# COMPARING GRAPH NEURAL NETWORK-BASED TRAFFIC SPEED PREDICTION FOR STATIC SENSOR AND FLOATING VEHICLE DATA SETS

**J VISAGIE[1]*, AJ HOFFMAN[1]** and J JOUBERT[1]**

[1]North-West University Potchefstroom School for Electrical, Electronic and Computer Engineering, Private Bag X6001, 2520 Potchefstroom, South Africa;
Email: *20270186@nwu.ac.za; **alwyn.hoffman@nwu.ac.za

## ABSTRACT

Traffic speed prediction using deep learning neural networks has been the topic of many studies, mostly using data sets that were collected from static sensors. Floating vehicle data offers a more flexible alternative, as it can be obtained for any roads travelled by GPS tracked vehicles. In this paper we compare the performance of leading traffic speed prediction techniques when applied to both static sensor and floating vehicle data sets. Data sets were collected for the road networks serving Johannesburg, representing South Africa's most congested roads. Based on prediction accuracy, training time and robustness the Graph WaveNet method produced the best results. We found that the static sensor and floating vehicle data sets, representing traffic movements on the same sets of roads, produced comparable results, providing evidence that static sensor data can be complemented and, in some cases, replaced by floating vehicle data. This will enable traffic speed prediction for roads where no static sensors are installed, resulting in significant cost savings. Extending traffic speed prediction to all major roads will result in improved traffic management strategies for the overall road network, leading to less congestion and an improved road user experience.

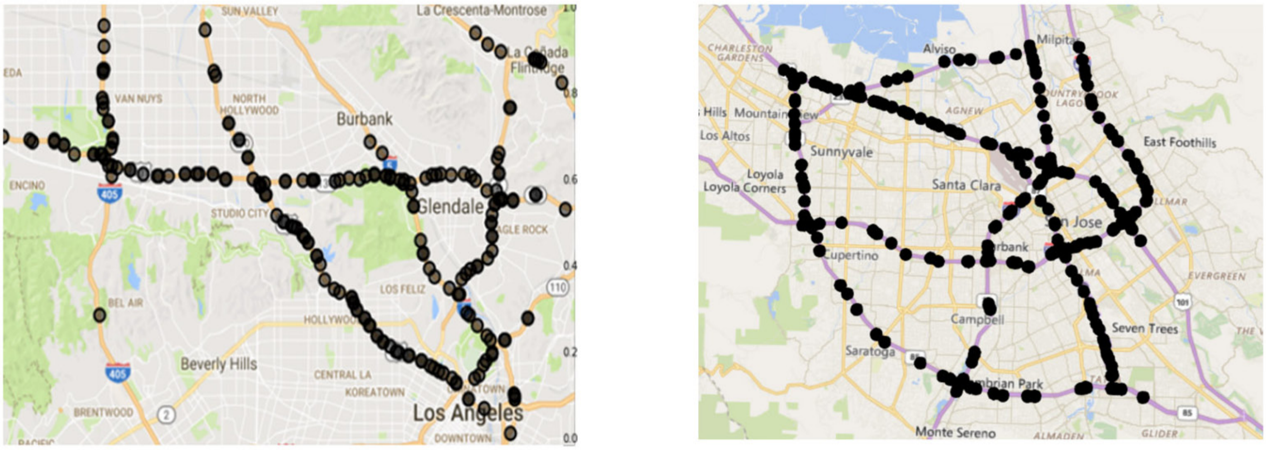
## 1. INTRODUCTION

Traffic speed prediction is an important part of managing metropolitan traffic networks (Kumar & Raubal, 2021). The high costs associated with congestion in big cities is one of the main reasons for the recent interest traffic speed prediction (Mena-Oreja & Gozalvez, 2020), (Polson & Sokolov, 2017). From a recent study the revenue lost in some Australian cities due to traffic congestion was approximately $16.5 billion (Tedjopurnomo, 2022). Traffic congestion can be alleviated by better traffic management (United States Department of Transportation, 2020) and can be accomplished by implementing an Intelligent Transportation System (ITS) to regulate traffic. An integral part of an ITS is traffic speed prediction (Wang et al., 2019).

Models like auto-regression and conventional supervised neural networks have been used for traffic speed forecasting but display limited accuracy according to (Wang, 2016). Deep learning neural networks have been one of the most successful techniques used for time series prediction and many approaches have been published (Mena-Oreja & Gozalvez, 2020). Speed prediction methods are evaluated using performance measures such as prediction accuracy, robustness under various conditions, computational intensity of the learning process and others (Mallick et al., 2021).

## 1.1 Historical Experimentation

Historic datasets play a vital role in developing models as well as improving existing models for traffic speed prediction. These datasets should include enough relevant information to train the model and to be considered a benchmark dataset for integration into other models or traffic networks. Two datasets namely the METR-LA and PEMS-BAY have been widely used as baseline datasets to evaluate methods used to predict future traffic speeds. The METR-LA dataset which is located primarily on the Los Angeles highway uses traffic data from 207 sensors across March 1$^{st}$, 2012, to June 30$^{th}$, 2012, and the PEMS-BAY dataset has 325 sensors collecting data across January 1$^{st}$, 2017, to May 31$^{st}$, 2017. For both datasets the data is captured in 5 minute intervals (Li et al., 2017). These two datasets have their sensors spread on certain road segments as shown in Figure 1 below.




