Estimating city wide hourly bicycle flow using a hybrid LSTM MDN

Marcus Skyum Myhrmann*[1] and Stefan Eriksen Mabit[1]

[1]Department of Technology, Management and Economics, DTU, Denmark

SHORT SUMMARY

This study proposes a novel method to improve the estimation of hourly bicycle flow. The specific model employed is a Long Short-Term Memory Mixture Density Network (LSTMMND). This model presents the additional upside approximates a distribution of cycling flow conditional on the input data and hence addresses potentially unobserved heterogeneity. In a case of city-wide bicycle flow in Copenhagen, the LSTMMDN yielded $\sim 75\%$ more accurate bicycle flow estimates than the calibration-based approach currenty used by transport agencies. The paper further quantifies the improvement of accident analyses brought on by the improved bicycle volume measures. The LSTMMDN estimates result in an improved model fit in a crash model, compared to other estimates for bicycle exposure, with all other variables unchanged. Overall, the results strongly indicate that investing in more advanced methods for bicycle volume estimation will benefit the quality of performance measures related to bicycle issues computed by transport agencies.

Keywords: Aggregation Bias, Bicycle crash risk, Bicycle volume estimation, Deep learning, Long Short-Term Memory Mixture Density Network

1. INTRODUCTION

In the past decades safety engineers, transport agencies, and researchers have investigated various aspects of bicycle accidents to identify the factors associated with bicycle crash risk (Aldred et al., 2018; Dozza, 2017; Janstrup, Møller, & Pilegaard, 2019; Rossetti et al., 2018; Twisk & Reurings, 2013; Vandenbulcke, Thomas, & Int Panis, 2014), and injury outcome(Fountas et al., 2021; Kaplan, Vavatsoulas, & Prato, 2014; Myhrmann, Janstrup, Møller, & Mabit, 2021) to make informed mitigating efforts and improve the safety of cycling. However, many such investigations do not account for the cyclist exposure in relation to the factors investigated (Dozza, 2017; Vandenbulcke et al., 2014).

An issue investigating risk factors in relation to bicycle traffic is the general lack of bicycle flow data. To address this issue, transportation agencies use calibration factors applied to the annual average daily cycling traffic (AADCT) or annual average weekday cycling traffic (AAWCT) in order to obtain reasonable hourly volume profiles, (Schrank, 2021). However, the estimates derived from calibration factors do not reflect variations in traffic external factors such as weather and bank holidays. This presents a significant issue considering the impact weather and other factors have on cyclist ridership (Böcker, Dijst, & Prillwitz, 2013; Nankervis, 1999; Nosal & Miranda-Moreno, 2014).

Detailed traffic flow data is essential to provide the most accurate accident analyses (Norros et al., 2016), and there is a need for improved hourly bicycle volume estimates based on the limited data sources available.

In order to benefit the work of traffic agencies concerning the computation of performance indicators and safety analyses conducted by safety engineers, this study is concerned with estimating historic hourly bicycle estimates, similar to Sekuła et al. (2018). The current study focuses on improving the hourly estimates of bicycle volumes derived from mean daily measures, the estimation of which are the focus of recently developed large scale models (Aled Davies, 2017; Kjems & Paag, 2019). As such, the model in this study contrasts short-term traffic flow forecasts. The model should estimate historic bicycle flow unconditional of previous traffic flow, where only the mean-expected daily traffic is available.

Several current state-of-the-art approaches for estimating traffic flow use some variation of the Long Short-Term Memory (LSTM) neural network (Hochreiter & Schmidhuber, 1997). This statistical framework has

been adopted in many recent studies forecasting short-term traffic states/flow (Chen et al., 2016; Cui et al., 2020; Duan, Lv, & Wang, 2016; Ma et al., 2015; Zhao et al., 2017).

To overcome the shortcomings of the calibration-based method, this study explores the use of a novel neural network approach to estimating hourly bicycle volumes conditional on weather conditions, temporal effects and road conditions. The method is a hybrid of an LSTM and a Mixture Density Network (Bishop, 1994). This hybrid introduces a Gaussian mixture model extension to the traditional LSTM, enabling the model to estimate a conditional bicycle flow distribution in contrast to the traditional conditional mean estimation. Furthermore, the LSTMMDN improves the estimation of the hourly bicycle volumes by treating them as random draws from a distribution, thus introducing variation across the network even on measurably similar roads. Finally, we illustrate how the improved bicycle flow estimates can improve accident analyses conducted to create a safer cycling environment.

2. METHODOLOGY

In this section we describe the LSTMMDN framework used to derive the conditional bicycle flow distribution for this study along with the specific setup used in the investigation of the Copenhagen use-case.

