails of the 20% largest importance ratios. If the resultant estimate of the shape parameter of the Pareto distribution, k , is larger than 0.7 this would be reason for concern (see ref. [26]) as it indicates a data point that is highly influential to the posterior distribution and therefore has the potential to negatively affect future predictions. Figure 9 shows that estimates of k for all data points are less than or close to 0.5, indicating that our model’s out-of-sample prediction accuracy is high.

As part of our model checking process we sample 20 values from the posterior distribution and plot the respective predictions, y_{rep} , against the data, y , in Figure 10. The upper panel suggests that overall the predicted posterior density follows the shape of the data. The lower panel shows more detail for the interval $[0, 5]$ on the x -axis. It demonstrates

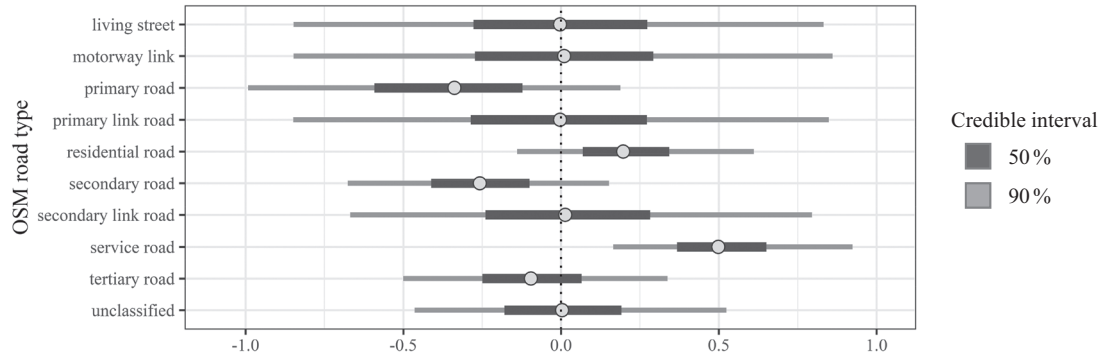


FIGURE 6 Posterior distributions of road type specific effects $\alpha_{1,[j]}$

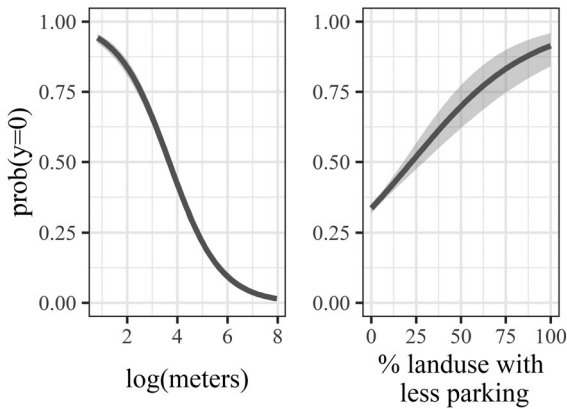


FIGURE 7 Conditional effects of predictors (at the mean) of the hurdle model component β_0

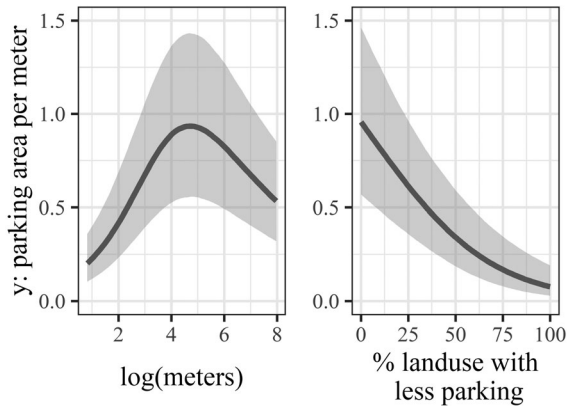


FIGURE 8 Conditional effects of predictors (at the mean) of the Gamma component β_1

that at low magnitudes of detected parking areas per metre, uncertainty is much increased. That is, predictions at low levels of parking areas are associated with higher levels of uncertainty.