(a) METR-LA

(b) PEMS-BAY

**Figure 1: All the sensors located on the road segments for (a) METR-LA and (b) PEMS-BAY datasets adapted (Li et al., 2017)**

To establish a performance reference conventional ARIMA (autoregressive integrated moving average) prediction methods as well as a DCRNN (diffusion convolutional recurrent neural network) method, a recurrent neural network that does not use a graph method to incorporate spatial dependencies, were applied to the data. The prediction results achieved with these methods were compared against the results produced by the Graph WaveNet method (Wu et al., 2019). These results for the METR-LA and PEMS-BAY datasets are displayed in Table 1 and Table 2 and are adapted from various sources that used different prediction techniques. The metrics used in testing to measure the results are the Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE).

As seen in Table 1 and Table 2, the Graph WaveNet technique achieved the best results on the METR-LA and PEMS-BAY datasets. By using these datasets as benchmarks, the performance of the Graph WaveNet technique when applied to South African datasets can be measured.

**Table 1: Results for the METR-LA dataset (Papers with code, 2024a)**

| Models | 15 min | | | 30 min | | | 60 min | | |
|---|---|---|---|---|---|---|---|---|---|
| | MAE | RMSE | MAPE | MAE | RMSE | MAPE | MAE | RMSE | MAPE |
| ARIMA | 3.99 | 8.21 | 9.60% | 5.15 | 10.45 | 12.70% | 6.90 | 13.23 | 17.40% |
| DCRNN | 2.77 | 5.38 | 7.30% | 3.15 | 6.45 | 8.80% | 3.60 | 7.60 | 10.50% |
| Graph WaveNet | 2.69 | 5.15 | 6.90% | 3.07 | 6.22 | 8.37% | 3.53 | 7.37 | 10.01% |

**Table 2: Results for the PEMS-BAY dataset (Papers with code, 2024b)**

| Models | 15 min | | | 30 min | | | 60 min | | |
|---|---|---|---|---|---|---|---|---|---|
| | MAE | RMSE | MAPE | MAE | RMSE | MAPE | MAE | RMSE | MAPE |
| ARIMA | 1.62 | 3.30 | 3.50% | 2.33 | 4.76 | 5.40% | 3.38 | 6.50 | 8.30% |
| DCRNN | 1.38 | 2.95 | 2.90% | 1.74 | 3.97 | 3.90% | 2.07 | 4.74 | 4.90% |
| Graph WaveNet | 1.30 | 2.74 | 2.73% | 1.63 | 3.70 | 3.67% | 1.95 | 4.52 | 4.63% |

## 2. RESEARCH PROBLEM STATEMENT

Traffic congestion is problem that incurs many unnecessary costs. These costs can be reduced by managing traffic more effectively using Intelligent Transportation Systems (ITSs). An integral part of these ITSs involves traffic speed prediction. Deep learning neural networks have been proven to be successful at predicting traffic speeds using static sensor data. The problem is that not all roads can be equipped with static sensors due to budget constraints. The ability to gather speed information per road segment using vehicle GPS data makes it possible to extend traffic speed prediction to roads not equipped with static sensors. Traffic speed data gathered from GPS tracking data is called floating vehicle data (FVD).

The aim of this paper is to compare the Graph WaveNet technique on static sensor and floating vehicle datasets gathered on South African roads. The Graph WaveNet technique produced satisfactory results when tested on the California Transportation Agencies (CalTrans) Performance Measurement System (PEMS) dataset as well as the Los Angeles Metropolitan Transportation Authority (METR-LA) dataset, which are regarded as benchmark datasets.

The aim was to determine if the same techniques that provide satisfactory results on the benchmark static sensor datasets will also be useful when applied to a static sensor dataset collected from SANRAL static sensors as well as on a floating vehicle dataset acquired from the company INRIX, both datasets being recorded at the same locations and for similar time frames. These results could indicate whether static sensor datasets can be complemented by floating vehicle datasets to achieve wider coverage when implementing intelligent traffic management.

## 3. DATA

The datasets that were used in testing are the floating vehicle INRIX dataset and the static sensor SANRAL dataset. The METR-LA and PEMS-BAY datasets were collected mainly on highways, thus only data collected on highways were used in the INRIX and SANRAL datasets.

## 3.1 INRIX Data

An INRIX road segment is a basic element of a road network used to measure the speeds and detect incidents on a road. INRIX uses two types of segments, namely XD Segments and TMC segments. TMC segments are an encoding of TMC location tables onto the linear road network and were used in this experiment. TMC tables are typically created and maintained by various third-party entities at a country level. The international standards body, TISA, reviews and certifies these location tables. The INRIX dataset used for this study consists of 357 road segments in the Johannesburg area as shown in Figure 2.

