Gone in 30 Days! Predictions for Car Import Planning

Emanuel Lacic, Matthias Traub, Tomislav Duricic, Eva Haslauer and Elisabeth Lex

Abstract: A challenge for importers in the automobile industry is adjusting to rapidly changing market demands. In this work, we describe a practical study of car import planning based on the monthly car registrations in Austria. We model the task as a data driven forecasting problem and we implement four different prediction approaches. One utilizes a seasonal ARIMA model, while the other is based on LSTM-RNN and both compared to a linear and seasonal baselines. In our experiments, we evaluate the 33 different brands by predicting the number of registrations for the next month and for the year to come.

ACM CCS: Information systems → Information systems applications → Decision support systems → Expert systems

Keywords: Automotive industry; Data-driven expert systems; Car brand recommendation; Linear methods; Nonlinear methods; Deep learning; Customer demand

1 Introduction

The automotive industry, one of the largest private investors in research and development[1], has been facing an unstable and changing market demand [6, 9]. Especially as the competition between car manufacturers increases, it has become increasingly challenging to anticipate customers' demand. On the one hand, customers ask for more individualized vehicles and a larger product variety. On the other hand, policy measures for reducing emission levels or increasing tax revenue also impact the vehicle choice [18, 28]. To cover high production expenses while still satisfying local demands, many automobile manufacturers assemble vehicles at scale economies and offload them to automobile importers and car dealers [20, 36]. The problem with this approach is that this limits importers' and car dealerships' ability to respond to dynamic market demands. For instance, they would implement official distribution strategies and maintain an inventory of certain models or brands at predetermined volumes each year. However, as reported in [36], such manufacturing and distribution strategies are contributing to regional differences in pricing, availability of specific models and vehicle specifications in general. In order to suppress and regulate the impact of these strategies, it is crucial for importers and car dealerships to know and plan for the real market demand. Studies tried to model the effect of external factors like personal income and the general market situation [1] and were able to predict the general market behaviour but could not apply this on a regional and brand based level.

The present work. In a joint collaboration with Porsche Austria, we present a real-world application of predicting vehicle registrations for car importers and dealerships in the Austrian market. Specifically, the task is to predict the number of car registrations of a specific car brand for the next month as well as for the following year. This information in turn allows to make timely decisions on how many cars of a specific brand should be imported and distributed. For this purpose, we utilize monthly registration data of 33 individual car brands made available by Porsche Austria. These car registrations are further divided into three categories based on the number of monthly registration. This lets us estimate the popularity of a particular brand. Respectively, we present and compare the efficacy against a simple linear baseline and a seasonal baseline within the MASE of two forecasting approaches for car import planning: (i) Long Short-Term Memory Recurrent Neural Network (LSTM-RNN), and (ii) Seasonal Autoregressive Integrated Moving Average (SARIMA). In our experiments, we study two application scenarios. Firstly, in a short-term setting, the aim is to predict the number of cars of a specific brand, which will be registered in the next month. This information is especially valuable for car dealerships as they can still adjust the official

[1] According to the European Commission: `https://ec.europa.eu/growth/sectors/automotive_en` (accessed on January 3rd, 2017)

distribution strategies of the automobile manufacturers when the market suddenly becomes volatile. Secondly, predicting the car registrations for the following year not only helps for a timely import planning, but also enables to provide feedback to the automobile manufacturers to adjust their distribution strategy beforehand. Having a detailed prediction of the upcoming month further helps to better plan and design marketing strategies and also new car model introductions. As shown by Pauwels et. al. [24], marketing strategies that increase short-term sales and the introduction of a new product do have an impact on the long-term sales performance.

2 Related Work

Since the 1930s until today it has been essential to find models and predict changes in car demand [27, 34]. This has been of interest for entities like governments, automobile manufacturers, car dealerships, environmental protection groups or public transport authorities [32, 36]. Early studies tended to use regression models with economic indices, such as the GDP or world fuel prices, in combination with manufacturer-specific variables like car prices and maintenance costs [5, 26]. Other studies tried to investigate and model the effect of external factors [1, 24]. As the availability of information sources can be an issue, recent studies also looked into utilizing data from the Web [11]. All these approaches aimed to predict the general market behaviour. However, for our specific use case we are interested in the local market behaviour of Austria and the distinct car brands.

At present, we identify three main lines of research that are related to our work: (i) forecasting problems using linear models, (ii) forecasting problems using nonlinear models, and (iii) comparative studies of both.

Linear models. The publication from Box and Jenkins in 1970 [3] has probably had one of the biggest impacts on the theory and practice of modern time-series analysis. They improved on early formulations of the Autoregressive Moving Average (ARMA) models [33] by introducing *differencing* to handle non-stationary time-series data. In stationary time-series, statistical properties such as mean, variance, and autocorrelation are all constant over time. Nonstationarity can be recognized by observing a trend or seasonality in the data. A time-series needs to be stationary in order to properly fit an ARMA model. The introduction of *differencing* had the purpose to bridge this gap by stationarizing the time-series. The ability to work with both stationary and non-stationary data has popularized the use of ARIMA models and their extensions in many areas of science and industry like [8, 30], just to name a few.