Investigating further, we compute the average predictive error $y - y_{rep}$, based on 100 draws from the posterior distribution. Overall the bulk of the average predictive errors of each observation is close to zero, with a mean of 0.03 and a median of -0.7 . The latter is negative because for some observations the

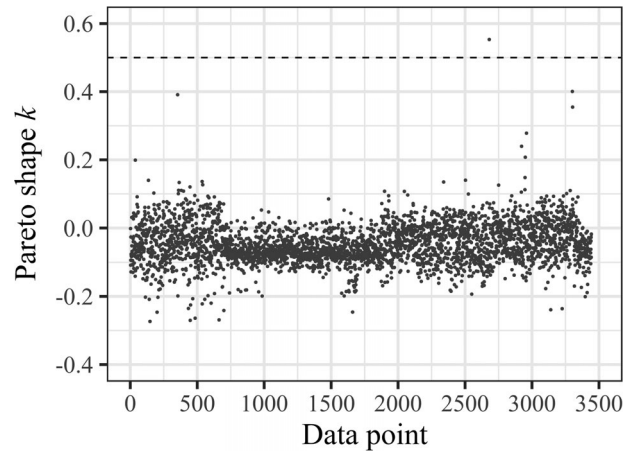


FIGURE 9 Pareto smoothed importance sampling diagnostic plot

model predicts positive values for the detected parking areas, which are in fact zero. Investigating further, we plot it against the data y in Figure 11. This shows that the average predictive error systematically increases as y increases, suggesting that as the magnitude of the detected parking area increases, the model tends to under-predict the size of the detected parking areas. However, this is not too surprising as the bulk of the observations are near one, the mean of y is 1.28, and at this range the average predictive error is close to zero, suggesting that for the majority of observations the model's predictive performance is good.

5 | DISCUSSION

This paper presents a novel approach to predicting parking areas at the individual street level using aerial imagery, OSM road network information, land use data and Bayesian regression modelling. One advantage of aerial imagery is capturing a large study area in a short time, which gives a consistent dataset for the time of the flight. Compared to existing approaches, the proposed processing chain not only provides data on dedicated parking areas on publicly accessible ground. Those on private property are also mapped. Additionally, the fusion step identifies those places that are regularly occupied by parked vehicles

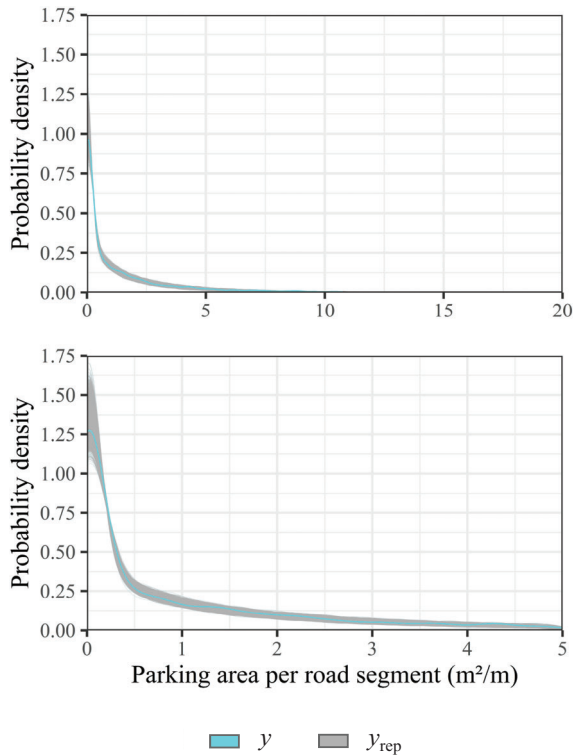


FIGURE 10 Posterior distributions of the predicted parking area y_{rep} against the smoothed data y . The upper panel depicts the entire distribution. The lower panel zooms into the peak of the distribution to better show the increase in uncertainty for values close to zero

but not marked as parking lots. Taken together, this increases the completeness of the inventory. Furthermore, georeferencing of aerial images is much more precise than using imagery from moving cars. This makes it easier to link the features extracted from remotely sensed imagery to other sources like OSM. Since the features used in the Gamma hurdle model to describe the road network and land use are available for many geographic regions, predictions on parking area at the street level can be made for those areas lacking aerial imagery but that are structurally comparable.