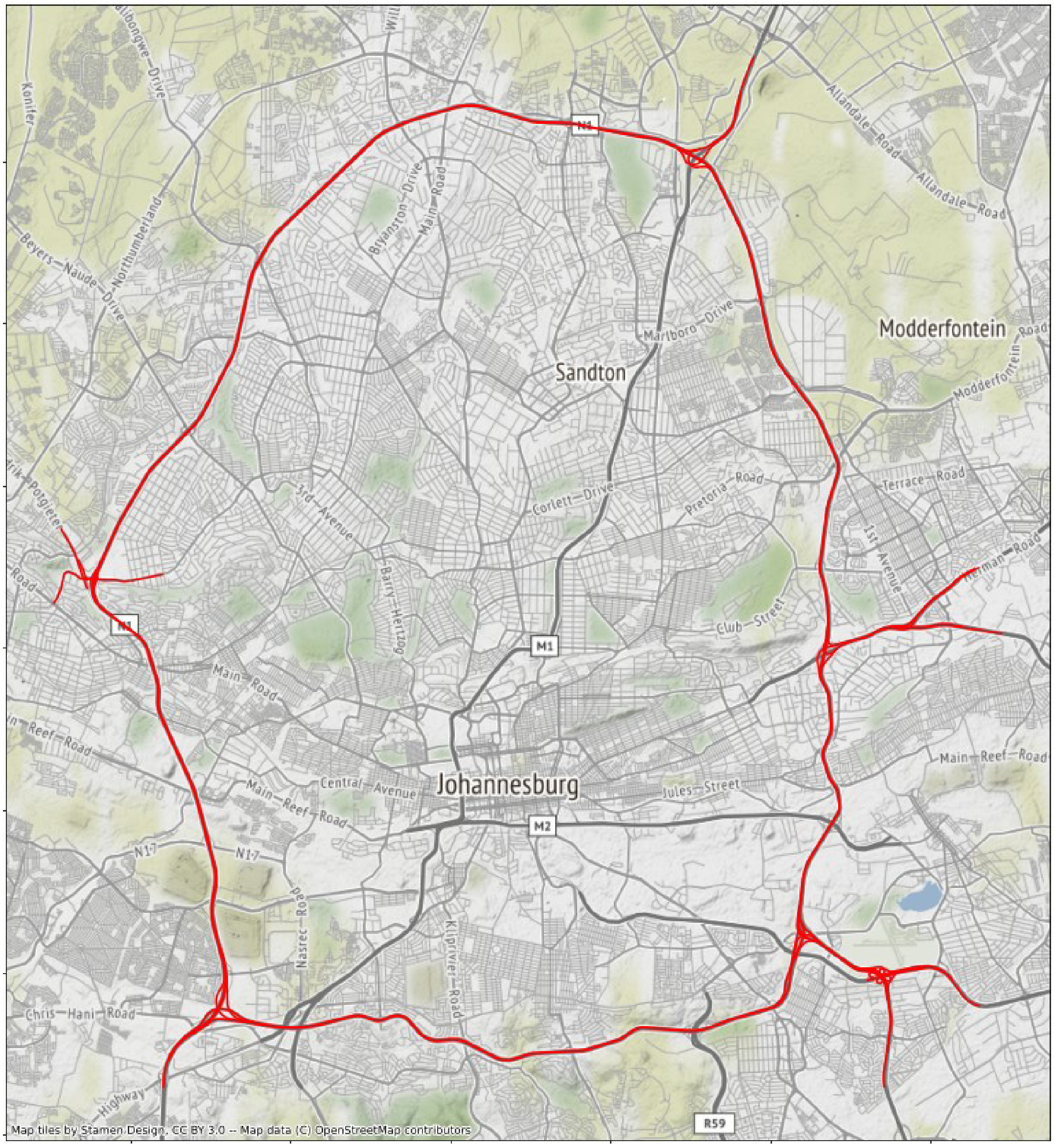



**Figure 2: INRIX Johannesburg Road Segments used for this study**

The data received from INRIX uses 5-minute sampling intervals and is based on tracking data obtained from approximately 10% of all the vehicles travelling on the respective road segments at that time. It had to be formatted to be suitable for training the Graph WaveNet model. Examples of the received and formatted data can be seen in Figure 3 and Figure 4.

| | timestamp | code | type | speed | average | reference | traveltimeminutes | speedbucket |
|---|---|---|---|---|---|---|---|---|
| 0 | 2022-07-26 22:00:00 | A17-00102ZA | TMC | 85.0 | 92 | 92 | 1.688 | 3.0 |
| 1 | 2022-07-26 22:00:00 | A17+00103ZA | TMC | 100.0 | 101 | 101 | 1.334 | 3.0 |
| 2 | 2022-07-26 22:00:00 | A17-00103ZA | TMC | 82.0 | 89 | 89 | 0.878 | 3.0 |
| 3 | 2022-07-26 22:00:00 | A17+00104ZA | TMC | 121.0 | 92 | 92 | 0.316 | 3.0 |
| 4 | 2022-07-26 22:00:00 | A17-00104ZA | TMC | 84.0 | 97 | 97 | 0.637 | 2.0 |