LSTMMDN

The LSTMMDN is variation of the classical MDN presented in Bishop (1994), relying on an LSTM (Hochreiter & Schmidhuber, 1997) to estimate the mean, average and mixture coefficient, conditional on the input data X and a Gaussian Mixture Model (GMM) output, which is used to approximate the city-wide bicycle flow distribution. The conditional density computed in the output layer of the LSTMMDN is described in Equation (1)

$$p(y|X) = \sum_{i=1}^{A} \alpha_i(X) \frac{1}{\nu_i(x)^{1/2}} e^{-\frac{(t-\mu_i(x))^2}{2\nu_i(x)^2}} \tag{1}$$

where A is the number of mixture components, $\alpha_i(X)$ are the mixing coefficients dependent on the input data, $\mu_i(x)$ is the centre of the kernel, i.e. the conditional average, and $\nu_i(x)$ the associated variance. The benefit of the Mixture Density type Network for estimating city-wide hourly bicycle flow is derived from it approximating a conditional density function for the hourly bicycle flow. This allows for bicycle flow estimates to vary across different road sections, even in under equal measurable conditions, as all estimates are treated as draws from the conditional density function. This further addresses unobserved heterogeneity in hourly bicycle flow estimates.

Substituting the artificial neural network (ANN) in the traditional MDN proposed in Bishop (1994) is motivated by the LSTMs superior capability of handling nonlinear time series problems (Hochreiter & Schmidhuber, 1997). The LSTM learns to represent temporal data by introducing a memory cell and sub-processes, referred to as gates. There are three such gates in the LSTM cell: the input gate, the forget gate, and the output gate. The gates each handle different tasks, involving what information to keep from the previous cell state, what new input to consider, and which to include in the cell state. In Figure 1 there is a visual illustration of the LSTM cell layout, and the computations performed in the LSTM cell for each time step in the temporal sequence are shown in Equations (2) to (7).

$$f_t = \sigma(W_{fx}x_t + W_{fh}h_{t-1} + b_f) \tag{2}$$
$$i_t = \sigma(W_{ix}x_t + W_{ih}h_{t-1} + b_i) \tag{3}$$
$$\tilde{C}_t = tanh(W_{Cx}x_t + W_{Ch}h_{t-1} + b_C) \tag{4}$$
$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{5}$$
$$o_t = \sigma(W_{ox}x_t + W_{oh}h_{t-1} + b_o) \tag{6}$$
$$h_t = tanh(C_t) \times o_t \tag{7}$$

[1] https://upload.wikimedia.org/wikipedia/commons/3/3b/The_LSTM_cell.png

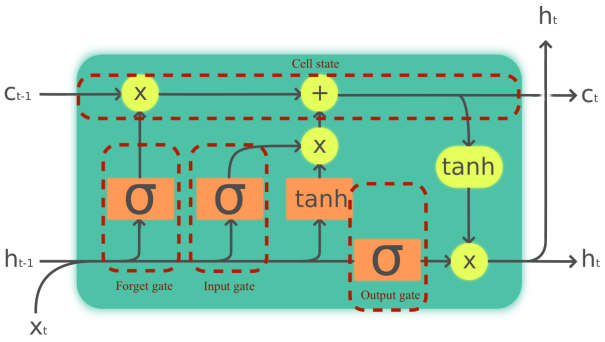

Figure 1: Graphical illustration of LSTM cell structure[1]: Orange squares mark layers with respective activation functions, yellow circles mark point wise operation, where x and + icons indicate respective element-wise multiplication and addition of tensors in the LSTM cell.

Here $\sigma(\cdot)$ and $\tanh(\cdot)$ are the respective activation functions, $W_{fx}, W_{fh}, W_{ix}, W_{ih}, W_{ox}, W_{oh}, W_{Cx}$ and W_{Ch} are the weight matrices of the respective gates f_t, i_t, o_t and the memory cell C_t, in the LSTM cell. b_f, b_i, b_o and b_C are intercept/bias terms of the respective gates, and h_t represents the hidden state at the time step t.

The LSTMMDN is estimated using back-propagation (Hastie, Tibshirani, & Friedman, 2009) in order to minimise the negative log-likelihood loss function in Equation (8).

$$\log \mathscr{L}(\Theta) = \frac{1}{N}\sum_{j}^{N} -\ln\{\sum_{i=1}^{A} \alpha_i(X_j)\phi_i(y_j|X_j)\} \tag{8}$$

Model configuration

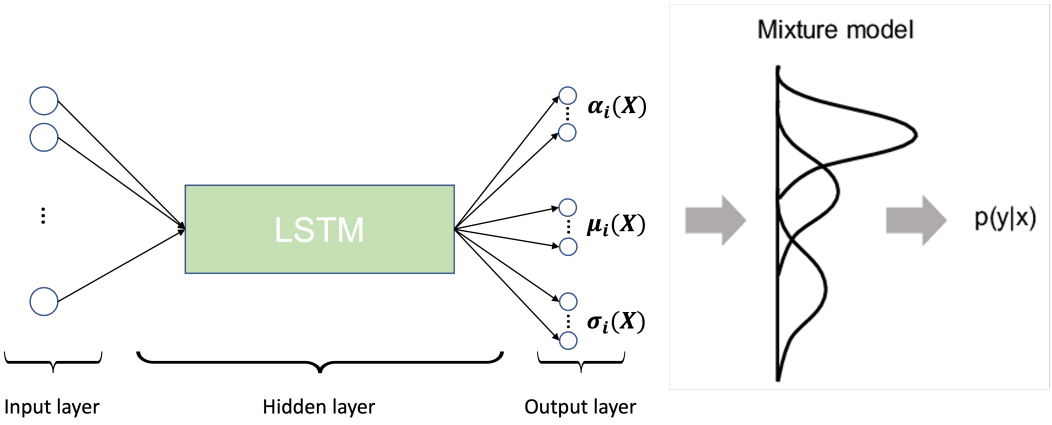

Figure 2: Graphical illustration of the LSTMMDN setup used in the study.