Our approach of segmenting traffic areas in aerial images is already critically discussed [20, 31]. In the context of this study, it is important to consider that the data comprise detected parking areas, which are generally expected to be smaller than the actual areas. One reason is occlusion and shadowing caused by buildings and vegetation; another is that the ground truth data, which are used to train the neural network for the segmentation of traffic areas in individual aerial images, classifies a pixel as a parking area only, if it can be unambiguously recognized as such. Since smaller predicted parking areas often reflect on-street parking and designated parking areas are typically clearly visible from aerial imagery, we expect under-detection to decrease as parking area size increases. This hypothesis is supported by the fusion scheme: an area that is not exclusively dedicated to parking traffic is only classified as parking area if it has a vehicle on it. If in the majority of the observations there are no vehicles parked on such a multi-use area, the area may be classified as road or access way by the fusion scheme.

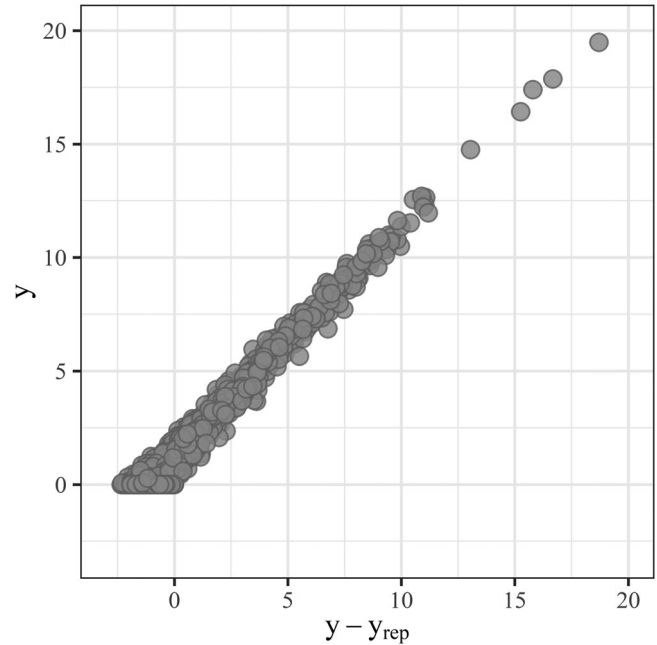


FIGURE 11 Average predictive error of the Gamma hurdle model

The process of variable extraction also presents various aspects that need to be considered. First, accuracy and completeness of both data sources, but especially of OSM, as it is a volunteered geographic information project, are limited. An assessment of data quality is difficult as no official database is available for comparison. Still, investigations show that the quality of OSM data is high in regions with a high population density and with numerous OSM contributors [32, 33], which is the case for our study area. Second, measurement errors for road length may arise, if roads consist of several lines. After using the GIS linking function, around 60% of line segments were merged, but the process fails, if the endpoints of lines do not coincide. Shorter roads may therefore be slightly over-represented in our data. Third, when buffering roads to combine them with land use data, comparing line features with aerial images to estimate road widths shows that not all lines are located in the middle of a road leading to buffered roads being misplaced. There are also roads that have varying widths, for example, due to different number of lanes in different sections or irregular roadside parking. In this case, the overlapping areas of the buffered roads with the land use polygons may be displaced. Moreover, the width of residential roads varied a lot within the study area, so that an average value is chosen. The implication for our data is, that there is some uncertainty about the correct assignment of land use, which is one of the variables in the model.

In the following we focus on the discussion of the predictive model for detected parking areas. First and foremost, the parameters of Bayesian hurdle Gamma regression model make sense and support the initial hypothesis: longer streets are associated with higher predicted parking areas and with a lower likelihood of observing zero values; and higher percentages of land use like wooded areas, agriculture and water surfaces are associated with smaller parking areas and a higher probability of observing

zero values. The parameters associated with variation in parking area due to different street types also make sense. Service and residential roads are associated with larger detected parking area per metre of road length. This results also reflects the fact that these two road types make up 64% and 22% of the data, respectively. Instead, parameters of street types with few observations and/or little association are shrank towards zero by the hierarchical model [34].