Nonlinear models. In contrast to linear models, which assume that the future value of a time-series is linearly related to past observations, nonlinear models like Artificial Neural Networks (ANN) are increasingly gaining momentum. Especially in recent years, recurrent neural networks (RNNs), a type of deep neural networks, which are *deep* in the temporal dimension, have been extensively used for forecasting problems [13, 25]. A RNN is an extension of a conventional feed-forward neural network, which can handle sequences of variable length. This is achieved through a recurrent hidden state whose activation at each time step depends on the activation of the previous one. However, it has been observed that in general, it is difficult to train an RNN to successfully learn the long-term dependency within a sequence (e.g., due to the vanishing or exploding gradient problem) [23]. One effective solution for this problem is to use a Gated Recurrent Unit (GRU) [7] or a Long Short-Term Memory (LSTM) [14] recurrent neural network. In this work, we focus on the LSTM-RNN, which is based on memory cells instead of neurons. By applying a nonlinear function and adding a forget gate, each cell maintains a memory of a specific point in time. Thus, each cell can decide whether to keep or discard its existing memory. Previous work has shown (e.g., [10, 29]) that such an approach is especially beneficial when modelling forecasting problems as it allows to learn long-term dependencies.

Comparative studies. Many works focus on comparing RNNs with a standard stationary ARIMA approach. For instance, the authors of [31] showed that ANNs are better suited than an ARMA or ARIMA model to make long-term water inflow predictions for dam reservoirs when using monthly data gathered for a period of 42 years. In [35], the authors showed that an LSTM neural network outperforms an ARIMA approach for predicting wind power. They used a year's worth of data where the wind speed was recorded every 15 minutes. In case of predicting traffic speed, the authors of [17] showed on a 1-month traffic speed data with the updating frequency of 2 minutes that an LSTM neural network offers the best performance when compared to an ARIMA model. Similar results were presented in [19], where the authors aimed to predict the daily value of bitcoins using the data from past 3 years. A general conclusion could be drawn that approaches based on ANNs typically outperform ARIMA based approaches when enough observations are provided. However, we found that this does not apply in our setting with a smaller amount of available data.

3 Methodology

We formulate the task of predicting the number of car brand registrations as follows: let $[x_1, x_2, ..., x_n]$ be the monthly registrations extracted for a specific brand. We state that $x_i \in \mathbb{N}, (1 \leq i \leq n)$ is the number of new registrations a car brand x had at month i. To plan the number of vehicles to import, we seek a model

M that given a historical monthly registrations time-series $x = [x_1, x_2, ..., x_{r-1}, x_r], (1 \leq r \leq n)$ provides the forecast $y_{r+1} = M(x)$, where $y \in \mathbb{R}$ is the number of car brand registrations for the next month. In order to get the output for several months ahead (e.g., also to predict y_{r+2}), we would incrementally extend the historical monthly registration time-series using the previously predicted values (e.g., $y_{r+2} = M(x)$ where $x = [x_1, x_2, ..., x_{r-1}, x_r, y_{r+1}]$). This way, we can say how many cars of a specific brand should be planned for import based on the past registration data.

3.1 Baseline

For comparison, we first introduce a naive baseline approach, which uses a simple linear regression model to predict the upcoming registrations. In this approach, only the last 12 months of the training data are used to train the model. In the case of the short-term prediction, the baseline is retrained at every step, where at each step, one month is added to the training set. This is further explained in the following Section 4.1 where we also introduce a seasonal baseline that is contained within the utilized MASE evaluation metric.

3.2 Long Short-Term Memory (LSTM)

Recurrent neural networks (RNN) are an extension of conventional feed-forward networks, with the difference that they can handle time-series of variable-length. This is achieved through a recurrent hidden state, whose activation at each time-step is dependent on the previous state. As already described in Section 2, RNNs can be seen as deep architectures in the sense that they are unrolled in time and that each layer shares the same model parameters. In this work, we consequently utilize a Long Short-Term Memory (LSTM) neural network [14] to successfully learn possible long-term dependencies. Specifically, we utilize an LSTM-RNN with one hidden layer combined with a standard stochastic gradient descent (SGD) for optimization and we apply Nesterov's momentum [22]. The internal weights of our model are initialized with Xavier initialization [12]. To implement the LSTM-RNN we utilize the open-source Deeplearning4j[2] Java library.