A visualisation of the LSTMMDN applied in this study is shown in Figure 2. It is configured using the activation functions described in Equations (2) to (7), various setups of k hidden neurons in each gate, and an internal dropout rate of 20% dropout in the LSTM cell (Srivastava et al., 2014). Meanwhile, we use a linear $a(x) = x$ in the output layer. The Mixture output layer itself we set up to consist of A mixtures, thus receiving input from $A \times 3$ nodes from the networks output layer shown in Figure 2. Finally, model estimation/training is done using the Adaptive Moment Estimation (Adam) algorithm (Kingma & Ba, 2015), using a training data set comprised of 70% of the total data. To avoid over-fitting the model, we implement an early stopping criteria such that the iterative optimisation of the model does not improve over $L = 100$ iterations when evaluated on a validation set comprised of 10% of the total data, not in the training data.

Minor headings

New paragraphs are not indented, but are preceded by a line of space. Also, please avoid using footnotes or splitting tables over two (or more) pages.

3. RESULTS

Data and experimental setup

The case studied in this paper, is that of bicycle flow in Copenhagen in the period 2017-2020. The LSTM-MDN is trained and evaluated using aggregated hourly bicycle volumes reported from bicycle monitoring stations in Copenhagen from 2017-2020, weather data provided by the Danish Meteorological Institute (DMI) and information on the Danish bank holidays. A total of 37 bicycle monitoring stations were active in the period. These are marked by the blue dots in Figure 3. The 37 stations showed considerable variation in active monitoring days, ranging from 9-731 days, with the median number of active days being 40. Overall, a total of 64,664 hourly bicycle volumes were recorded in the period. For each year, for each monitoring station, and Annual Average Weekday Cycling Traffic estimate is included in the data reports.

DMI weather data for Copenhagen in the period 2017-2020 was recorded at the main weather station located at the green dot in Figure 3 and obtained using their Open Data API (DMI, 2021). The variables considered in the model are 10-minute averages of the air temperature, pressure, wind speed, wind gusts, wind direction, precipitation levels, visibility and snow volume.

Finally, we include temporal data such as hour of the day, day of the week, week of the year, and whether days are public holidays, as this would be assumed to strongly influence cyclist ridership.

In total, the data consist of 64,664 observations of hourly bicycle volumes and $\approx 388,000$ weather and temporal measurements at 10-minute intervals. The hourly aggregated bicycle volumes reported at each station are used as response variables to train the model, while the AAWCT at the monitoring station, 10-minute weather data and holiday status of a day are used as predictors. The 10-minute interval predictors $\boldsymbol{x}_t$ are grouped such that a sequence of six 10-minute interval observations $\boldsymbol{x}_t$ are paired with the matching response (i.e. the accumulated bicycle volume of the hour) as shown below.

$$y_T(\boldsymbol{X}_T): \quad \boldsymbol{X}_t = \begin{pmatrix} \boldsymbol{x}_{t1} \\ \boldsymbol{x}_{t2} \\ \boldsymbol{x}_{t3} \\ \boldsymbol{x}_{t4} \\ \boldsymbol{x}_{t5} \\ \boldsymbol{x}_{t6} \end{pmatrix} = \begin{pmatrix} (x_{1,t1}, x_{2,t1}, \ldots, x_{D,t1}) \\ (x_{1,t2}, x_{2,t2}, \ldots, x_{D,t2}) \\ (x_{1,t3}, x_{2,t3}, \ldots, x_{D,t3}) \\ (x_{1,t4}, x_{2,t4}, \ldots, x_{D,t4}) \\ (x_{1,t5}, x_{2,t5}, \ldots, x_{D,t5}) \\ (x_{1,t6}, x_{2,t6}, \ldots, x_{D,t6}) \end{pmatrix} \tag{9}$$

Where $t1, t2, t3, t4, t5, t6$ indicate the first to sixth 10-minute intervals of the input sequence matching the aggregated hourly cycling volume, T indicates the full hour considered, and D is the number of predictors used.

Bicycle flow estimation

We assess the LSTMMDN's performance and compare it against 1) other neural network architectures and 2) the profiling method based on calibration factors currently employed by the Danish road directorate (the Seasonal Variation Factors, SVF), using a test data set containing the remaining 20% of the total data.