**Figure 3: INRIX Received Data Format**

| | A17P19364ZA | A17-00487ZA | A17+00487ZA | ... | A17N08477ZA | A17N08478ZA | A17N08455ZA |
|---|---|---|---|---|---|---|---|
| **2022-02-07 05:20:00** | 87.0 | 90.0 | 19.0 | ... | 69.0 | 53.0 | 21.0 |
| **2022-02-07 05:25:00** | 85.0 | 80.0 | 19.0 | ... | 71.0 | 56.0 | 24.0 |
| **2022-02-07 05:30:00** | 85.0 | 79.0 | 19.0 | ... | 72.0 | 56.0 | 16.0 |
| **2022-02-07 05:35:00** | 84.0 | 92.0 | 19.0 | ... | 80.0 | 42.0 | 11.0 |
| **2022-02-07 05:40:00** | 85.0 | 101.0 | 23.0 | ... | 76.0 | 29.0 | 14.0 |

**Figure 4: INRIX Formatted Data**

## 3.2 SANRAL Data

For the SANRAL dataset, CTO (Comprehensive Traffic Observations) data gathered from static sensors that exist along the same road segments as for the INRIX data was used. The data that was made available by SANRAL is for the first 4 months of 2022, thus that timeframe was also used for the INRIX data. One important difference is that the SANRAL data used a 15-minute sampling interval. While the Sanral speed data was differentiated between passenger and heavy vehicles, we lumped these together, as the INRIX data did not differentiate between vehicle classes.

The SANRAL data suffered from a significant fraction of missing values. As a prediction model requires complete time series for training purposes, we calculated weekly speed profiles for each SANRAL sensor. Missing values were replaced by the corresponding values for the same day of week and time of day obtained from the weekly profiles for the respective sensors. This ensured that the replacement data would not result in discontinuities when training the model.

Similar to the INRIX data, the SANRAL data also had to be formatted from the received format to the required format for use in the Graph WaveNet model. Samples of the received and formatted data can be seen in Figure 5 and Figure 6. Only 28 of the SANRAL static sensors coincided with the available INRIX data, as shown in Figure 7.

| STATION_ID | LANE_... | COUNT_DATE | COUNT_TIME | TOTAL_VEH | AVG_SPEED |
|---|---|---|---|---|---|
| 17445 | 1 | 01-JAN-22 | 01-JAN-22 12:0... | 13 | 86 |
| 22897 | 2 | 21-JAN-22 | 21-JAN-22 08:3... | 106 | 108 |
| 21006 | 3 | 16-APR-22 | 16-APR-22 09:0... | 46 | 115 |
| 23545 | 6 | 20-MAR-22 | 20-MAR-22 07:... | 1 | 122 |

**Figure 5: SANRAL Received Data Format**

| Datetime_Index | 17445 | 17446 | 17450 | 22820 |
|---|---|---|---|---|
| 2022-01-01T00:00:00.000 | 86.16666... | 86.45833... | 82.31578... | 81.05263... |
| 2022-01-01T00:15:00.000 | 91.2 | 92.04 | 81.28125 | 79.4375 |
| 2022-01-01T00:30:00.000 | 82 | 82.85714... | 83.14634... | 81.56097... |
| 2022-01-01T00:45:00.000 | 86.27586... | 87 | 84.18421... | 83 |
| 2022-01-01T01:00:00.000 | 88.76190... | 89.57142... | 91.07894... | 89.34210... |
| 2022-01-01T01:15:00.000 | 89.45161... | 89.70967... | 90.4 | 88.78 |
| 2022-01-01T01:30:00.000 | 92.76666... | 93.03333... | 85.08108... | 85.02702... |
| 2022-01-01T01:45:00.000 | 91.71428... | 92.02380... | 88.42307... | 85.84615... |