Model configuration. When constructing a network, there are still many hyperparameters, which impact the final performance of the trained network. For instance, we utilize the l_2 norm and dropout [38] for regularization as it makes the trained model less sensitive to noisy data. Additionally, we add a bias to the forget gate of the memory cell as it was also found to improve the performance of the network [37]. As proposed by Mikolov (2012) [21], we also clip the gradients on a per-element basis. For each gradient g, we set it to $sign(g) * max(maxAllowedValue, |g|)$. This means that if a parameter gradient has an absolute value greater than the defined threshold, we truncate it. This results in eight different hyperparameters for which we define the following search space: (1) the number of hidden units in a layer $n_h \in [10, 300]$, (2) learning rate $\eta \in [0.0001, 0.1]$, (3) momentum $\mu \in [0.1, 0.99]$, (4) dropout $p \in [0.1, 0.9]$, (5) $L2$ regularization $L_2 \in [0.0001, 0.1]$, (6) a gradient normalization threshold $g \in [0.5, 1000]$, (7) the length of forward and backward truncated backpropagation through time $k \in [1, 46]$ as well as (8) a forget gate bias $b_f \in [0.5, 5.0]$. We search for the optimal values setting using random search, as it was shown to be a more efficient strategy for parameter optimization than grid search [2]. This was carried out as an initial step using the extracted train and validation set (see later in Section 4). Every search iteration lasted 200 epochs and the whole process has been ran for 14 hours on an IBM System x3550 server with two 2.0 GHz six-core Intel(R) Xeon(R) E5-2620 processors and 128 GB of RAM. This resulted in 169 different model configurations, which were evaluated on the validation set using the RMSE metric (see later in Section 4.1). The resulting optimal hyperparameter values were: $n_h = 232$, $\eta = 0.0121$, $\mu = 0.6568$, $p = 0.6597$, $L_2 = 0.0368$, $g = 588.225$, $k = 33$ and $b_f = 0.9773$.

3.3 SARIMA

SARIMA is an Autoregressive Integrated Moving Average (ARIMA) model that considers seasonality in the time-series data. ARIMA is a linear model and based on the estimated time-series shown in Figure 1, we assume that linear time-dependence is an important aspect of our data. To identify an ARIMA model means to determine the orders (number of time lags) of the Autoregressive (AR) and Moving Average (MA) components, as well as the degree of differencing. As already described in Section 2, differencing refers to a transformation applied to time-series data in order to make it stationary (if necessary). This model can then be denoted as ARIMA(p, d, q), where p is the order of the non-seasonal AR component, d is the number of non-seasonal differencing and q is the order of the non-seasonal MA component. This notation does not consider the case when seasonality is involved. In this setting, the model is extended as ARIMA$(p, d, q) \times (P, D, Q)_s$ and thus named seasonal-ARIMA or SARIMA for short. The seasonal model is then determined by the before mentioned p, d and q parameters as well as four additional ones: (i) the seasonal AR order P, (ii) the seasonal differencing D, (iii) the seasonal MA order Q and (iv) a seasonal period s. The formal notation of SARIMA is defined in [4]. For simulation and prediction with Seasonal ARIMA models we utilize 'sarima', an open-source R package[3].

Model configuration. To identify the best SARIMA model for our data, first, the parameters p, d, q, P, D,

[2] `http://deeplearning4j.org/lstm`

[3] https://CRAN.R-project.org/package=sarima

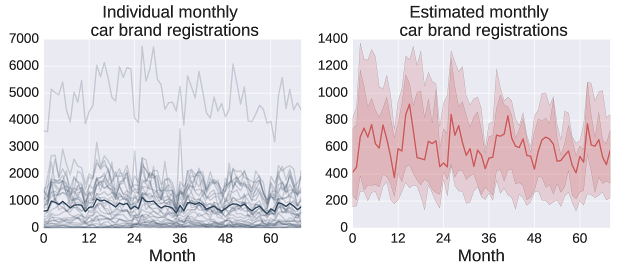


Figure 1: Monthly car brand registrations used for training. Left plot shows individual monthly car brand registrations while the right plot shows a linear trend estimation of the time-series data based on the median and with 75% and 95% confidence bands.

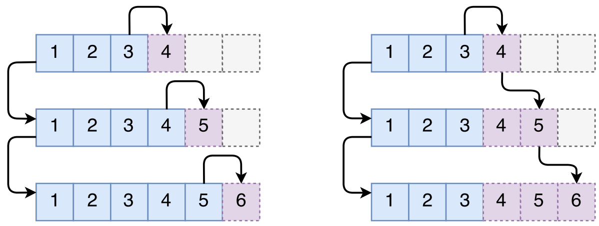


Figure 2: Evaluation procedure for the next month prediction (i.e., short-term prediction on the left) compared to predicting the next year to come (i.e., long-term prediction on the right). The difference lies in how the historical monthly registration time-series is constructed before it is used to predict the number of vehicles to import for the next month.

Q and s need to be determined. For this task, we apply the step-wise algorithm for traversing the model space as proposed by Hyndman and Khandakar [15]. The algorithm conducts a search through the model space within the provided constraint values and returns the best ARIMA model with respect to the Akaike information criterion function with a correction for finite sample sizes (AICc). This results in an ARIMA model with a parameter combination that outputs the lowest AICc values. In our experiments, we define a search space of the parameter combinations as follows: $p \in [0, 5]$, $d \in [0, 2]$, $q \in [0, 5]$, $P \in [0, 2]$, $D \in [0, 1]$, $Q \in [0, 1]$ and we set s to 12 (i.e., a whole year). Note that each car brand yields a model with different parameter configuration.