Four models for the estimation are being compared in various setups, the ANN setup proposed in Sekuła et al. (2018), LSTMs and LSTMMDNs in various setups, and the SVF-based method used by the Danish government.

The performance measures are the average mse ($\widehat{mse}$) and the average negative log-likelihood ($-\log \hat{\mathscr{L}}$) of 100 posterior draws from the LSTMMDN, the mse of the conditional average (mse_μ, and the negative

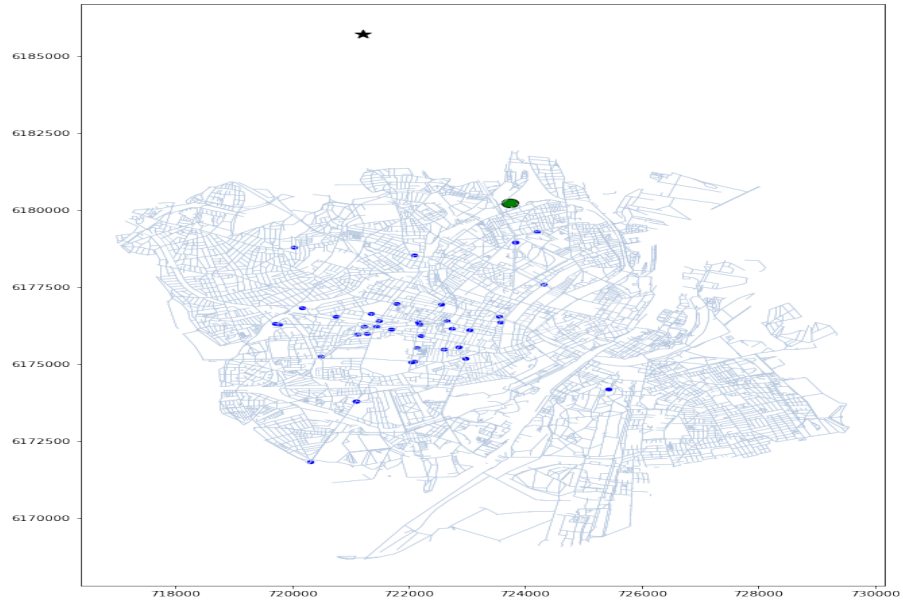

Figure 3: Map of the measurement stations and weather stations in Copenhagen

log-likelihood of the conditional average and $(-\log \mathscr{L}_\mu)$ for all models (also using 100 posterior draws from the LSTMMDN).

The results are shown in Table 1.

Table 1: Goodness-of-fit measures related to various LSTMMDN structures and the SVFs, determined on the test data.

model specification	$-\log \mathscr{L}_\mu$	$-\log \hat{\mathscr{L}}$	MSE_μ	$\hat{MSE}$	Trainable parameters
ANN: (Sekuła et al., 2018)	5772	-	0.108	-	162,025
LSTM: $L(k=32)$	5851	-	0.129	-	6,561
LSTM: $L(k=64)$	5792	-	0.114	-	21,313
LSTM: $L(k=32)\times 6$	5832	-	0.128	-	6,733
LSTM: $L(k=64)\times 6$	5777	-	0.110	-	21,645
LSTMMDN: $L(k=32)\times G(A=6)$	5805	5929	0.119	0.226	7,122
LSTMMDN: $L(k=32)\times G(A=8)$	5811	5933	0.121	0.234	7,320
LSTMMDN: $L(k=64)\times G(A=6)$	5753	5864	0.102	0.195	22,418
LSTMMDN: $L(k=64)\times G(A=8)$	5777	5872	0.109	0.201	22,808
SVF method	6570	-	0.377	-	-

As we see in Table 1, the LSTMMDNs outperforms its counterpart LSTMs in similar setups, i.e. the same number of LSTM nodes k and similar output dense layer. The best performing LSTMMDN setup has $k=64$ computational nodes in each LSTM gate and $A=6$ mixtures to describe the city-wide cycling flow distribution. This model also proves superior to the model the ANN model from Sekuła et al. (2018), while one having $1/8$ of the estimable parameters. As seen in Table 1 all LSTMMDNs are superior to the SVF calibration method employed by the Danish road directorate. Based on the MSE_μ we see that that the best LSTMMDN produces $\sim 75\%$ more accurate bicycle flow estimates than the calibration-based method.

In comparing the LSTMMDN and SVF calibration approach further, we only use the best LSTMMDN (LSTMMDN: $L(k=64)\times G(A=6)$) and refer to this as the LSTMMDN.

LSTMMDN vs. SVF-calibration method

We see the the hourly bicycle flow estimates from the LSTMMDN superiority over the SVF-calibration method when visualised in an estimated vs. actual flow heat plot in Figure 4. The estimates produces by the LSTMMDN are much closer to the 45° line establishing a $x=y$ relation. Furthermore, it is evident that

the calibration method has a tendency to overestimate the hourly bicycle flow.

