EXPLOITING NEW SENSOR TECHNOLOGIES
FOR REAL-TIME PARKING PREDICTION IN URBAN AREAS

Eleni I. Vlahogianni, Ph.D.
Department of Transportation Planning & Engineering
School of Civil Engineering
National Technical University of Athens
5 Iroon Polytechniou Str, Zografou Campus, 15773 Athens Greece
Tel: +30 210 772 1369; Fax: +30 210 772 1454
e-mail: elenivl@central.ntua.gr

Kostanstinos Kepaptsoglou, Ph.D.
School of Rural and Surveying Engineering
National Technical University of Athens,
5 Iroon Polytechniou Str, Zografou Campus, 157 73 Athens Greece
Tel: +30 210 772 3167; Fax: +30 210 772 2404
e-mail: kkepap@central.ntua.gr

Vassileios Tsetsos, Ph.D.
Mobics Ltd
27 Kifissias Avenue, 11523 Athens Greece
Tel: +30 210 643 3525; Fax: +30 210 643 3407
e-mail: btsetsos@mobics.gr

Matthew G. Karlaftis, Ph.D.
Department of Transportation Planning & Engineering
School of Civil Engineering
National Technical University of Athens
5 Iroon Polytechniou Str, Zografou Campus, 15773 Athens Greece
Tel: +30 210 772 1280; Fax: +30 210 772 2404
e-mail: mgk@central.ntua.gr

Word Count: 6241
Tables: 2
Figures: 9

*corresponding author
ABSTRACT
This paper proposes a methodological framework - based on survival analysis and neural networks - to provide parking availability forecasts for extended prediction horizons. Two different types of predictions are provided: i. the probability of a free space to continue being free in subsequent time intervals, and ii. the short-term parking occupancy prediction in selected regions of an urban road network. The available data comes from a wide network of parking sensors installed on-street in the “smart” city of Santander, Spain. The sensor network is segmented in four different regions and, then, survival and neural network models are developed for each region separately. Findings show that the Weibull parametric models best describe the probability of a space continuing to be free in the forthcoming time intervals. Simple genetically optimized Multilayer Perceptrons accurately predict region parking occupancy up to 1 hour in the future by only exploiting 5 minute data. Finally, the real time, web based, implementation of the proposed parking prediction availability system is presented.

Keywords: Parking occupancy, parking duration, parking sensors, smart city, internet of things
INTRODUCTION AND BACKGROUND

Parking availability and provision are among the most important factors affecting both private car based trip decisions and traffic conditions in urban areas. Driver decisions are temporally dependent, implying that they are influenced by past experience, as well as real-time (on road) perceptions. Parking is such a case where prior knowledge on possible prevailing conditions (e.g. difficulty in finding a parking space, off-street parking costs, and so on) affects parking decisions. At the same time, vehicles in search of vacant parking spaces negatively impact traffic conditions and the environment. In such a context, parking information provision is a research area of particular interest since modern communication technologies offer alternative ways of delivering information to travelers in a timely and effective manner.

Systems developed for providing parking information and guidance have been proposed by researchers in the past (1-4); in such systems, information is usually disseminated by variable message signs or through the internet, cellular phones, PDA, and GIS technologies (2, 4). Teodorovic and Lucic (5) argue that although parking guidance systems may not affect the occupancy rate or average parking duration, drivers tend to greatly appreciate the information provided by such systems. This is the case since such systems significantly increase the probability of finding vacant parking spaces, mitigate frustration of those drivers/visitors unfamiliar with the city center, decrease queues in front of parking garages, decrease total vehicle-miles travelled (particularly in the city center), decrease average trip time, energy consumption, and air pollution.

The usefulness of such predictive parking information is straightforwardly understood. If all drivers act without information and make “uninformed” choices, they will probably resort to similar optimal decisions leading to induced long waiting times, queues and increased parking circling. On the other hand, dissemination of accurate and timely parking availability information may lead to improved driver decisions and parking searching (6). From a transport operator perspective, accurate parking availability forecasts may lead to better management of the system, congestion mitigation due to queue formation avoidance, and so on.

