

Time Series Predictions

Igor Ilic

Oct 25, 2019

Data Science Lab: Ryerson University

Overview

- Introduce Problem
- Review Literature
- Dataset Fetching and Cleaning
- Model Implementation
- Results
- Next Steps

(1/6) Introduce Problem

- Contacted by LAPD regarding highway traffic around LA Dodgers stadium
- 6 months worth of 5-minute interval data
- How many traffic control officers are needed on a given day?

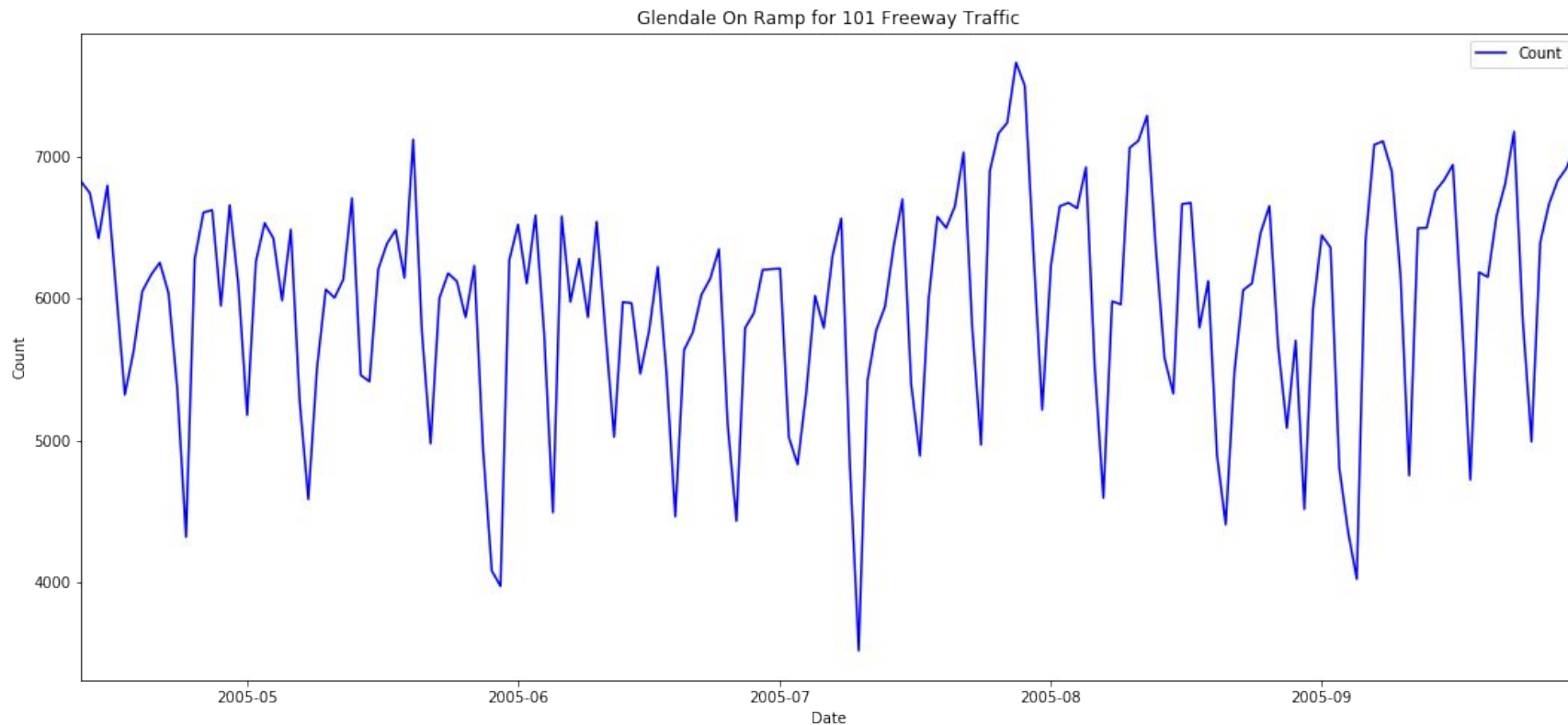
(2/6) Review Literature

- [1] O. Ozyegen, S. Mohammadjafari, E. Kavurmacioglu, J. Maidens and A. Basar. “Experimental Results on the Impact of Memory in Neural Networks for Spectrum Prediction in Land Mobile Radio Bands”.
- [2] S. Mohammadjafari, E. Kavurmacioglu, J. Maidens and A. Bener, "Neural Network Based Spectrum Prediction in Land Mobile Radio Bands for IoT deployments," *2019 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, Arlington, VA, USA, 2019, pp. 31-36.
- [3] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
- [4] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov. “Dropout: a simple way to prevent neural networks from overfitting,” *The journal of machine learning research*, 15(1), 1929-1958, 2014.