Looking at the predictive accuracy of the detected parking area model in Figures 10 and 11, we find that accuracy varies with the size of the detected parking area per metre. At mid to large ranges, performance is quite satisfactory, but at zero or small values prediction uncertainty increases. Thus, predictions for short roads and/or a high percentage of land use not for parking are associated with more uncertainty as zero or small detected parking areas are expected. Many observations fall into this lower range on the one hand because this is the nature of parking areas in Brunswick, on the other hand for reasons discussed above. The binary modelling of the zero observations with two predictors, the inclusion of which is based on model fit criteria, is relatively imprecise. A possible explanation is that street length and percentage land use not for parking contain little information on the probability of observing zero detected parking area per metre. Additional data on street signs and information of parking restrictions may provide useful information in this context, but are currently not available or in a digital processible format. To a lesser extent, this explanation also applies to the modelling of the continuous data in the Gamma model component, but this component uses additional information on road type, which helps explain variation in detected parking area per metre. These results suggest that it is particularly important to consider the underlying parameter uncertainty that is provided by the Bayesian model when making predictions. Information on the entire posterior distributions of the parameters is available, which means prediction uncertainty can be quantified and more importantly included in further modelling in a transparent manner, for example, when results are used in the context of an agent-based travel demand model, where these distributions help to calibrate the available parking areas. Finally, our model predicts detected parking area, but a conversion into the number of parking spaces is in principle possible by dividing total parking area by an average value of the area of a single parking space.

6 | CONCLUSION AND FUTURE WORK

This study is a proof of concept showing how open geodata in combination with remotely sensed data and Bayesian inference can be used to estimate a street level parking area prediction model for on-street parking. The model can be applied using information on road length, road type and land use as obtainable, for example, from OSM. The use cases are numerous: from creating parking constraints in travel demand models, to inventories for public administrations, to generating a data base for intelligent transport system (ITS) technologies that sup-

port parking. Even though the choice of available variables that can be used for prediction is highly limited, we show that it is in principle possible to predict parking area. Uncertainty in the prediction increases as the size of the detected parking areas per metre of street decreases. We therefore conclude that explicit consideration of the underlying parameter uncertainty when making predictions is imperative, something which is easily made possible by using a Bayesian inference in the estimation of the model as information on the entire posterior distribution of parameters is available. Going forward, we identify a need to investigate and quantify the contribution of the different sources of measurement errors as a result of data collection and variable creation in the predictive modelling process. It is also desirable to include more data describing the spatial and network structure in the estimation, for example from other cities, as it may help increase validity of the estimates in a broader context. Aside from model estimation, additional remote sensing data from these cities can be used to assess to which extent our approach is transferable to other regions. To ensure a good detection performance of parking areas in spite of specific natural and urban structures, it is recommended to perform a fine-tuning on a few samples of the new data. The field of application can be expanded by adapting the processing chain for high-resolution satellite images. As a result, we anticipate improved generalization capability even when the model is applied to structurally similar regions with no aerial imagery available.

AUTHOR CONTRIBUTIONS

J.H.: Conceptualization, methodology, project administration, software, writing – original draft preparation, writing – review and editing. A.K.: Formal analysis, methodology, software, writing – original draft preparation, writing – review and editing. M.L.D.: Formal analysis, visualization, writing – original draft preparation, writing – review and editing. N.M.: Methodology, software, writing – original draft preparation, writing – review and editing. C.H.: Methodology, software, writing – original draft preparation, writing – review and editing. F.K.: Conceptualization, software, supervision. M.H.: Conceptualization, supervision, writing – original draft preparation.

ACKNOWLEDGEMENT

The authors gratefully acknowledge funding by the German Aerospace Center (DLR) as part of the project *VMo4Orte: Integrated mobility for livable places*.

Open access funding enabled and organized by Projekt DEAL.








CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available on EOC Geoservice at <https://doi.org/10.15489/8zt0ay2yke58>, reference number [35].

ORCID

Jens Hellekes  <https://orcid.org/0000-0002-0080-3124>
 Ariane Kehlbacher  <https://orcid.org/0000-0003-3898-858X>
 María López Díaz  <https://orcid.org/0000-0002-7986-3970>
 Nina Merkle  <https://orcid.org/0000-0003-4177-1066>
 Corentin Henry  <https://orcid.org/0000-0002-4330-3058>
 Franz Kurz  <https://orcid.org/0000-0003-1718-0004>
 Matthias Heinrichs  <https://orcid.org/0000-0002-0175-2787>