**Figure 6: SANRAL Formatted Data Format**

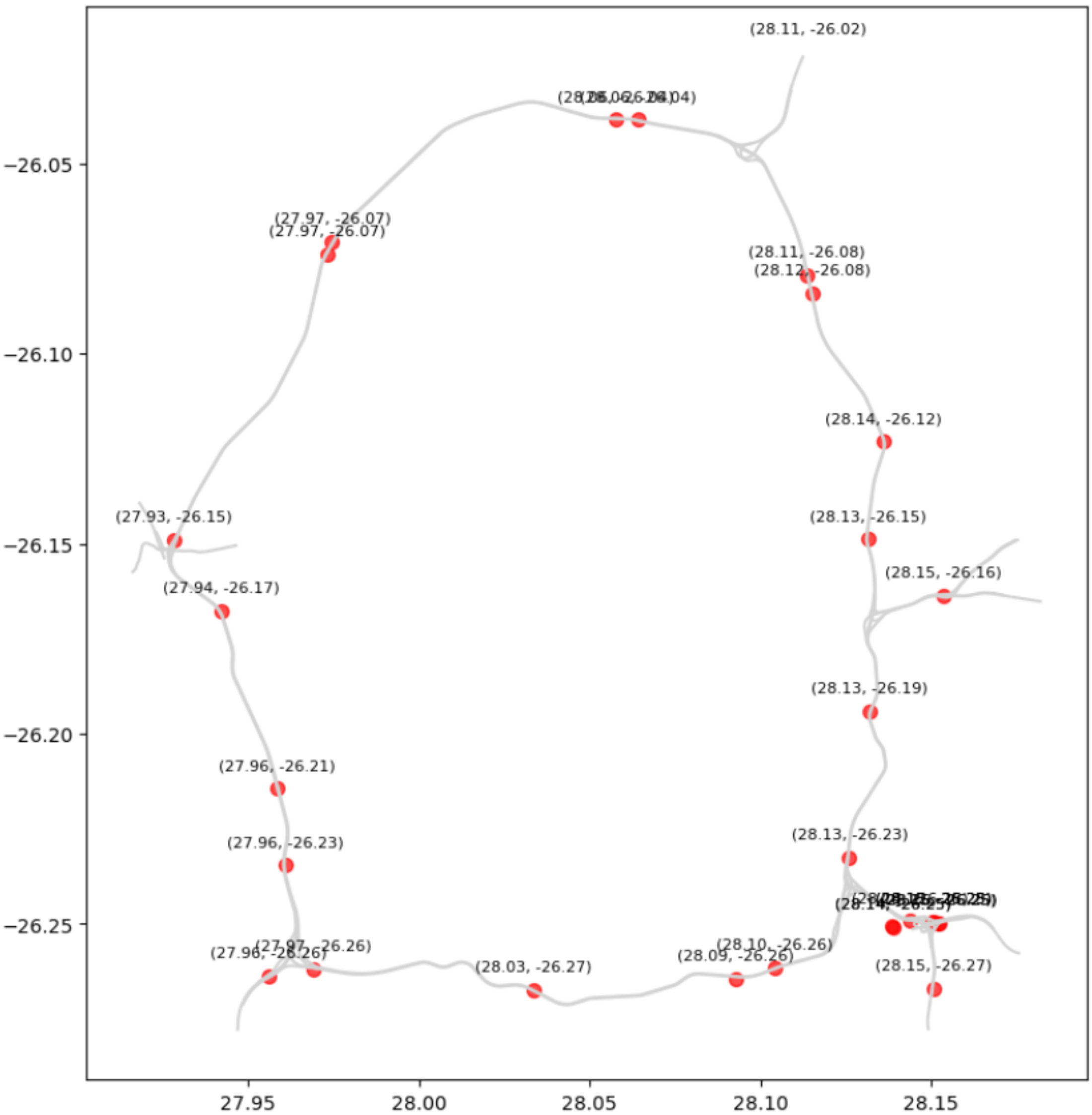



**Figure 7: SANRAL Johannesburg Static Sensor Stations**

## 4. TESTING

To be able to perform the computationally intensive calculations as required to train the Graph WaveNet model on the different datasets, access was obtained to the CHPC high-performance computing cluster. Firstly, training data was generated for the model by running a program that created a file in the required format from the SANRAL and INRIX datasets. 70% of the data is selected for training, 10% is used for validation testing and the remaining 10% is for testing purposes. Once all the files were created the model ran multiple training sequences multiple times by using the generated training data and performing sweeps to obtain the best hyperparameters for the sweeps over all the data. The hyperparameters used for the sweeps are displayed in Table 3.

**Table 3: Hyperparameter Sweep Variables**

| Parameter | Distribution | Min | Max |
|---|---|---|---|
| Nhid (number of hidden layers) | int_uniform | 10 | 110 |
| Batch_size | int_uniform | 10 | 100 |
| Learning_rate | uniform | 0.0001 | 0.01 |
| Dropout | uniform | 0.01 | 0.6 |
| Weight decay | uniform | 0.00001 | 0.001 |
| Epochs | uniform | - | 50 |

These sweeps were performed and visually showcased with an external tool named Weights & Biases (Weights & Biases, 2023), where multiple sweeps can be done with different parameters for specific experimentation purposes. The train loss is displayed in Figure 8 where multiple sweeps were completed to obtain the best epoch for hyperparameter selection. As seen in Figure 8 the sweeps produced improved results as the model was trained multiple times over 50 epochs.