- [5] D. P. Kingma, and J. Ba. "Adam: A method for stochastic optimization." *arXiv preprint arXiv:1412.6980*, 2014.
- [6] K. W. Hipel, A. I. McLeod and W. C. Lennox. “Advances in Box-Jenkins modeling: 1. Model construction,” *Water Resources Research*, 13(3), 567-575, 1977.
- [7] J. Chung, C. Gulcehre, K. Cho and Y. Bengio. “Empirical evaluation of gated recurrent neural networks on sequence modeling,” *arXiv preprint arXiv:1412.3555*, 2014.
- [8] Freeway Performance Measurement System (PeMS), “<http://pems.eecs.berkeley.edu/>.”

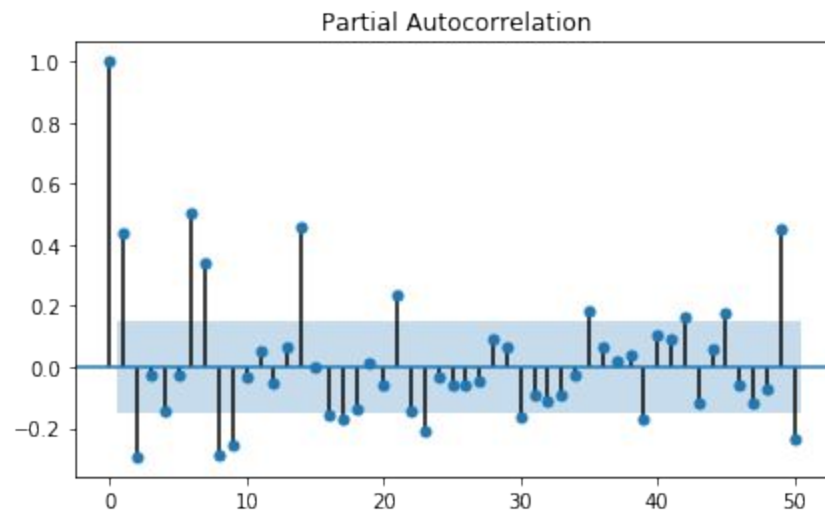
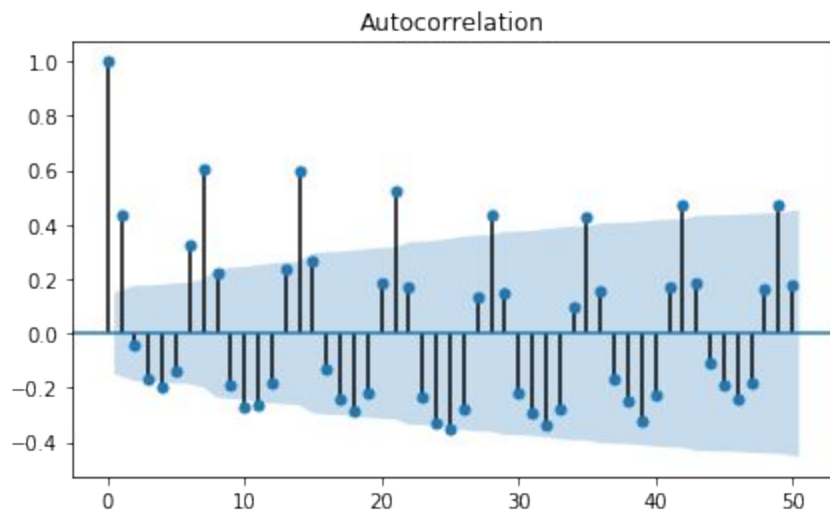
(3/6) Dataset Fetching and Cleaning

- Dodgers Loop Sensor Data Set
- Initial dataset was used to predict minute-by-minute impact on a ramp due to LA Dodgers Game
- High variance in minute-by-minute. We want to predict daily
- Filled missing datapoints with average 5-minute interval
- 5-minute intervals concatenated into daily intervals

(3/6) Dataset Fetching and Cleaning Cont.