REFERENCES

- Thigpen, C.G., Volker, J.M.: Repurposing the paving: The case of surplus residential parking in Davis, CA. *Cities* 70, 111–121 (2017). <https://doi.org/10.1016/j.cities.2017.06.020>
- Bonsall, P.W.: The changing face of parking-related data collection and analysis: The role of new technologies. *Transportation* 18(1), 83–106 (1991). <https://doi.org/10.1007/BF00150560>
- Coric, V., Gruteser, M.: Crowdsensing maps of on-street parking spaces. In: *Proceedings of the International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 115–122. Cambridge, MA, USA (2013). <https://doi.org/10.1109/DCOSS.2013.15>
- Yamada, S., Watanabe, Y., Kanamori, R., Sato, K., Takada, H.: Estimation method of parking space conditions using multiple 3D-LiDARs. *Int. J. ITS Res.* 20(2), 422–432 (2022). <https://doi.org/10.1007/s13177-022-00300-w>
- Di Mauro, D., Furnari, A., Patanè, G., Battiato, S., Farinella, G.M.: Estimating the occupancy status of parking areas by counting cars and non-empty stalls. *J. Visual Commun. Image Represent.* 62, 234–244 (2019). <https://doi.org/10.1016/j.jvcir.2019.05.015>
- Pozo, R.F., Gonzalez, A.B.R., Wilby, M.R., Vinagre Diaz, J.J., Matesanz, M.V.: Prediction of on-street parking level of service based on random undersampling decision trees. *IEEE Trans. Intell. Transp. Syst.* 23(7), 8327–8336 (2022). <https://doi.org/10.1109/TITS.2021.3077985>
- Waraich, R.A., Axhausen, K.W.: Agent-based parking choice model. *Transp. Res. Rec.* 2319(1), 39–46 (2012). <https://doi.org/10.3141/2319-05>
- Morency, C., Trépanier, M.: Characterizing Parking Spaces Using Travel Survey Data; 2008. <https://www.cirrelt.ca/documentstravail/cirrelt-2008-15.pdf> Accessed 10 October 2022
- Benenson, I., Martens, K., Birfir, S.: Parkagent: An agent-based model of parking in the city. *Comput. Environ. Urban Syst.* 32(6), 431–439 (2008). <https://doi.org/10.1016/j.compenvurbysys.2008.09.011>
- Kurz, F., Türmer, S., Meynberg, O., Rosenbaum, D., Runge, H., Reinartz, P., et al.: Low-cost optical camera systems for real-time mapping applications. *Photogramm. – Fernerkundung – Geoinf.* 2012(2), 159–176 (2012). <https://doi.org/10.1127/1432-8364/2012/0109>
- Cheng, G., Han, J., Lu, X.: Remote sensing image scene classification: Benchmark and state of the art. *Proc. IEEE* 105(10), 1865–1883 (2017). <https://doi.org/10.1109/JPROC.2017.2675998>
- Zhou, W., Newsam, S., Li, C., Shao, Z.: PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval. *ISPRS J. Photogramm. Remote Sens.* 145, 197–209 (2018). <https://doi.org/10.1016/j.isprsjprs.2018.01.004>
- Azimi, S., Henry, C., Sommer, L., Schumann, A., Vig, E.: SkyScapes – Fine-grained semantic understanding of aerial scenes. In: *Proceedings of the International Conference on Computer Vision (ICCV)*, p. 7392–7402. Seoul, South Korea (2019). <https://doi.org/10.1109/ICCV.2019.00749>
- Amato, G., Carrara, F., Falchi, F., Gennaro, C., Meghini, C., Vairo, C.: Deep learning for decentralized parking lot occupancy detection. *Expert Syst. Appl.* 72, 327–334 (2017). <https://doi.org/10.1016/j.eswa.2016.10.055>
- Drouyer, S.: Parking occupancy estimation on planetscope satellite images. In: *Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS)*, p. 1098–1101. Waikoloa, HI, USA (2020). <https://doi.org/10.1109/IGARSS39084.2020.9323104>
- Fraunhofer, I.A.O.: Using AI to measure the demand for parking space. Heilbronn, Germany (2021). <https://www.iiao.fraunhofer.de/en/press-and-media/latest-news/using-ai-to-measure-the-demand-for-parking-space.html> Accessed 10 October 2022
- Henry, C., Fraundorfer, F., Vig, E.: Aerial road segmentation in the presence of topological label noise. In: *Proceedings of the 25th International Conference on Pattern Recognition (ICPR)*. Piscataway, NJ, pp. 2336–2343. Milan, Italy (2021). <https://doi.org/10.1109/ICPR48806.2021.9412054>
- Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)* 2015. Cham, pp. 234–241. Munich, Germany (2015). https://doi.org/10.1007/978-3-319-24574-4_28
- Hazirbas, C., Ma, L., Domokos, C., Cremers, D.: FuseNet: Incorporating depth into semantic segmentation via fusion-based CNN architecture. In: *Proceedings of the Asian Conference on Computer Vision (ACCV)*. Cham, pp. 213–228. Taipei, Taiwan (2017). https://doi.org/10.1007/978-3-319-54181-5_14
- Henry, C., Hellekes, J., Merkle, N., Azimi, S., Kurz, F.: Citywide estimation of parking space using aerial imagery and osm data fusion with deep learning and fine-grained annotation. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* XLIII-B2-2021, 479–485 (2021). <https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-479-2021>
- Mosinska, A., Marquez-Neila, P., Kozinski, M., Fua, P.: Beyond the pixel-wise loss for topology-aware delineation. In: *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway, NJ, pp. 3136–3145. Salt Lake City, UT, USA (2018). <https://doi.org/10.1109/CVPR.2018.00331>
- Abinaya, N.S., Susan, D., Rakesh Kumar, S.: Naive Bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries. *Ecol. Inf.* 61, 101248 (2021). <https://doi.org/10.1016/j.ecoinf.2021.101248>
- Davison, A.C., Hinkley, D.V.: *Bootstrap Methods and their Application*. In: *Cambridge Series on Statistical and Probabilistic Mathematics*, vol. 1. Cambridge University Press, Cambridge (1997). <https://doi.org/10.1017/CBO9780511802843>
- Federal Agency for Cartography and Geodesy: Digital Landscape Model (DLM250) (2021). <https://gdz.bkg.bund.de/index.php/default/open-data/digitales-landschaftsmodell-1-250-000-ebenen-dlm250-ebenen.html> Accessed 10 October 2022
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B.: *Bayesian Data Analysis*. 3rd ed. CRC Press, Boca Raton (2013)
- Vehtari, A., Gelman, A., Gabry, J.: Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* 27(5), 1413–1432 (2017). <https://doi.org/10.1007/s11222-016-9696-4>
- Stan Development Team. Stan User's Guide: Version 2.27 (2021). https://mc-stan.org/docs/2_27/stan-users-guide-2_27.pdf Accessed 10 October 2022
- Stan Development Team. RStan: the R interface to Stan: R package version 2.21.1 (2021). <https://mc-stan.org/> Accessed 10 January 2022
- Bürkner P.-C.: brms: An R package for Bayesian multilevel models using Stan. *J. Stat. Soft.* 80(1), 1–28 (2017)
- Hoffman, M.D., Gelman, A.: The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* 15, 1593–1623 (2014). <https://www.jmlr.org/papers/volume15/hoffman14a/hoffman14a.pdf>
- Hellekes, J., Merkle, N., Lopez Diaz, M., Henry, C., Heinrichs, M., Azimi, S., et al.: Assimilation of parking space information derived from remote sensing data into a transport demand model. In: *ITS World Congress 2021: Book of Abstracts*, pp. 2579–2690. Hamburg, Germany (2021)
- Barrington-Leigh, C., Millard-Ball, A.: The world's user-generated road map is more than 80 % complete. *PLoS One* 12(8), e0180698 (2017). <https://doi.org/10.1371/journal.pone.0180698>
- Eckle, M., Albuquerque, J.P.: Quality assessment of remote mapping in openstreetmap for disaster management purposes. In: *Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM)*. Krystiansand, Norway (2015).

- http://idl.iscram.org/files/melanieeckle/2015/1218_MelanieEckle+JoaoPortodeAlbuquerque2015.pdf
34. Gelman, A.: Multilevel (hierarchical) modeling: What it can and cannot do. *Technometrics* 48(3), 432–435 (2006). <https://doi.org/10.1198/004017005000000661>
 35. German Aerospace Center (DLR), EOC Geoservice. UrMo Digital – Traffic Area Map (TAM) – Brunswick, Germany (2022). <https://geoservice.dlr.de/web/maps/eoc:urmotam> Accessed 10 October 2022

How to cite this article: Hellekes, J., Kehlbacher, A., Díaz, M.L., Merkle, N., Henry, C., Kurz, F., Heinrichs, M.: Parking space inventory from above: Detection on aerial images and estimation for unobserved regions. *IET Intell. Transp. Syst.* 1–13 (2022). <https://doi.org/10.1049/itr2.12322>