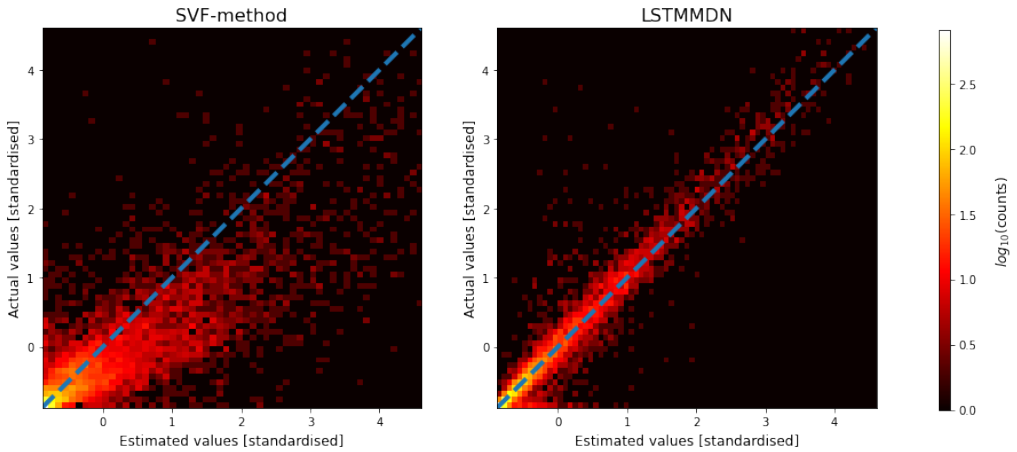

Figure 4: Heat maps comparing the prediction accuracy of the SVF method (left) and the LSTMMDN method (right) based on the reserved test data set.

Lastly, we show a continuous week of standardised hourly bicycle flow estimates for an arbitrary counting station during an arbitrary week in Figure 5. Here we see the actual hourly bicycle flow (blue), the LSTMMDN estimates (orange) and the calibration-based estimates (green). As expected, the LSTMMDN estimates follow the observed bicycle flow during this period much more better than the calibration-based estimates, representing peak traffic and weekend traffic much better. The most significant difference between the LSTMMDN and the calibration-method in Figure 5 is the discrepancy of the calibration method to estimate the traffic at weekends accurately.

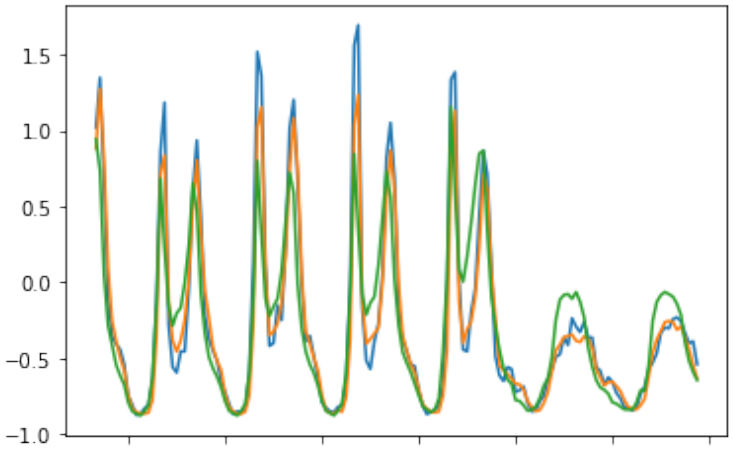

Figure 5: Visualisation of the standardised bicycle flow over 148 continuously registered hours from a randomly picked counting station. Blue: Actual counts of bicyclists from random counting station, orange: LSTMMDN traffic counts estimation of the same road section and week and time: SVF based estimation of same road section and time.

Impact of bicycle exposure on bicycle accident analyses

Cyclist safety analysis is a topic that concerns transport agencies which could benefit significantly from improved bicycle flow estimates.

We exemplify this by estimating three separate city-wide bicycle crash frequency models for Copenhagen between 2017 and 2020, only varying the exposure variable. The considered exposure variables are, AAWCT, the calibration based hourly flow estimates or the LSTMMDN hourly flow estimates. To estimate the city-wide crash frequency, we use a simple Poisson regressions. While it is not the most advanced model type, it should be sufficient to illustrate the effect of various exposure variables.

We consider for the response variable the aggregated number of bicycle crashes in Copenhagen during

any given hour in the period 2017-2020, and consider the following exogenous variables: Visibility $[m]$, Temperature $< 0°C$, Temperature $> 20°C$, Morning peak hours (7-9 weekday), Afternoon peak (15-17 weekday), Wind speed $< 5m/s$, Wind speed $> 9m/s$, Precipitation ($> 0mm$), Bank holidays, Bicycle flow.

The cyclist exposure is estimated hourly at each of the counting stations in Figure 3 given the AAWCT, weather and temporal information for the period 2017-2020, using respectively the calibration method and LSTMMDN. The estimates are then aggregated to approximate the city-wide cyclist exposure at any given hour. The AAWCT exposure is divided to approximate the annual average hourly cycling traffic.