(3/6) Dataset Fetching and Cleaning Cont.



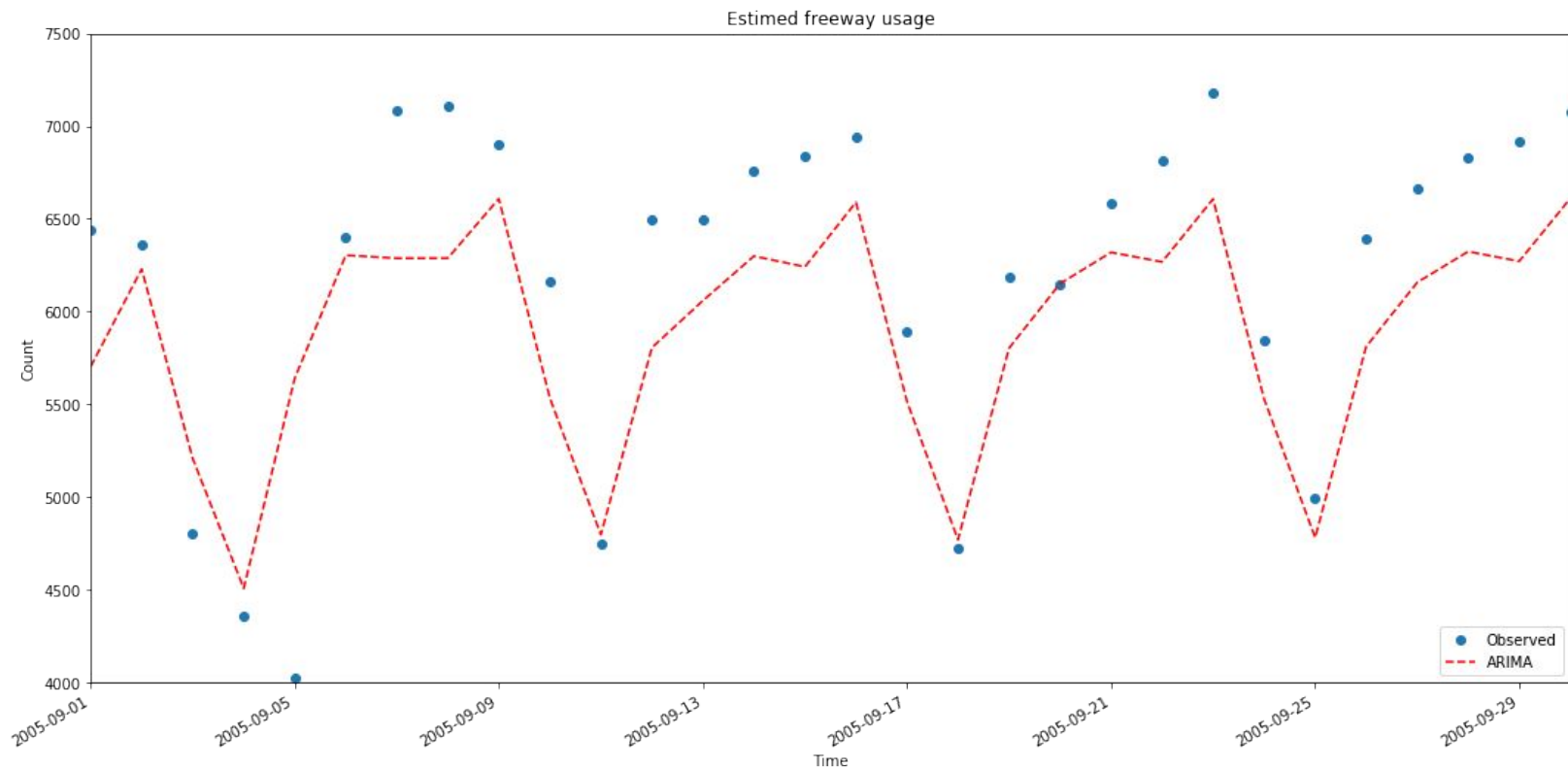
(4/6) Model Implementation

- SARIMA(3,0,1)(1,0,3)₇ - fit using Box Jenkins method.
- Long Short-Term Memory (LSTM)
- LSTM + Variants (Optimizer / Dense Network / L2 Regularization)
- Time Delayed Neural Network (TDNN)