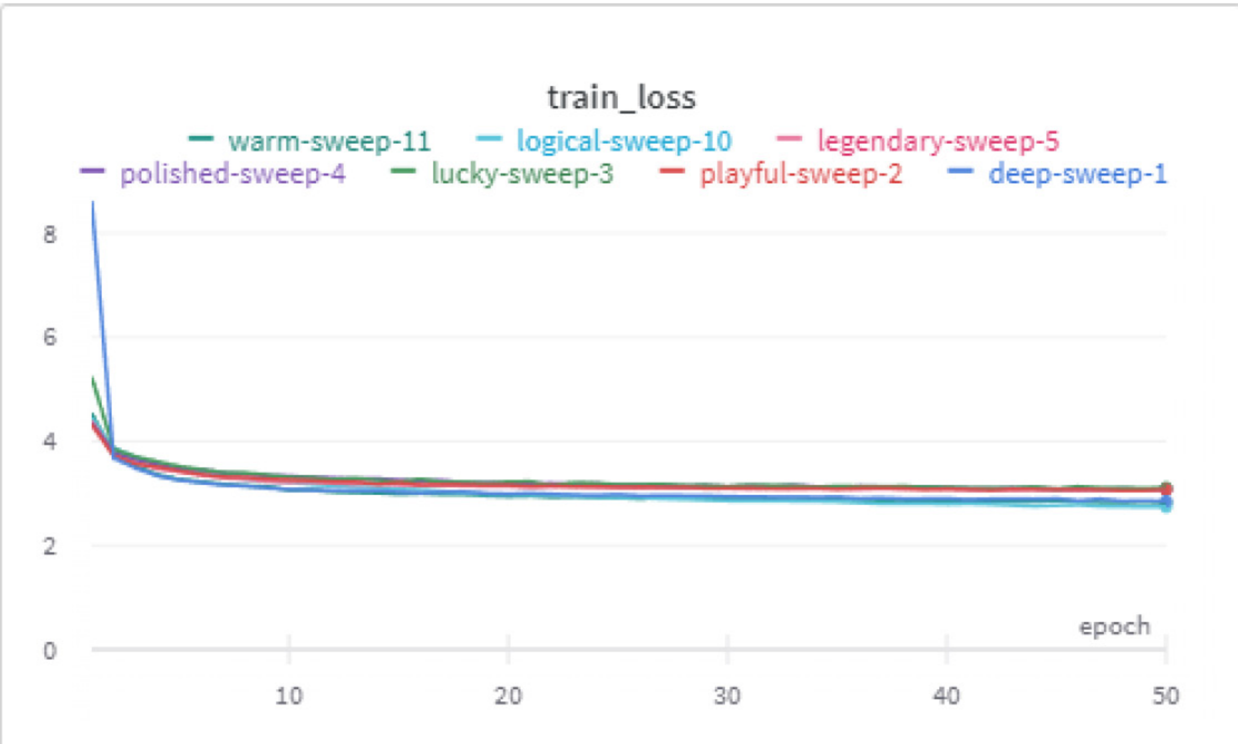



**Figure 8: SANRAL training loss for multiple sweeps**

After each epoch, the validation loss is calculated to determine the best performance of the model. The hyperparameters corresponding to the lowest validation loss value will be used for testing the model on the unseen data. A Bayes sweep was performed in all cases; this sweep method is based on a Gaussian process that utilizes the relationship between the model parameters and the model metric to optimize the probability for model improvement (Weights & Biases Docs, 2023). Other sweeping methods included a grid sweep that iterates over the parameter values and uses all possible combinations from these parameters. Another sweep method was the random sweep, that uses random parameter values for each iteration when doing multiple sweeps. As mentioned, the Bayes sweep was the best method to use for hyper parameter optimization.

After the Graph WaveNet model was trained on the datasets the best hyperparameters were determined, the model was tested with using code provided by (Zhan, 2023) to obtain the performance results displayed below. Model performance was based on the values of the MAE (mean absolute error), RMSE (root mean square error), and MAPE (mean absolute percentage error). The MAE is the mean of the absolute values of the residuals (the differences between the predicted values and the target values from the dataset). The RMSE value is the square root of the sum of the squared error values calculated from the residuals. The MAPE is the MAE expressed as percentage of the actual values are calculated. To illustrate the progressively improving model during training the MAPE is displayed in Figure 9.

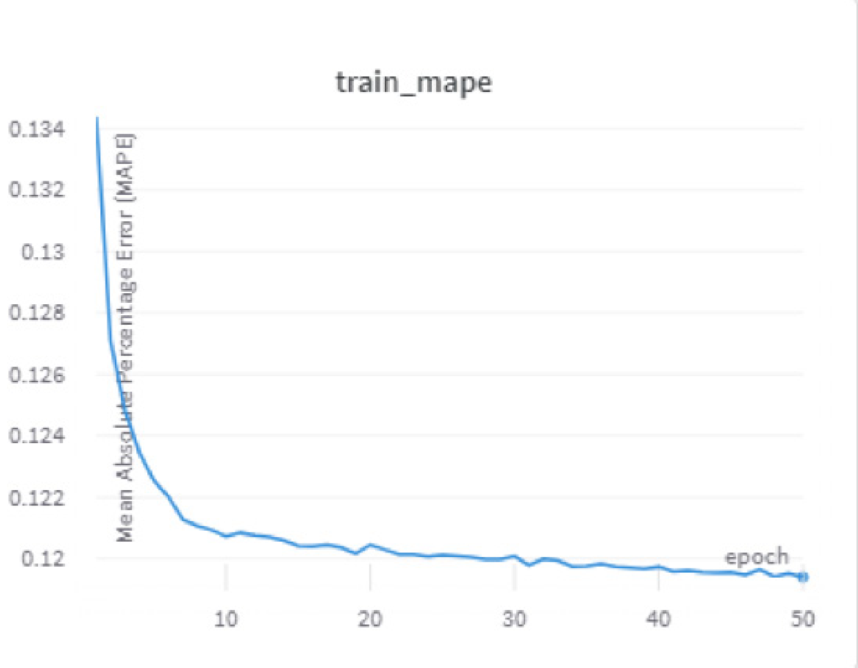



**Figure 9: SANRAL training MAPE for multiple sweeps**

A typical prediction result is displayed in Figure 10 below. The training data is displayed in green, the actual results for the testing period in blue and the predicted results for the testing period in red. It can be observed that the prediction closely follows the actual speed behaviour.