Parking modeling has been a topic of interest since the 1970s. An early review by Young et al. (7) reported three categories of parking related models: driver behavior with respect to parking (parking choice), optimal positioning of parking lots (parking allocation), and interaction of parking operations with other transportation system elements and infrastructures (parking interaction). A decade later, Arnott and Rowse (8) indicated that parking has been investigated from several perspectives including parking patterns, impacts on traffic, off-street parking technologies, parking policy, choice and location models, as well as economic models describing parking conditions. Parking availability prediction on the other hand, was introduced in recent years, coupled with modern capabilities on data collection and processing, along with ITS based exploitation of such information (a fact already indicated since the early 1990s (9)). In the early 2000s, David et al. (10) proposed a model for event-oriented forecasting of parking occupancy based on standardized daily distribution occupancy rate curves and on-line data obtained from parking lots equipped with detectors. On-street occupancy estimation in the absence of detectors was investigated by David and Keller (11). An event driven model was developed for that purpose using historical socioeconomic and parking specific data; the model was successfully validated for the city of Munich, Germany. Teodorovic and Lucic (5) developed a system based on fuzzy logic, simulation and optimization models, which decides whether to accept or reject new parking requests in real-time, according to estimated availability for
parking lots. Martens and Benenson (12) exploited agent-based modeling for representing parking behavior by integrating the effect of considering real-time and expected parking availability, prices, and parking enforcement. Caicedo (13) developed a discrete choice model for combining online and historical data for real-time, off-street parking availability prediction; this model was later used by Caicedo et al. (3). A fuzzy logic model for estimating the uncertainty of peak period parking availability in park-and-ride facilities was proposed by Chen et al. (14). Recently, Fubuyasi et al. (15) developed ParkPGH, a system for predicting parking availability in eight Pittsburgh parking facilities using historical and real-time data.

Modern technological advances have revolutionized the ways of monitoring and recording transportation operations and data; these have, however, been exploited for parking prediction to a lesser extent, and have mainly focused on off-street parking facilities. The objective of this paper is to exploit statistical and computational intelligence methods for developing a methodology that can be used for multiple step ahead on-street parking prediction in “smart” urban areas. This work takes advantage of massive real-time parking data availability, obtained by an extended parking sensor network available in the “smart” city of Santander, Spain. Models are developed for predicting expected parking occupancy along with the probability of finding vacant parking spaces. Traditional survival analysis models as well as neural network models are developed. The methodology is evaluated and a real-time, web-based system exploiting the proposed prediction models, for the city of Santander, Spain, is presented.

SMART CITIES AND PARKING PREDICTION CHALLENGES

Cities are characterized as “smart” when their transportation and communication (ICT) infrastructures along with their human and social capital investments cooperate and actively support sustainable growth and high quality of life, through participatory action and engagement while preserving natural resources (16). Indeed, as noted by Komninos (17), innovation and use of ICT for improving capacity of infrastructures are key elements of “smart” cities. A novel type of “smart” city infrastructure, applicable to the transportation sector is the so-called Internet of Things (IoT). IoT consists of a variety of devices or objects – such as Radio-Frequency IDentification (RFID) tags, sensors, actuators, mobile phones, and so on – which, through unique addressing schemes, are able to interact with each other and cooperate with their neighbors to reach common goals (18). By continuously collecting, analyzing and redistributing transportation information, IoT networks can offer valuable, real time information to both travelers and operators, and thus support and improve the operations of ITS, traffic and public transportation systems. Although one can trace several reasons that may prevent IoT to be fully developed in urban environments, this unique technological paradigm is expected to substantially support sustainable development of future smart cities (19).

The SmartSantander project is such a case of an IoT architecture deployed in the city of Santander, to achieve a massive deployment of sensors and network communications in order to provide efficient and equitable transportation and other services to citizens (http://www.smartsantander.eu/). The IoT network of the SmartSantander project consists of (20):

- IoT nodes: these are responsible for sensing and collecting information from the natural and socioeconomic environment and activities (temperature, CO, noise, light, car presence etc.).
• **Repeaters**: These nodes are placed high above ground in street lights, semaphores, information panels, etc, and behave as forwarding nodes of IoT information.

• **Gateways**: Both IoT nodes and repeaters are configured to send all the information to gateways. Once information is received by gateways it can be either stored locally or sent to central processing units through different interfaces (for example, WiFi, GPRS/UMTS or Ethernet).