4 Experimental Setup

Our study is carried out on car brand registration data, which was provided by Porsche Austria [4]. In total, we look at the monthly registrations of 33 individual car brands beginning with 2010 up to the end of 2016. During this period, four new car brands have just been introduced (i.e., Dacia, Mini, Smart and Tesla) and therefore, for them, a smaller set of data points is available. As seen in Figure 1, there are differences in popularity between the individual brands. As such, we further divide the car brands into three categories based on the median number of monthly registrations (i.e., $[0, 100)$, $[100, 1000)$ and $[1000, \infty)$). Categorizing the car brands in such a way enables us to more easily compare the prediction performance between the individual brands. For instance, having a prediction error of 401 vehicles may seem high. However, if that is the case for Volkswagen (VW) (i.e., with monthly car registrations in the thousands), we eventually get on average a percentage error of 8.59% and then such a number starts to look more tolerable for the car dealerships. The specific car brands and their corresponding category can be seen in Table 1 of Section 5 where we report our evaluation results. The table is sorted by the median number of monthly registrations within the training set where Tesla has the lowest median number of registrations (i.e., 1) and VW by far the most (i.e., $4,705$). Due to space restrictions, we actually leave out the results of car brands in Table 1 from Category 1 (Subaru), Category 2 (Suzuki, Nissan, Citroen, Kia and Volvo) and Category 3 (Hyundai, Ford, und Opel) that show a similar performance to other brands from the same category.

4.1 Evaluation protocol

Every car brand contains one data point (i.e., the number of car registrations) for every month of the data collection period. We divide these monthly values of each brand into a train, validation and test set. The training set contains the first 45 months for each car brand (i.e., a brand has a sequence of only 45 registration measurements). The following 24 months are used for the validation set and the most recent 12 data points (i.e., one year worth of brand registrations) represent the test set. To evaluate the three prediction approaches, we use the train and validation set to find the best model configuration for LSTM-RNN and SARIMA as described in Section 3.2. The individual training data, as well as the linear trend estimation, can be seen in Figure 1. It shows the effect of seasonality every 12 months and the different characteristics of each brand as well as the different amplitudes.

In our experiments, we focus on two scenarios: (1) predicting how many cars of a specific brand will be registered next month (i.e. a short-term prediction), and (2) predicting how many cars will be registered in the following year (i.e. a long-term prediction). Specifically, as shown in Figure 2 the difference in the evaluation procedure lies in how the historical monthly registration time-series is constructed. By predicting next month's number of car brand registrations, car dealerships can timely react and adjust the official distribution strategies when the market suddenly becomes volatile. To evaluate a short-term prediction, we always use the complete history to anticipate how many cars of a specific brand should be imported next month. We start with the com-

[4] This data can also be extracted from the Austrian institute for statistics: `http://www.statistik.at/web_en/statistics/EnergyEnvironmentInnovationMobility/transport/road/registration_of_new_vehicles/index.html`

plete train and validation set to make the first prediction and compare it to the first data point of the test set. Afterwards, we add the first data point of the test set to the history and predict the second data point of the test set. We then continue in such a manner until the whole test set is evaluated. In case of the long-term prediction, we use the complete train and validation set to predict the whole test set. In such a setting, each predicted data point is used as an input to predict the following one. As already stated in Section 3, we incrementally extend the monthly registration time-series from the training set with previously predicted values in order to get the next one.

Evaluation metrics. For our study, we report our prediction performance using the following four evaluation metrics: (1) Mean Absolute Error (MAE), (2) Root-Mean-Square Error (RMSE), (3) Mean Absolute Percentage Error (MAPE) and, (4) Mean Absolute Scaled Error (MASE).

By using MAE, we measure how far away on average was a predicted value from the expected one:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|p_t - e_t|$$

The RMSE is a well-known quadratic scoring rule which measures the average magnitude of the error between predicted and expected values:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{N}(p_t - e_t)^2}{N}}$$

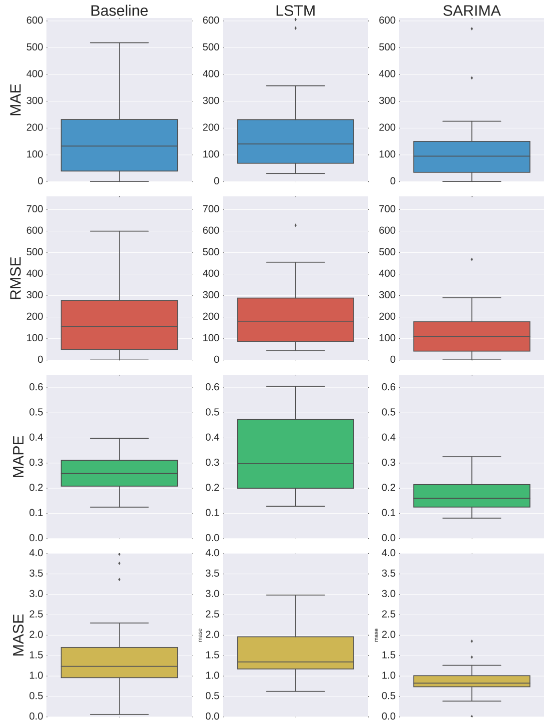
The RMSE gives a relatively high weight to large prediction errors which means that it is most useful when large errors are particularly undesirable, as is the case when planning car imports. The MAPE is expressed in generic percentage terms to show how big the error is when compared to the expected values:

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}|\frac{p_t - e_t}{e_t}|$$

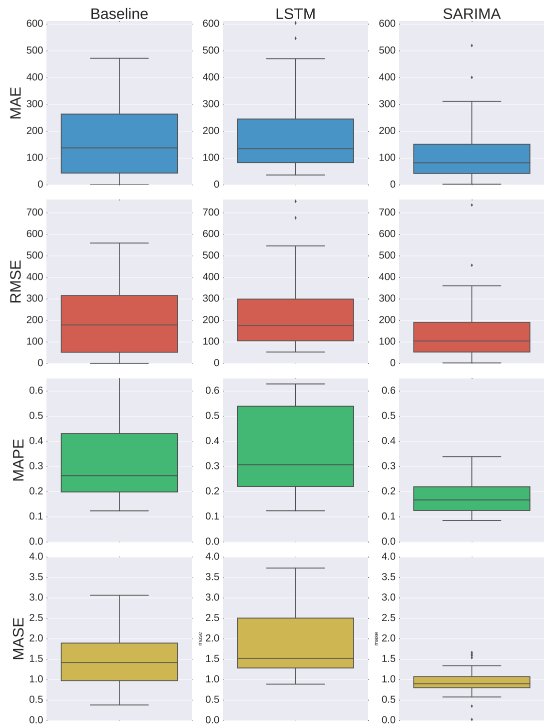
This is useful when we want to compare the prediction accuracy on cars, which occur in completely different categories based on the number of monthly registrations. The MASE is a measure of forecast accuracy which compares the MAE of the actual forecast with the MAE_{naive} produced by a naive forecast [16]:

$$MASE = \frac{1}{T}\sum_{t=1}^{N}\left(\frac{|e_t|}{\frac{1}{T-m}\sum_{t=m+1}^{T}|p_t - p_{t-m}|}\right)$$

When forecasts are based on seasonal data, the naive forecast sets each prediction to the last observed value of the same season (1 year before). The error measure is independent of the scale of the data an can therefore be used to compare forecasts across various car brands with different scales. In contrast to the MAPE, the mean absolute scaled error takes negative and positive errors equally into account as well as large and small forecast errors.



(a) Short-term prediction.

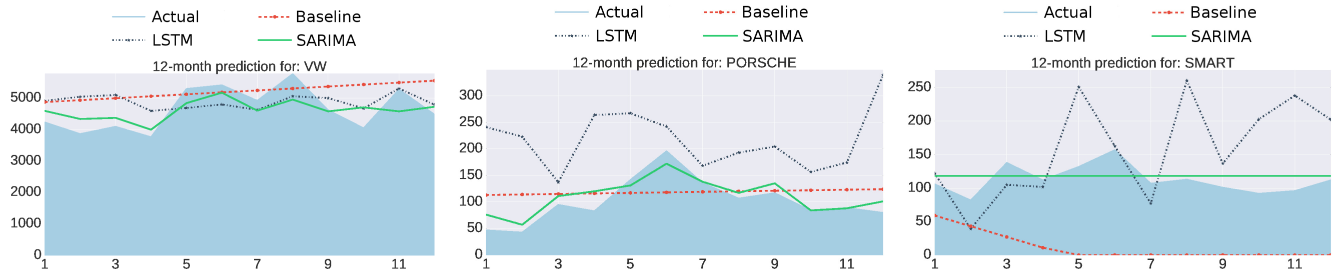


(b) Long-term prediction.

Figure 3: Evaluation results of the linear baseline, LSTM-RNN and SARIMA in terms of MAE, RMSE, MAPE, MASE. We report the results for the task of predicting the number of car brands for the next month (a) and the whole following year (b).

5 Results

The overall performance of the simple baseline, the LSTM-RNN and the SARIMA approach is given in Figure 3. In general, the SARIMA approach outperforms



(a) (b) (c)

Figure 4: The examples show the long-term evaluation results from Table 1 for Volkswagen, Porsche and Smart comparing forecasted with actual values. Brands from the 3rd category like Volkswagen (a) should favor lower MAPE, whereas for the ones from the 1st category, like Porsche (b), a small MAE is more important. As seen by the example of Smart (c), there are cases when SARIMA cannot find a seasonal trend and the predictions converge to a mean value.

both the baseline and the LSTM-RNN approach. With respect to our two evaluation scenarios, in our case, short-term predictions (Figure 3.a) are slightly more accurate than long-term predictions (Figure 3.b). Interestingly enough, the differences between the short and long-term predictions are quite small, which suggests that the utilized approaches can be reliably used in both settings given our datasets.

In the reported MAPE boxplots, we cut off four distinctive outliers (i.e., Lancia, Tesla, Jaguar and Chevrolet) since the value differences would distort the plots. The brands Lancia and Chevrolet exhibit the worst performance in the evaluated test set. The reason for this is that they basically vanished from the Austrian car market in the evaluated period. The MASE plot also shows that SARIMA performs much better than the linear regression and the LSTMN-RNN. However, in some cases the linear baseline is performing better than the SARIMA and the implicit seasonal naive baseline. The linear baseline was able to model the diminishment of Lancia and Chevrolet and performed best. SARIMA was competitive, yet, the LSTM-RNN showed a much larger prediction error in both the next month and the following year prediction. This is due to the fact that the data we used for training had no examples of car brands, which were gradually vanishing from the market. As such, the network could not successfully recognize the gradual decrease. Actually, we generally argue that since we have a small number of monthly data for only 33 car brands, i.e., 33 different time-series examples, makes it really hard to successfully train an LSTM-RNN for this kind of a problem. Due to such low amounts of extracted car brand registration data, the SARIMA approach positions itself as the better choice to be used in our system for car import planning.