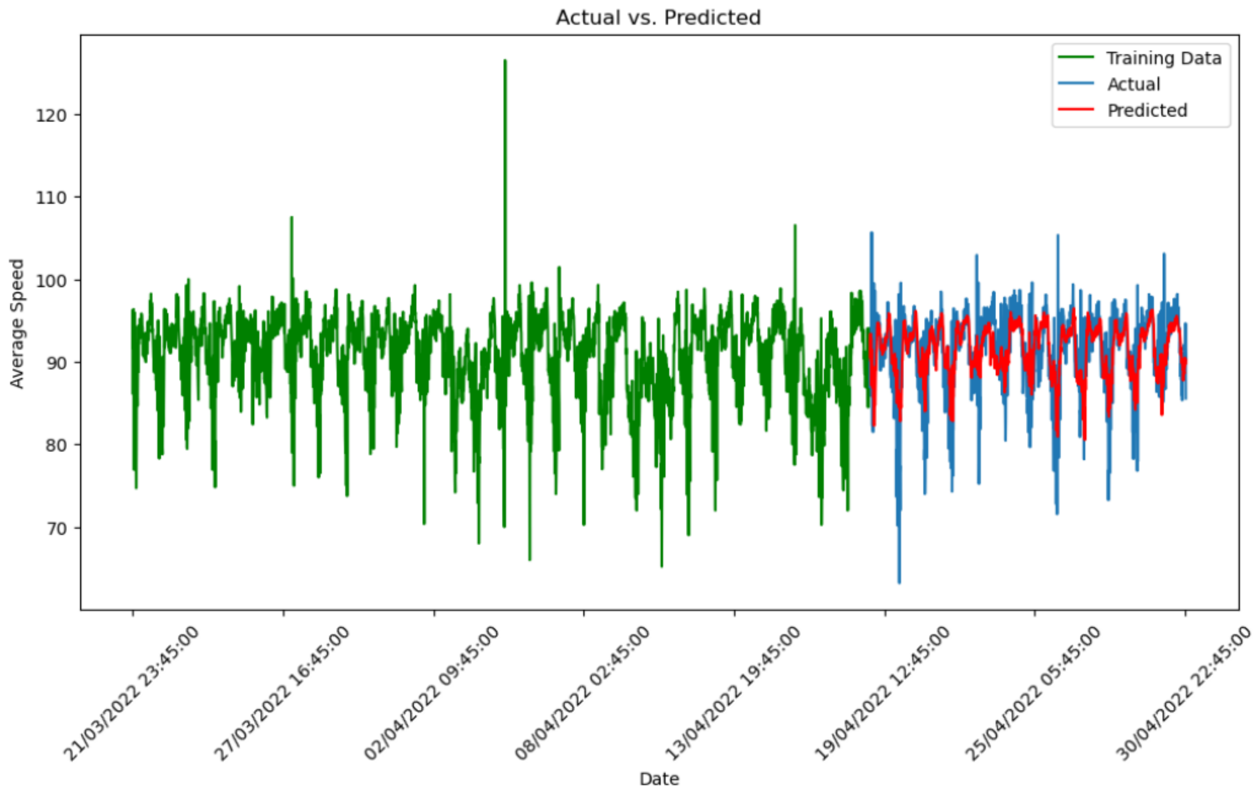



**Figure 10: Predicted speed for SANRAL station 17445**

## 5. RESULTS

The same performance metrics used for the METR-LA and PEMS-BAY datasets, namely, the Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE), were also used for the INRIX and SANRAL datasets.

### 5.1 INRIX Data Results

Similar to the prediction results for the METR-LA and PEMS-BAY datasets, we produced predictions over 15-minute, 30 minute and 60 minute intervals for the INRIX dataset, i.e. a maximum prediction horizon equal to 12 sampling periods. The results obtained by applying the Graph WaveNet model to the INRIX data can be seen in Table 4.

**Table 4: INRIX Graph WaveNet Results**

| Dataset | 15 min | | | 30 min | | | 60 min | | |
|---|---|---|---|---|---|---|---|---|---|
| | MAE | RMSE | MAPE | MAE | RMSE | MAPE | MAE | RMSE | MAPE |
| INRIX JHB | 7.4225 | 10.2261 | 10.29% | 7.6923 | 10.6453 | 11.02% | 7.917 | 11.0624 | 11.73% |

### 5.2 SANRAL Data Results

For the SANRAL data, the sampling period is 15 minutes, thus prediction over a horizon of 12 sampling periods, as for the INRIX data, results in a 3-hour prediction horizon. The results can be seen in Table 5, where 45-minute, 90-minute and 180-minute predictions were made by the Graph WaveNet model.

**Table 5: SANRAL Graph WaveNet Results**

| Dataset | 45 min | | | 90 min | | | 180 min | | |
|---|---|---|---|---|---|---|---|---|---|
| | MAE | RMSE | MAPE | MAE | RMSE | MAPE | MAE | RMSE | MAPE |
| SANRAL JHB | 2.7812 | 5.1486 | 3.92% | 3.0596 | 5.8622 | 4.46% | 3.3437 | 6.3891 | 4.94% |

### 5.3 Comparisons

This section compares the Graph WaveNet results for the South African data sets with the results obtained from the METR-LA and PEMS-BAY datasets. The comparisons include the MAE, RMSE, and MAPE values from all the datasets and can be seen in Table 6. In all cases the prediction horizons are defined over 3, 6 and 12 sampling periods.

It is observed that the prediction accuracy for the SANRAL data is similar to that obtained for the PEMS-BAY data, which is also based on static sensors. The predictions obtained for the INRIX floating vehicle data are somewhat less accurate. This could be expected as, in contrast to the other data sets, the INRIX data is based on only a fraction of the total vehicle population traveling on the respective roads during the period of data collection. It is however still close to the accuracy obtained for the METR-LA data and can thus be regarded as good enough for practical use.