SmartSantander and similar testbeds fuse high resolution datasets stemming from both static and mobile sensors. Such data refer to macroscopic traffic flow in road sections of interest, detailed parking information in real-time in urban areas, as well as high resolution transit information. In the case examined in this paper, having parking sensors collecting real time information on their occupancy provides accurate high resolution parking information. New technologies can replace common parking metrics of average duration, turnover rate and occupancy in extended time windows of 1 to 3 hours (which are manually collected in small scale regions) with accurate, both aggregate and disaggregate information, on occupancy and parking duration. In this framework there are several questions that may arise; this paper focuses on the following:

• Is it possible to predict parking occupancy using time series modeling approaches based on data collected from an IoT network of sensors?

• How accurate are parking occupancy predictions produced by such models in relation to the predictive horizon?

• What are the statistical properties of parking space duration and how can we predict the probability of having free parking spaces in the area of interest?

**MODELING PARKING AVAILABILITY BASED ON SENSOR DATA**

**Methodology**

The architecture of the proposed parking occupancy prediction system is presented in FIGURE 1. In each sliding time window $T$, parking efficiency is defined by the following metrics:

- **Duration ($\bar{D}$)** of free parking space: The time period that a slot is free,
- **Occupancy ($\bar{O}$)**: The percentage of parking slots occupied during a predefined time period.

The free space duration is an indication of how frequently a parking space becomes available, whereas occupancy relates the parking accumulation, meaning the number of parked vehicles in the study area at any specific time period, to the parking capacity.

The proposed methodology has two modules. The first module is a real-time time series occupancy prediction scheme based on recurrent artificial neural networks. Simple yet flexible memory mechanisms will be applied in order to replicate the temporal dynamics of parking occupancy $\bar{O}$. The model is presented with past information of occupancy ($\bar{O}_{t-1}, \ldots, \bar{O}_{t-(m-1)}$) to predict occupancy in one step ahead $\bar{O}_t$. The characteristics of the model’s memory (the time delay $\tau$ and the dimension $m$) will be evaluated through a nonlinear analysis of the dynamics of parking occupancy. The use of exogenous variables such as traffic volume will be evaluated relative to the
improvement on the accuracy of predictions they impose. Moreover, the use of more than one neural network models with relation to the type of day (weekends, weekdays) will be evaluated.

![Functional architecture of the proposed parking occupancy prediction scheme.](image)

**FIGURE 1.** Functional architecture of the proposed parking occupancy prediction scheme.

When considering short-term prediction systems that operate in real-time and in an “intelligent” technology-based environment, the effectiveness depends, mostly, on predicting traffic information in a timely manner (21), (22). Real-time system effectiveness depends both on the results and on the time in which these are produced (23). The computational time for making a prediction mainly depends on the functional form of the prediction system; empirical results show that data-driven prediction systems that include recursive data-search algorithms exhibit ‘best’ prediction accuracy but need extensive computational time for convergence at acceptable results (21). For this, the final structure of the prediction model is kept as simple and flexible as possible. The trade-off between simplicity and efficiency is studied in a preliminary stage of analysis using genetic algorithms to optimize the structural and learning parameters of the different models.

Nevertheless, to ensure the optimal performance of the prediction system the proposed approach may encompass a real-time training strategy to account for new (and unobserved) patterns. The uncertainty in training time is treated by controlling the extent of training data in each category of neural network models. During real-time operation of the prediction system data where inaccurate predictions are detected (mean relative percept error exceeding a pre-specified threshold for example) are saved for retraining. The system will automatically prompt for retraining when the error systematically increases.

The second is a static approach for estimating the probability of finding available parking space with relation to a series of factors such the traffic volume entering the project’s area, the type of day (weekday, weekend), and the time period (peak, off-peak, morning evening). The modeling will be based on survival analysis. Parametric hazard based modeling may be developed under a variety of functional forms. The model will be also presented with the predicted value of occupancy as provided by the first module. The output of the model (the probability of finding an available parking space and the anticipated number of available space) can be visually
depicted using various graph methods (such as heat maps), to provide useful information to users.