Due to missing observations in the weather data, the weather, and therefore bicycle flow estimates, in the period 2017-2020 becomes imbalanced with an over-representation of summer days. Therefore, stratified bootstrapping is performed to sample three full years of estimated bicycle flow and weather data from the four years of data, conditional on the year's month, day, and daytime. The three years bootstrap data includes 2104 bicycle crashes in 26,232 hours, meaning 0.08 accidents per hour. Summary statistics of the data used in the bicycle crash frequency model are shown in Table 2.

Variable	Mean
Visibility	27,519 m
Bank holiday	0.034
Exposure(LSTMMDN)	5430 cyclists/hour
Exposure(Exposure)	6051 cyclists/hour
Exposure(AAWCT)	6079 cyclists/hour
Morning Peak/ Afternoon peak	0.089
Temperature $< 0°C$	0.032
Temperature $> 20°C$	0.075
Wind speed $< 5m/s$	0.102
Wind speed $> 9m/s$	0.008
Precipitation $> 0mm$	0.429

Table 2: Summary statistics of the data used in the Poisson model

The three models are estimated using Maximum Likelihood estimation, and the resulting parameter estimates and goodness of fit measures are shown in Table 3.

	Model 1 (AAWCT based exposure)		Model 2 (SVF-calibration based exposure)		Model 3 (LSTMMDN based exposure)	
No. Obs	26,232		26,232		26,232	
Estimated parameters	10		10		10	
log-likelihood	−7,142		−6,874		−6,740	
Deviance	10,287		9,750		9,483	
χ^2	27,524		27,704		29,810	
Variables	Parameter estimates	p-value	Parameter estimates	p-value	Parameter estimates	p-value
Intercept	−23.390	< 0.001	−7.958	< 0.001	−10.740	< 0.001
Visibility (for one log change)	0.071	0.029	−0.068	0.037	−0.062	0.055
Bank Holiday	−0.857	< 0.001	−0.809	< 0.001	−0.096	0.609
log(Exposure)	2.267	< 0.001	0.697	< 0.001	1.006	< 0.001
Morning peak	1.098	< 0.001	0.210	0.002	0.247	< 0.001
Afternoon peak	1.194	0.003	0.296	< 0.001	0.284	< 0.001
Temperature $< 0°C$	−0.512	< 0.001	−0.113	0.519	−0.027	0.877
Temperature $> 20°C$	0.422	0.774	0.075	0.270	0.224	40.001
Wind speed $< 5m/s$	−0.023	0.946	0.181	0.025	0.136	0.090
Wind speed $> 9m/s$	−0.018	0.001	−0.058	0.824	−0.030	0.908
Precipitation	0.147	0.001	0.095	0.031	0.124	0.005

Table 3: Estimates for the Poisson regressions with varying exposure variables (AAWCT, calibration based hourly cycling volume and LSTMMDN model based hourly cycling volume).

The goodness-of-fit measures clearly indicate that the model using LSTMMDN exposure estimates is the superior of the models. The resulting log-likelihood is 5.5% higher in Model 3 than in Model 1. Having only varied the exposure variable in the three models, this highlights that the exposure estimates used in models have a substantial impact on the conclusions drawn.

Furthermore, the estimates in Table 3 reveal that conclusions regarding various variable impacts on the

bicycle crash frequency would vary depending on which of the three model variations is employed. For example both low temperatures and high wind speeds would lower the crash risk in Model 1, contrasting the findings of the other two models.

4. DISCUSSION AND CONCLUSIONS

The current approach used by transport agencies to estimate hourly bicycle traffic relies on calibration factors that do not to account for weather and other factors' impacts on cyclist ridership. This study aims to amend this issue, proposing a LSTMMDN model to estimate city-wide hourly bicycle volumes in Copenhagen conditional on the mean-daily traffic, weather and temporal effects.

The proposed approach significantly outperforms the calibration factor method. Conditional on the model setup, the LSTMMDN yields $\sim 75\%$ more accurate estimates of the hourly bicycle volume in Copenhagen. Furthermore, the method introduces a framework that allows for variation in the bicycle volume estimates across the city network even under similar conditions and on measurably similar roads. This is achieved by the model approximating a conditional distribution for the hourly bicycle volume instead of only approximating the conditional average.

The LSTMMDN generated bicycle flow estimates were subsequently used to quantify the improvements that better exposure data can have on accident analyses. The results indicate that improving the accuracy of bicycle volume estimates will result in better crash risk models.

On a side-note, the crash model and the variables used for those models are very simplistic and quantify the impact that better bicycle volume estimates can have on the accident models that are crucial to making informed decisions to increase cycling safety. Many studies investigate the factors related to bicycle crashes and the outcomes thereof (Aldred et al., 2018; Fountas et al., 2021; Janstrup et al., 2019; Kaplan & Prato, 2013; Kim et al., 2007; Myhrmann et al., 2021; Schepers et al., 2020), and this study does not try to conduct a deep risk analysis of bicycle crashes. Nevertheless, the results of the three Poisson regressions indicate that the bicycle flow estimates affect the output. This result makes it plausible that the same would be the case in more advanced models. However, a definitive clarification would require an in-depth analysis and is the subject of future studies.