**Table 6: Comparison of INRIX, SANRAL, METR-LA and PEMS-BAY Graph WaveNet Results**

| Dataset | 3 Sampling Periods | | | 6 Sampling Periods | | | 12 Sampling Periods | | |
|---|---|---|---|---|---|---|---|---|---|
| | MAE | RMSE | MAPE | MAE | RMSE | MAPE | MAE | RMSE | MAPE |
| INRIX | 7.4225 | 10.2261 | 10.29% | 7.6923 | 10.6453 | 11.02% | 7.917 | 11.0624 | 11.73% |
| SANRAL | 2.7812 | 5.1486 | 3.92% | 3.0596 | 5.8622 | 4.46% | 3.3437 | 6.3891 | 4.94% |
| METR-LA | 2.69 | 5.15 | 6.90% | 3.07 | 6.22 | 8.37% | 3.53 | 7.37 | 10.01% |
| PEMS-BAY | 1.30 | 2.74 | 2.73% | 1.63 | 3.70 | 3.67% | 1.95 | 4.52 | 4.63% |

## 6. CONCLUSIONS AND FUTURE WORK

The Graph WaveNet technique performed similarly on the static SANRAL dataset compared to the benchmark PEMS-BAY and METR-LA datasets but did not fare as well on the INRIX floating vehicle dataset. This may be the result of data gathering methods influencing the results. The static sensors record data for all vehicles that pass over the respective roads, while the floating vehicle data is extracted from only a fraction of all the vehicles. It would be interesting to study the effect of the percentage of vehicles used to extract floating vehicle data on the accuracy of traffic speed predictions.

While the floating vehicle data results are worse than the static sensor data results, it offers the benefit that it can be gathered on a larger scale compared to using static sensors at no extra cost in terms of installed infrastructure. Thus, it could prove to be useful as a supplement to static sensor data, enabling traffic speed prediction for roads where no static sensors are installed, and resulting in significant cost savings. Extending traffic speed prediction to all major roads will result in improved traffic management strategies for the overall road network, leading to less congestion and an improved road user experience.

Since this research was performed, some new techniques have emerged in the traffic speed prediction space. Future work will involve testing such techniques to determine if they perform better on floating traffic data than the Graph WaveNet technique. It could also be useful to compare the results for the INRIX data set to results obtained for other floating vehicle datasets using multiple prediction techniques to obtain further insights.

## 7. REFERENCES

Kumar, N & Raubal, M. 2021. Applications of deep learning in congestion detection, prediction and alleviation: A survey, *Transportation Research Part C: Emerging Technologies,* 133:1-27. Doi: 10.1016/j.trc.2021.103432.

Li, Y et al. 2017. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting, CoRR, abs/1707.01926. Doi: 10.48550/arXiv.1707.01926.

Metr-la benchmark (traffic prediction). The latest in Machine Learning. Available at: https://paperswithcode.com/sota/traffic-prediction-on-metr-la. Accessed 23 January 2024.

Mena-Oreja, J & Gozalvez, J. 2020. A comprehensive evaluation of deep learning-based techniques for traffic prediction, IEEE Access, 8:91188-91212.
Doi: 10.1109/access.2020.2994415.

Mallick, T et al. 2021. Transfer learning with graph neural networks for short-term highway traffic forecasting, 2020. 25th International Conference on Pattern Recognition (ICPR) [Preprint]. Doi: 10.1109/icpr48806.2021.9413270.

PEMS-bay benchmark (Traffic prediction). The latest in Machine Learning. Available at: https://paperswithcode.com/sota/traffic-prediction-on-pems-bay. Accessed 23 January 2024.

Polson, NG & Sokolov, VO. 2017. Deep learning for short-term traffic flow prediction, *Transportation Research Part C: Emerging Technologies*, 79:1-17.
Doi: 10.1016/j.trc.2017.02.024.

Tedjopurnomo, DAZB. 2022. A Survey on Modern Deep Neural Networkfor Traffic Prediction: Trends, Methodsand Challenges. *IEEE Transactions on Knowledge and Data Engineering,* 34(4):1544-1561.

Wang, JQG. 2016. Traffic Speed Prediction and Congestion Source Exploration: A Deep Learning Method. *2016 IEEE 16th International Conference on Data Mining (ICDM).* Barcelona.

Wang, Y et al. 2019. Enhancing transportation systems via Deep Learning: A Survey, *Transportation Research Part C: Emerging Technologies*, 99:144-163.
Doi: 10.1016/j.trc.2018.12.004.

*Weights & Biases*. 2023. 23 July. Available at: https://wandb.ai/site.

*Weights & Biases Docs*. 2023. 10 October. Available at:
https://docs.wandb.ai/guides/sweeps/define-sweep-configuration#method.

Wu, Z et al. 2019. Graph WaveNet for deep spatial-temporal graph modeling, Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence [Preprint]. Doi: 10.24963/ijcai.2019/264.

Zhan, NN. 2023. 15 August. *Graph-WaveNet*. Available at:
https://github.com/nnzhan/Graph-WaveNet.