**Neural Network Models for Time Series Prediction**

Time series modeling is a popular approach for making predictions in transportation problems (24). This approach is suitable for analyzing parking occupancy due to the temporal structure of the performance measures of parking systems. A common prediction strategy implemented in transportation problems is based on the autoregressive moving average family of models. These models are relatively straightforward mathematically and easy to produce; however, they are severely constrained by stationarity and linearity, characteristics that most frequently violated in real transportation time series. Treating non-stationarity and nonlinearity may lead to a tedious process without achieving the desired levels of accuracy in predictions and modeling reliability (25). Neural Networks (NNs) for time series provides a good alternative, as they relax many of these constraints (25), and also appear to provide short-term forecasting models that are more adaptable to sudden shifts in the data (27). A recent study on traffic time series prediction has shown the structural equivalencies between nonlinear univariate and multivariate ARIMA models with exogenous variables and dynamic forms of Multilayer Perceptrons (MLPs) (27). The simplest of all is the NAR(p) structure of order p. In general, The MLP presented with p lagged values of parking occupancy \( O_k \) \( (k=1,\ldots,n) \) may act as a predictor of the form (28):

\[
\hat{O}_t = h(O_{t-1,\ldots,O_{t-p}) = \sum_{i=1}^{p} W_i f \left( \sum_{j=1}^{p} w_j O_{t-j} + \theta_i \right) \tag{1}
\]

Where \( f(\cdot) \) is a smooth monotonic function, \( W_i \) and \( w_j \) are the weights of the connections (synapses) - coefficients - estimated through learning (training) thereby obtaining an estimate of the nonlinear approximation \( \hat{h} \). The network converges to an estimate of \( \hat{h} \) by minimizing the residuals \( \sum_k \left( O_k - \hat{O}_k \right)^2 \) using a learning algorithm (usually back-propagation).

For a time series consideration in a NN framework, the MLP should be modified to account for the time sequence of events under study. This is usually accomplished by adding memory structures in than MLP that retain the effect of past information to the system and use it during learning. The memory is accomplished using local - at an intra-neuron level - and global - between neurons of different layers - recurrent connections in a neural network. Memory mechanisms may be of a simple tap delay form realizing structures:

\[
O(t) = \{O(t-\tau),\ldots,O(t-(m-1)\tau)\} \tag{2}
\]

where \( \tau \) is the delay and \( m \) the dimension of the horizon of past information introduced to the model or of other more complex mathematical forms (Gamma memory and so on). The networks with memory usually require cumbersome and slow learning procedures that may not be always stable (27). To avoid this, static MLPs can be externally modified to represent the temporal characteristics of transportation time series (e.g. parking occupancy) in a manner resembling the common statistical prediction techniques. The introduction of such data inputs in
MLPs that are unchanged in their internal structural logic may conceptually approximate very complex and multivariate statistical structures with equal efficiency as classical MLPS.

**Hazard-Based Free Parking Duration Modeling**

Hazard-based duration modeling deals with the statistical representation of time to event data, a very frequent form of data in transportation problems; typical example of such transportation data is the time to clear and incident (29), (30), the time until the end of congested phenomena (31), (32), the time until the end of transit vehicle repair (33), the time to an activity (34), (35), the time to household evacuation under the emergence of a physical disaster (36), the time for a pedestrian to cross a signalized intersection approach (37), the time to vehicle transaction (38) and so on. Extensive review on the hazard based transportation modeling applications may be found in Hensher and Mannering (39) and Bhat (40). Moreover, methodological, computational and estimation issues in duration modeling with focus on transportation problems may be found in Washington et al. (25).

Parametric hazard-based modeling is based on two concepts: the survival function and the hazard function. Let \( T \) be a non-negative random variable representing the time a vehicle occupy a space, the survival function is defined as the probability that \( T \) is of length at least \( t \) (i.e. a parking space is occupied at least \( t \) min) and is given by:

\[
S(t) = P(T > t) = 1 - F(t), \quad 0 < t < \infty, \tag{3}
\]

where \( F(t) \) is the cumulative probability. The survival function \( S(t) \) can have a variety of shapes following certain restrictions; it is bounded by 0 and 1 and, as \( T \) cannot be negative, \( S(0) = 1 \). Moreover, the larger the \( t \), \( S \) never increases (and usually decreases).

For continuous survival data, the hazard function specifies the instantaneous failure rate at \( T=t \) conditional upon survival to time \( t \) and defined as (25):

\[
h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}, \tag{4}
\]

where \( f(t) \) is the probability density function:

\[
f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt} \tag{5}
\]

By combining (2) and (3) we get:

\[
h(t) = -\frac{d}{dt}\log S(t) \tag{6}
\]

The hazard function \( h(t) \) is always nonnegative and, unlike survival functions, has no upper bound. Parametric hazard based models may have a range of different functional forms aiming at modeling different distributional characteristics. The hazard rate \( h(t|X) \) with covariates is given by:

\[
h(t|X) = h_b(t) \exp(\beta X) \tag{7}
\]

where \( h_b(t) \) is the baseline hazard rate, \( \beta \) is a vector of estimated parameters and \( X \) is a vector of covariates.
IMPLEMENTATION AND RESULTS