Looking at evaluation results of individual car brands in Table 1[5], we can see that by using SARIMA we can achieve the best performance for all three categories in terms of MAE, RMSE, MAPE and MASE. BMW was actually the only car brand on which the LSTM-RNN performed better than SARIMA. The results of SARIMA are statistically significantly better in all four error metrics with p values < 0.05. Of particular interest is the third category of car brands, where the average monthly vehicle count is the highest. In such cases, when the number of new brand registrations is in the thousands, a low MAPE is especially favorable. For example as seen in Table 1 and Figure 4.a, Volkswagen (VW) has on average around 4,000 car registrations each month and a low MAPE value such as 8.59% is most favorable, as the opposite would mean much higher extra costs when compared to the other two categories. On the contrary, when looking at Porsche within the 1st category of Table 1, we see that the MAPE when using the SARIMA approach is 19.10% for the long-term prediction. With Porsche being a premium brand, an overall smaller MAE and RMSE value is a much better indicator for the performance of the utilized approach since the number of registrations is comparatively low. This can be seen even better in Figure 4.b, where the results for the long-term predictions are really close to the ground truth.

Discussion. While SARIMA performs best most of the time, there are cases where a seasonal trend cannot be found and the predictions eventually converge to a mean value and remain the same. This was the case for three different car brands in our experiments, namely, Smart, Jaguar and Land Rover. As seen in Figure 4.c for Smart, the predicted mean value still produces the smallest error but is arguably useful as a prediction output. In general SARIMA scores better values in terms of MASE compared to LSTM-RNN where the naive seasonal baseline is nearly always better. Nonetheless, in some cases the naive seasonal baseline is better than SARIMA. In this respect we also explored the performance of a weighted hybrid (e.g., $M_{Hyb} = \alpha \cdot M_{RNN} + \beta \cdot M_{SARIMA}$) to tackle the shortcomings of the utilized methods. We found that there are certain settings, where such a combination can be useful (e.g., Jaguar, Dacia, Peugeot and BMW), but it would still be needed to investigate further.

[5] Please note that we only report the long-term prediction, i.e., prediction results of the next 12 months for individual car brands. This is because the performance between the evaluated approaches on the reported car brands is mostly the same.

Brand	LSTM				SARIMA			
	MAE	*RMSE*	*MAPE*	*MASE*	*MAE*	*RMSE*	*MAPE*	*MASE*
Category 1								
Tesla	77.19	93.44	2.2538	2.693	26.25	31.59	.7577	.916
Jaguar	67.08	75.69	1.3248	1.445	48.83	56.30	.5879	1.052
Smart	65.22	82.61	.6139	1.294	17.83	20.86	.1655	.354
Lancia	123.82	138.09	134.2800	1.891	2.75	2.78	2.0000	.029
Jeep	135.24	152.14	.5052	3.731	32.5	42.24	.1407	.897
Porsche	115.76	134.68	1.5735	7.938	15.25	18.77	.1910	1.046
Category 2								
Land Rover	55.25	65.53	.3074	1.287	35.58	45.32	.1855	.829
Mini	91.89	99.54	.3980	2.583	57.83	70.35	.2198	1.625
Alfa Romeo	69.44	90.53	.6285	1.673	43.00	55.45	.3342	1.036
Honda	62.37	83.44	.3194	1.198	58.17	65.65	.2951	1.117
Mitsubishi	113.83	131.80	.5398	1.521	43.25	62.00	.1882	.578
Chevrolet	37.13	53.32	25.3580	9.617	6.58	7.37	5.1000	1.591
Dacia	162.48	222.73	.2328	1.102	124.75	162.20	.1881	.846
Toyota	187.81	213.26	.4011	1.526	90.00	102.82	.2009	.731
Mazda	121.54	145.11	.1412	1.257	82.75	104.71	.0966	.856
Category 3								
Peugeot	131.13	154.23	.1478	1.026	103.17	143.45	.1175	.807
Mercedes	367.28	419.89	.2908	1.996	138.67	161.15	.1127	.754
Fiat	259.87	290.59	.2283	1.494	102.00	127.76	.0960	.586
Seat	246.04	311.76	.1978	.930	238.50	266.32	.2009	.901
BMW	293.52	340.53	.1732	.891	311.83	361.36	.1825	.946
Renault	323.74	374.42	.2207	1.375	204.00	236.58	.1255	.866
Audi	198.98	260.31	.1243	1.055	151.75	191.20	.0989	.804
Skoda	425.15	479.02	.2347	1.914	170.75	208.86	.1009	.769
VW	604.02	676.81	.1375	1.619	401.25	456.72	.0859	1.075

Table 1: Results of the LSTM and SARIMA approaches when predicting the number of vehicles of each car brand for the next 12 months. Car brands of the 3rd category have the best MAPE results, which is important if we consider their average number of monthly brand registrations. In most of the cases SARIMA has a lower MASE and therefore is preferable to LSTM-RNN. These results area also statistically significantly better in all four error metrics with p values < 0.05.