CONCLUSIONS

Overall, the impact of better cyclist exposure estimates is non-negligible, as revealed in a simple-crash risk model comparison. The results revealed significant improvements by simply addressing the exposure estimates used in the model, with all other variables held constant. Also, the interpretation of variables changed in light of better exposure estimates hinting potentially at previously erroneous inference from crash risk analyses caused by the lack of decent exposure estimates. And considering the recent development of models to estimate daily cycling traffic in the city networks such as the Cynemon model for London and COMPASS in Copenhagen, the model proposed out this study should be a welcomed extension in order to optimally represent the hourly bicycle traffic conditional on the proposed mean-daily estimates and other external factors.

ACKNOWLEDGEMENT

We thank Filipe Rodrigues and Mads Paulsen for valuable feedback on an earlier version of the paper process.

REFERENCES

Aldred, R., Goodman, A., Gulliver, J., & Woodcock, J. (2018, 8). Cycling injury risk in London: A case-control study exploring the impact of cycle volumes, motor vehicle volumes, and road characteristics including speed limits. *Accident Analysis & Prevention*, *117*, 75–84. Retrieved from https://www.sciencedirect.com/science/article/

pii/S0001457518301076?via%3Dihub doi: 10.1016/J.AAP.2018.03.003

Aled Davies. (2017). *Cynemon-Cycling Network Model for London Aled Davies-TfL Planning* (Tech. Rep.). Transport for London. Retrieved from https://www.ucl.ac.uk/transport/sites/transport/files/Davies_slides.pdf

Bishop, C. M. (1994). *Mixture Density Networks* (Tech. Rep.). Department of Computer Science and Applied Mathematics, Aston University. Retrieved from http://www.ncrg.aston.ac.uk/

Böcker, L., Dijst, M., & Prillwitz, J. (2013, 1). Impact of Everyday Weather on Individual Daily Travel Behaviours in Perspective: A Literature Review. *Transport Reviews*, *33*(1), 71–91. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/01441647.2012.747114 doi: 10.1080/01441647.2012.747114

Chen, Y. Y., Lv, Y., Li, Z., & Wang, F. Y. (2016). Long short-Term memory model for traffic congestion prediction with online open data. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 132–137. doi: 10.1109/ITSC.2016.7795543

Cui, Z., Ke, R., Pu, Z., & Wang, Y. (2020). Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values. *Transportation Research Part C: Emerging Technologies*, *118*(March 2019), 102674. Retrieved from https://doi.org/10.1016/j.trc.2020.102674 doi: 10.1016/j.trc.2020.102674

DMI. (2021). *Danish Meteorological Institute - Open Data - DMI Open Data - Confluence.* Retrieved from https://confluence.govcloud.dk/display/FDAPI/Danish+Meteorological+Institute+-+Open+Data

Dozza, M. (2017, 8). Crash risk: How cycling flow can help explain crash data. *Accident Analysis & Prevention*, *105*, 21–29. Retrieved from https://www.sciencedirect.com/science/article/pii/S0001457516301464 doi: 10.1016/J.AAP.2016.04.033

Duan, Y., Lv, Y., & Wang, F. Y. (2016). Travel time prediction with LSTM neural network. *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*, 1053–1058. doi: 10.1109/ITSC.2016.7795686

Fountas, G., Fonzone, A., Olowosegun, A., & McTigue, C. (2021, 12). Addressing unobserved heterogeneity in the analysis of bicycle crash injuries in Scotland: A correlated random parameters ordered probit approach with heterogeneity in means. *Analytic Methods in Accident Research*, *32*, 100181. doi: 10.1016/J.AMAR.2021.100181

Hastie, T., Tibshirani, R., & Friedman, J. (2009). Neural Networks. In *The elements of statistical learning data mining, inference, and prediction* (pp. 389–416). Retrieved from http://link.springer.com/10.1007/978-0-387-84858-7_11 doi: 10.1007/978-0-387-84858-7{_}11

Hochreiter, S., & Schmidhuber, J. (1997, 11). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. Retrieved from https://www.mitpressjournals.org/doi/abs/10.1162/neco.1997.9.8.1735 doi: 10.1162/neco.1997.9.8.1735

Janstrup, K. H., Møller, M., & Pilegaard, N. (2019, 5). A clustering approach to integrate traffic safety in road maintenance prioritization. *Traffic Injury Prevention*, *20*(4), 442–448. Retrieved from https://www.tandfonline.com/doi/full/10.1080/15389588.2019.1580700 doi: 10.1080/15389588.2019.1580700

Kaplan, S., & Prato, C. G. (2013). Cyclist-motorist crash patterns in Denmark: a latent class clustering approach. *Traffic injury prevention*, *14*(7), 725–733. doi: 10.1080/15389588.2012.759654