Dataset
Parked on information comes from a network of 253 sensors located in the area of Santander (FIGURE 2). The available dataset consists of time series of the state of each sensor every 1 minute for the month of April 2013. With this information, it is possible to straightforwardly calculate parking metrics such as parking accumulation, occupancy, duration and so on. In this study we focus on: i) parking occupancy (%) of a specific area, which is the percentage of parking spaces that are occupied by vehicles within a time interval; ii) the turnover rate (veh/space/t), which is equal to the number of vehicles per parking space in a time interval; and, iii) the duration that a parking space is not occupied by vehicles. For correctly analyzing the parking characteristics the entire area covered by the parking sensors is further segmented into four (4) regions; this segmentation is done based on drivers’ feasible cruising paths while searching for available parking spaces. The specific regions are depicted in FIGURE 2.

FIGURE 2: Parking sensors location in the area of Santander (map accessed at: www.ploigos.gr/es)

FIGURE 3 and 4 show the time series of parking occupancy (%) and the turnover rate (veh/space/h) per region for a typical week. Further statistical testing shows that there exist differences in the mean of occupancy and turnover rate between weekdays and weekends for all regions. Moreover, high variability in average parking occupancy for Region I in the morning and afternoon is observed; setting 85% as the critical occupancy, we may distinguish between high occupancy and low occupancy periods within a day. High/low occupancy periods for the other parking regions have an occupancy threshold of 90%.
FIGURE 3: Hourly evolution of parking occupancy (%) and turnover rate (veh/space/h) for Regions I and II for a typical week.

FIGURE 4: Hourly evolution of parking occupancy (%) and turnover rate (veh/space/h) for Regions III and IV for a typical week.
Duration of occupied spaces for all regions is best described by a Weibull distribution; the survival curve has the following form: 
\[ S(t) = \exp(-\lambda t^p) \]
with parameters \( \lambda \) and \( p \) to be estimated. FIGURE 5 shows the survival probabilities for the duration of occupied parking spaces per region. Higher probabilities after 60 minutes on average are expected in Region I and II when compared to those of Regions III and IV, where the probability of a space being occupied after 60 minutes drops below 0.5.

![Survival curves](image)

**FIGURE 5**: Survival probabilities for the duration of occupied parking spaces per region.

**Survival Analysis of Free Parking Spaces**

Hazard based free parking space duration models are developed for each parking region. Based on preliminary analysis, a Weibull survival function is used. Several forms of Weibull regression models are tested with independent variables being the TypeDay (weekend/weekday), the Period of day (high/low occupancy period) and the Weekday (1 to 7, from Monday to Sunday). Results are seen in TABLE 1.

**TABLE 1**: Survival analysis results.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Region I</th>
<th>Region II</th>
<th>Region III</th>
<th>Region IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-2.98</td>
<td>-2.90</td>
<td>-2.65</td>
<td>-2.06</td>
</tr>
<tr>
<td>TypeDay</td>
<td>0.43</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Period</td>
<td>-</td>
<td>0.13</td>
<td>0.15</td>
<td>-</td>
</tr>
<tr>
<td>Weekday</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.04</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>770.86</td>
<td>-1764.64</td>
<td>-1932.55</td>
<td>-4367.82</td>
</tr>
<tr>
<td>LR chi2(1)</td>
<td>10.15</td>
<td>9.78</td>
<td>8.28</td>
<td>14.36</td>
</tr>
<tr>
<td>Prob &gt; chi2</td>
<td>0.000</td>
<td>0.002</td>
<td>0.002</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

As can be observed, different independent variables are significant in each parking region when modeling free parking space duration. In Region I, free space
durations are dependent on the type of day (weekend or weekdays), whereas in Region II and III, free space durations depend on the period of the day. Finally, in Region IV, free space durations change with regard to the daily evolution of parking demand. Interesting results arise from observing the survival probability curves (FIGURE 6) that show whether the phenomenon will continue to be observed in the forthcoming intervals, given it has lasted up to time $t$. For example, in Region I there is a probability of a parking space continuing to be free after 5 minutes of 0.76 for weekdays; this increases to 0.83 for weekends.