6 Conclusion

In this paper, we discuss how the problem of car import planning is tackled at Porsche Austria. We showed how to model the prediction task and evaluated four different approaches: (i) a simple linear baseline, (ii) an implicit naive seasonal baseline within MASE, (iii) a SARIMA model and, (iv) an LSTM recurrent neural network. Our experimental results reveal that in most settings, the SARIMA approach showed the best performance. We assume that this is the case due to a low number of data points and time-series examples, which are available to train the LSTM-RNN. By comparing the short-term and long-term predictions, we conclude that all three approaches can be used for a long-term car import planning as the error differences are insignificant.

One limitation of our work is the small number of data points that are available for the car import data of the Austrian market. By incorporating the information of neighbouring countries we plan to investigate if a better performance of the RNN approach can be achieved. We didn't investigate the influence of additional data features for our prediction task. As such, for future work we also plan to extract additional features like the registration numbers of used cars and investigate different methods (e.g., seasonal ARIMA with exogenous parameters), which allow for additional input features. Finally, we found that there are certain settings, where a weighted combination of the utilized approaches can lead to even better performance results. We are further investigating the problem of learning the weighting parameters in order to achieve a meaningful prediction improvement.

In the end we have shown that a linear model can be used to predict the upcoming car registrations of a car brand for one month up until a whole year with adequate accuracy for planning marketing strategies and new product introductions. These insights can be applied to other domains which have similar market behaviour. Further, we showed that the usage of a non-linear model like an LSTM-RNN did not perform best due to a low number of data points and this trade-off between higher complexity and higher computational costs vs. performance and accuracy increase needs to be considered.

Acknowledgement

The authors would like to thank Dieter Theiler and Dominik Kowald for their help and valuable comments on this work. This work is supported by the Know-Center, which is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies.

Literature

[1] Adda, J., Cooper, R.: The dynamics of car sales: A discrete choice approach. Tech. rep., National Bureau of Economic Research (2000)

[2] Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization. Journal of Machine Learning Research 13(Feb), 281–305 (2012)

[3] Box, G.E., Jenkins, G.M.: Time series analysis: Forecasting and control (1970)

[4] Brockwell, P.J., Davis, R.A.: Introduction to time series and forecasting. Springer Science & Business Media (2006)

[5] Carlson, R.L., Umble, M.M.: Statistical demand functions for automobiles and their use for forecasting in an energy crisis. Journal of Business (1980)

[6] Childerhouse, P., Disney, S.M., Towill, D.R.: On the impact of order volatility in the european automotive sector. Int. Journal of Production Economics (2008)

[7] Cho, K., van Merriënboer, B., Gülçehre, Ç., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y.: Learning phrase representations using rnn encoder–decoder for statistical machine translation. In: Proc. of EMNLP'14. ACL (Oct 2014)

[8] Ediger, V.Å., Akar, S.: Arima forecasting of primary energy demand by fuel in turkey. Energy Policy 35(3), 1701 – 1708 (2007)

[9] Elkins, D.A., Huang, N., Alden, J.M.: Agile manufacturing systems in the automotive industry. Int. Journal of Production Economics (2004)

[10] Eyben, F., Weninger, F., Squartini, S., Schuller, B.: Real-life voice activity detection with lstm recurrent neural networks and an application to hollywood movies. In: Proc. of ICASSP'13. IEEE (2013)

[11] Fantazzini, D., Toktamysova, Z.: Forecasting german car sales using google data and multivariate models. Int. Journal of Production Economics (2015)

[12] Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: Aistats. vol. 9, pp. 249–256 (2010)

[13] Graves, A.: Sequence transduction with recurrent neural networks. arXiv preprint arXiv:1211.3711 (2012)

[14] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural computation 9(8), 1735–1780 (1997)

[15] Hyndman, R., Khandakar, Y.: Automatic time series forecasting: The forecast package for r. Journal of Statistical Software 27(1), 1–22 (2008)

[16] Hyndman, R.J., Koehler, A.B.: Another look at measures of forecast accuracy. International Journal of Forecasting 22(4), 679 – 688 (2006)

[17] Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y.: Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transportation Research Part C: Emerging Technologies 54, 187–197 (2015)

[18] Mandell, S.: Policies towards a more efficient car fleet. Energy Policy (2009)

[19] McNally, S.: Predicting the price of bitcoin using machine learning (2016)

[20] Miemczyk, J., Holweg, M.: Building cars to customer order – what does it mean for inbound logistics operations? Journal of Business Logistics (2004)

[21] Mikolov, T.: Statistical language models based on neural networks. Presentation at Google, Mountain View, 2nd April (2012)

[22] Nesterov, Y.: A method of solving a convex programming problem with convergence rate o (1/k2). In: Soviet Mathematics Doklady. vol. 27, pp. 372–376 (1983)