Kaplan, S., Vavatsoulas, K., & Prato, C. G. (2014, 9). Aggravating and mitigating factors associated with cyclist injury severity in Denmark. *Journal of Safety Research*, *50*, 75–82. Retrieved from https://www.sciencedirect.com/science/article/pii/S0022437514000437?via%3Dihub doi: 10.1016/j.jsr.2014.03.012

Kim, J.-K., Kim, S., Ulfarsson, G. F., & Porrello, L. A. (2007, 3). Bicyclist injury severities in bicycle–motor vehicle accidents. *Accident Analysis & Prevention*, *39*(2), 238–251. Retrieved from https://www.sciencedirect.com/science/article/pii/

S000145750600128X doi: 10.1016/J.AAP.2006.07.002

Kingma, D. P., & Ba, J. L. (2015, 12). Adam: A method for stochastic optimization. In *3rd international conference on learning representations, iclr 2015 - conference track proceedings* (p. 15). International Conference on Learning Representations, ICLR.

Kjems, S., & Paag, H. (2019). COMPASS: Ny trafikmodel for hovedstadsområdet. In *Proceedings from the annual transport conference at aalborg university* (p. 4). Copenhagen. Retrieved from www.trafikdage.dk/artikelarkiv

Ma, X., Tao, Z., Wang, Y., Yu, H., & Wang, Y. (2015, 5). Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies*, *54*, 187–197. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S0968090X15000935 doi: 10.1016/j.trc.2015.03.014

Myhrmann, M. S., Janstrup, K. H., Møller, M., & Mabit, S. E. (2021, 1). Factors influencing the injury severity of single-bicycle crashes. *Accident Analysis and Prevention*, *149*, 105875. doi: 10.1016/j.aap.2020.105875

Nankervis, M. (1999, 8). The effect of weather and climate on bicycle commuting. *Transportation Research Part A: Policy and Practice*, *33*(6), 417–431. doi: 10.1016/S0965-8564(98)00022-6

Norros, I., Kuusela, P., Innamaa, S., Pilli-Sihvola, E., & Rajamäki, R. (2016). The Palm distribution of traffic conditions and its application to accident risk assessment. *Analytic Methods in Accident Research*, *12*(March), 48–65. Retrieved from http://dx.doi.org/10.1016/j.amar.2016.10.002 doi: 10.1016/j.amar.2016.10.002

Nosal, T., & Miranda-Moreno, L. F. (2014). The effect of weather on the use of North American bicycle facilities: A multi-city analysis using automatic counts. *Transportation Research Part A: Policy and Practice*, *66*(1), 213–225. doi: 10.1016/j.tra.2014.04.012

Rossetti, T., Guevara, C. A., Galilea, P., & Hurtubia, R. (2018, 5). Modeling safety as a perceptual latent variable to assess cycling infrastructure. *Transportation Research Part A: Policy and Practice*, *111*, 252–265. doi: 10.1016/j.tra.2018.03.019

Schepers, P., de Geus, B., van Cauwenberg, J., Ampe, T., & Engbers, C. (2020). The perception of bicycle crashes with and without motor vehicles: Which crash types do older and middle-aged cyclists fear most? *Transportation Research Part F: Traffic Psychology and Behaviour*, *71*, 157–167. Retrieved from https://doi.org/10.1016/j.trf.2020.03.021 doi: 10.1016/j.trf.2020.03.021

Schrank, D. (2021). *2021 Urban Mobility Report – Appendix A: Methodology* (Tech. Rep.). Texas A&M Transportation Institute. Retrieved from http://mobility.tamu.edu/umr/congestion-data/

Sekuła, P., Marković, N., Vander Laan, Z., & Sadabadi, K. F. (2018). Estimating historical hourly traffic volumes via machine learning and vehicle probe data: A Maryland case study. *Transportation Research Part C: Emerging Technologies*, *97*(December 2017), 147–158. Retrieved from https://www.sciencedirect.com/science/article/pii/S0968090X18314773?dgcid=rss_sd_all doi: 10.1016/J.TRC.2018.10.012

Srivastava, N., Hinton, G., Krizhevsky, A., & Salakhutdinov, R. (2014). *Dropout: A Simple Way to Prevent Neural Networks from Overfitting* (Vol. 15; Tech. Rep.). Department of Computer Science University of Toronto.

Twisk, D. A. M., & Reurings, M. (2013). An epidemiological study of the risk of cycling in the dark: The role of visual perception, conspicuity and alcohol use. *Accident Analysis and Prevention*, *60*, 134–140. doi: 10.1016/j.aap.2013.08.015

Vandenbulcke, G., Thomas, I., & Int Panis, L. (2014, 1). Predicting cycling accident risk in Brussels: A spatial case–control approach. *Accident Analysis & Prevention*, *62*, 341–357. Retrieved from https://www.sciencedirect.com/science/article/pii/S0001457513002686 doi: 10.1016/J.AAP.2013.07.001

Zhao, Z., Chen, W., Wu, X., Chen, P. C., & Liu, J. (2017). LSTM network: A deep learning approach for short-term traffic forecast. *IET Image Processing*, *11*(1), 68–75. doi: 10.1049/iet-its.2016.0208