![Survival curves based on the models developed for free space durations in the four regions under study.](image)

The same probability for Region II spaces is 0.77 for high occupancy periods and increases to 0.79 in low occupancy periods. In Region III the probability of a space to be free after 5 minutes is 0.69 and 0.73 for high and low occupancy periods respectively. Finally, in Region IV the differences of the same probability are small but significant the probability ranges from 0.63 in Mondays and 0.69 in Sundays. The above survival curves may be used as a predictive tool for parking availability.

**Parking Occupancy Prediction**

A unique MLP for each parking region is developed in order to predict the overall occupancy (%) of parking spaces using past information on the evolution of occupancy. The input space is introduced with information on occupancy with a look back time window ranging from 5 to 10 minutes. For the prediction system to be useful, multiple steps ahead predictions are required. Here, the prediction horizon of the developed models is extended from 1 step to 60 steps ahead (that is, from 1 to 60 minutes ahead). Following previous research on short term prediction using neural networks in transportation problems, the neural networks are optimized with respect to their structural and learning properties using genetic algorithms. Training is conducted using a simple back propagation algorithm with genetically optimized learning rate and momentum. The structure of the hidden layer (number of hidden
units) is also optimized using genetic algorithms. The model’s look back time window is also genetically optimized. The available data is divided into the training (60%), cross-validation (20%) and test set (20%) used for training the models, evaluating the training of the models, and testing the generalization power of the models, respectively. The genetic optimization showed that an MLP of 8 hidden layers and a look back time window of 5 minutes in the past may be efficiently used to predict parking occupancy (%) up to 60 steps in the future with high accuracy. Prediction results (test set) are seen in TABLE 2 with respect to the following forecasting metrics:

- **Mean Absolute Error (MAE):** \[ \frac{1}{N} \sum_{i=1}^{N} |\hat{y}_{i\tau} - y_{i\tau}| \]
- **Root mean squared error (RMSE):** \[ \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\hat{y}_{i\tau} - y_{i\tau})^2} \]
- **Mean absolute percentage error (MAPE):** \[ \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\hat{y}_{i\tau} - y_{i\tau}}{y_{i\tau}} \right| \]
- **Root relative squared error (RRSE):** \[ \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_{i\tau} - y_{i\tau})^2}{N}} / \sqrt{\frac{\sum_{i=1}^{N} (\hat{y}_{i\tau} - y_{i\tau})^2}{N}} \]

where \( y_{i\tau} \) and \( \hat{y}_{i\tau} \) is the actual and predicted value with \( i = 1, \ldots, N \), \( \tau \) is the time step, \( N \) is the number of samples, \( y_{i\tau} \) is the last known value relative to the prediction step. Results seem encouraging particularly for longer prediction horizons. For example, the models are able to predict each region’s parking occupancy 15 minutes into the future with less than 4.2% MAPE.

**TABLE 2: Results for different prediction horizons (prediction step \( \tau \) equals to 1 minute) for each parking region (test set).**

<table>
<thead>
<tr>
<th></th>
<th>Prediction Horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 min</td>
</tr>
<tr>
<td><strong>Region 1</strong></td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>0.011</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.018</td>
</tr>
<tr>
<td>MAPE</td>
<td>1.374</td>
</tr>
<tr>
<td>RRSE</td>
<td>92.057</td>
</tr>
<tr>
<td><strong>Region 2</strong></td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>0.007</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.010</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.816</td>
</tr>
<tr>
<td>RRSE</td>
<td>76.843</td>
</tr>
<tr>
<td><strong>Region 3</strong></td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>0.007</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.010</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.702</td>
</tr>
<tr>
<td>RRSE</td>
<td>83.768</td>
</tr>
<tr>
<td><strong>Region 4</strong></td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>0.011</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.016</td>
</tr>
<tr>
<td>MAPE</td>
<td>1.237</td>
</tr>
<tr>
<td>RRSE</td>
<td>87.012</td>
</tr>
</tbody>
</table>
Interestingly, all models perform well and produce predictions that are more accurate than a naïve prediction (e.g. taking the last known value relative to the step of prediction) as seen from the RRSE. The larger the predictive horizon the more accurate the MLP becomes relative to a naïve prediction. FIGURE 7 shows a graphical representation of the MAPE evolution with respect to different prediction horizons. As can be observed, on average, the errors produced by the MLPs developed are very satisfactory, particularly when compared to other prediction problems met in transportation (24). Interestingly, parking occupancy predictability is different between the four Regions. More accurate predictions are accomplished in Regions 2 and 3, whereas the least accurate predictions are accomplished in Region 1.