[23] Pascanu, R., Mikolov, T., Bengio, Y.: On the difficulty of training recurrent neural networks. ICML (3) 28, 1310–1318 (2013)

[24] Pauwels, K., Silva-Risso, J., Srinivasan, S., Hanssens, D.M.: New products, sales promotions, and firm value: The case of the automobile industry. Journal of marketing 68(4), 142–156 (2004)

[25] Robinson, A.J.: An application of recurrent nets to phone probability estimation. IEEE transactions on Neural Networks 5(2), 298–305 (1994)

[26] Romilly, P., Song, H., Liu, X.: Modelling and forecasting car ownership in britain: a cointegration and general to specific approach. Journal of Transport Economics and Policy pp. 165–185 (1998)

[27] Rudd, E.: The relationship between the national income and vehicle registrations. road research laboratory. RN 1518 (1951)

[28] Ryan, L., Ferreira, S., Convery, F.: The impact of fiscal and other measures on new passenger car sales and co 2 emissions intensity: evidence from europe. Energy Economics 31(3), 365–374 (2009)

[29] Sak, H., Senior, A., Beaufays, F.: Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition. arXiv preprint arXiv:1402.1128 (2014)

[30] Tse, R.Y.: An application of the arima model to real-estate prices in hong kong. Journal of Property Finance 8(2), 152–163 (1997)

[31] Valipour, M., Banihabib, M.E., Behbahani, S.M.R.: Comparison of the arma, arima, and the autoregressive artificial neural network models in forecasting the monthly inflow of dez dam reservoir. Journal of hydrology 476, 433–441 (2013)

[32] Whelan, G.: Modelling car ownership in great britain. Transportation Research Part A: Policy and Practice 41(3), 205–219 (2007)

[33] Whitle, P.: Hypothesis Testing in Time Series Analysis. Almqvist & Wiksells (1951)

[34] de Wolff, P.: The demand for passenger cars in the united states. Econometrica: Journal of the Econometric Society pp. 113–129 (1938)

[35] Wu, W., Chen, K., Qiao, Y., Lu, Z.: Probabilistic short-term wind power forecasting based on deep neural networks. In: Proc. of PMAPS'16. IEEE (2016)

[36] Yeung, G., Mok, V.: Manufacturing and distribution strategies, distribution channels, and transaction costs: The case of parallel imported automobiles. Managerial and Decision Economics 34(1), 44–58 (2013)

[37] Zaremba, W.: An empirical exploration of recurrent network architectures (2015)

[38] Zaremba, W., Sutskever, I., Vinyals, O.: Recurrent neural network regularization. arXiv preprint arXiv:1409.2329 (2014)

M.Sc. Emanuel Lacic
Emanuel Lacic is a Software Engineer and Scientific Researcher in the Social Computing division at the Know-Center. He holds a M.Sc. and B.Sc. in Software Engineering and Information Systems from the University of Zagreb and, is currently a PhD student at the Graz University of Technology. His main research focus lies on Recommender Systems, Deep Learning and Social Network Analysis.

Address: Know Center GmbH, Inffeldgasse 13, AT-8010 Graz, E-Mail: elacic@know-center.at

Dipl. Ing. Matthias Traub
Matthias Traub studied Information and Computer Engineering at the Technical University of Graz. He was a research assistant of the department of Intelligent Information Systems at Joanneum Research (2013-2014). Since 2014 he is a member of the Social Computing department at the Know Center in Graz, working in the field of data driven systems with the focus on recommender systems.

Address: Know Center GmbH, Inffeldgasse 13, AT-8010 Graz, E-Mail: mtraub@know-center.at

M.Sc. Tomislav Duricic
Tomislav Duricic is a Software Engineer and a Junior Scientific Researcher in the Social Computing team at the Know-Center. He holds a Master's degree in Software Engineering and Information Systems from the University of Zagreb and is currently a PhD student at the Graz University of Technology. His current research interests are in the fields of Complex Network Analysis and Recommender Systems.

Address: Know Center GmbH, Inffeldgasse 13, AT-8010 Graz, E-Mail: tduricic@know-center.at

Dr. Eva Haslauer
Since October 2016 Eva Haslauer ist working as a Projectmanager at Porsche Austria GmbH & CO OG in Salzburg. She works in the Digital Department and is responsible for digital projects in the realm of connected car, used car sale and sales forecast. She holds a PhD in Applied Geoinformatics from the University of Salzburg and worked as PostDoc in Innovation Management at the GI Department.

Address: Porsche Austria GmbH & Co OG, Louise-Piech Strasse 2, AT-5020 Salzburg, E-Mail: eva.haslauer@porsche.co.at

Ass.-Prof. Dr. Elisabeth Lex
Elisabeth Lex is assistant professor at Graz University of Technology and head of the Social Computing area at Know-Center. Her research focus is on Social Computing, Recommender Systems and Data Science and she has published 60+ papers in these fields. She earned her PhD in Text Mining in the Social Web at Graz University of Technology, Austria, advised by Michael Granitzer and Stefanie Lindstaedt.

Address: Technical University of Graz, AT-8010 Graz, E-Mail: elex@tugraz.at