![Graph showing MAPE for different regions](image)

**FIGURE 7:** Mean Absolute Percent Error (MAPE) for multiple step ahead neural network parking occupancy prediction models (test set).

It is to note that MLPs manage to produce predictions up to 60 minutes ahead with acceptable levels of accuracy. Although this may hint a rather smooth temporal evolution of 1 minute parking occupancy data, the efficiency of these models may prove to be critical in delivering useful information services to travelers in urban areas.

**REAL WORLD APPLICATION**

The above models have been implemented in the MITOS system (Multi-Input TranspOrt planning System). MITOS is a system developed within SmartSantander project ([http://mitos.smartsantander.eu](http://mitos.smartsantander.eu)), in the city of Santander, Spain. MITOS goal is to deliver novel Intelligent Transportation Services (ITS) to smart city citizens and capture perceived user experience. The fact that an IoT (Internet of Things) infrastructure is available, provides new ways of quantifying the effect that ITS may have on the daily transportation of the commuters and the environment. The MITOS platform provides an integrated multi-modal transportation guide, which allows citizens and visitors to optimally choose their trips and get accurate information and guidance before and during their trip. The MITOS platform includes the following services:

- Stop/Line/POI Survey and Search for public transport,
- Route Guide,
- Estimation of Next Bus arrival,
• Parking monitoring and short-term parking occupancy prediction.

The architecture of the MITOS platform is seen in FIGURE 8. The services and applications deployed and demonstrated in the context of MITOS include a Web portal for transportation information, which is used as a city guide for citizens and tourists, and a mobile application for advanced transportation services. The last provides all type of information required by commuters and enables intuitive ways for searching routes and geo-referenced information based on the personal preferences and abilities of each user.

![FIGURE 8: Architecture of MITOS platform.](image)

An added value of the MITOS platform comes from the exploitation of various diverse sensor sources such as:
• Participatory sensors: users will act as sensors that provide relevant traffic and travel information in the form of free text or predefined messages (e.g. “heavy traffic”, “too much noise”, etc.) and/or image.
• Environmental data sensors, measuring noise, temperature and CO/CO2 emissions
• On-street parking space occupancy sensors.
• Traffic occupancy sensors
• On-vehicle devices (GPS) installed in buses.

The implementation of a parking prediction service involves retrieving parking sensor data every minute (through the SmartSantander framework APIs) and storing them in a relational database (MITOS DB). Next, the algorithms described earlier are executed, and a continuous stream of predictions is generated. Those predicted values are stored back in the database, so that related user requests on parking availability can be served as fast as possible. Following user requests for parking occupancy, data are retrieved and visualized on a digital Web map (FIGURE 8) and on a mobile phone (FIGURE 9). The visualization involves data easily perceived by the end users, such as estimated occupancy and estimated number of free parking slots in the short-term future (5, 10 and 15 minutes ahead).
FIGURE 9: Screenshot of Web parking prediction implementation.

FIGURE 10: Screenshot of mobile parking prediction implementation.
CONCLUSIONS
Driver decisions increasingly rely on technologies and information services that deliver high resolution transportation information. Parking availability prediction is a critical problem that has been recently investigated by researchers with the use of advanced algorithms combined with novel technological capabilities in data collection and processing. In this paper we exploited high resolution parking occupancy data and developed and tested a system for short and longer term parking availability prediction in urban areas.

The system encompasses two modules; the first applies survival analysis to predict the probability that a parking space will be free in the following time intervals. The second module introduces neural networks for the prediction of the time series of parking occupancy in different regions of an urban network. Findings show that the duration of free parking space follows a Weibull distribution. Moreover, the neural networks adequately captures the temporal evolution of parking occupancy and may accurately predict occupancy up to one hour ahead. The system is successfully incorporated within an innovative routing service (MITOS) for Santander, Spain, using real time data from the city’s “smart” infrastructure.

ACKNOWLEDGEMENTS
This work is part of the MITOS experiment, incorporated in the SmartSantander project and funded by the European Union Seventh Framework Programme (FP7).

REFERENCES
16. Caragliu, A; Del Bo, C. & Nijkamp, P. Smart cities in Europe. Serie Research Memoranda 0048 (VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics), 2009.