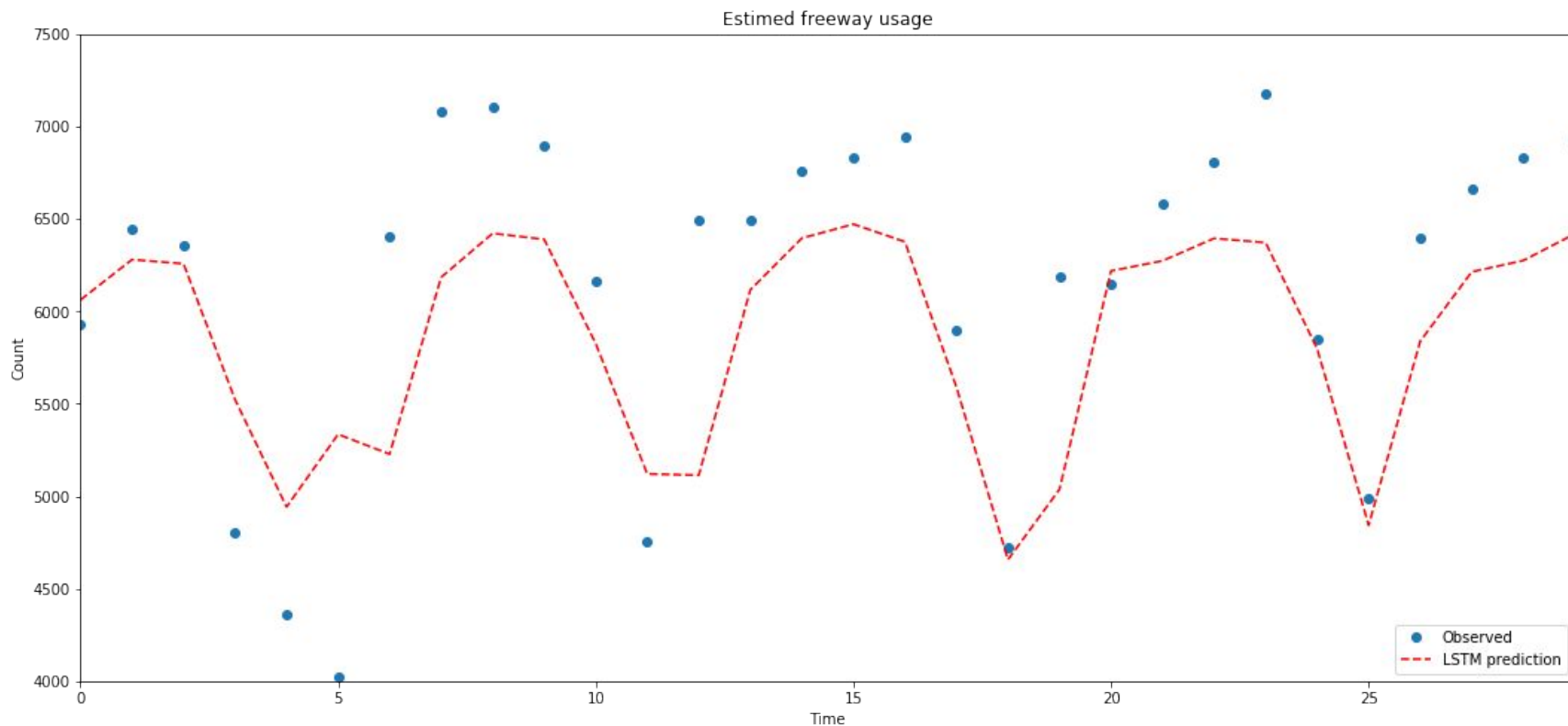
(5/6) Results

| Model | Test Set RMSE (cars/day over 30 day period) |
|-------------------|---|
| SARIMA | 551.42 |
| LSTM | 628.13 |
| LSTM with Dense | 475.46 |
| LSTM with Dropout | 547.76 |
| TDNN | 599.07 |

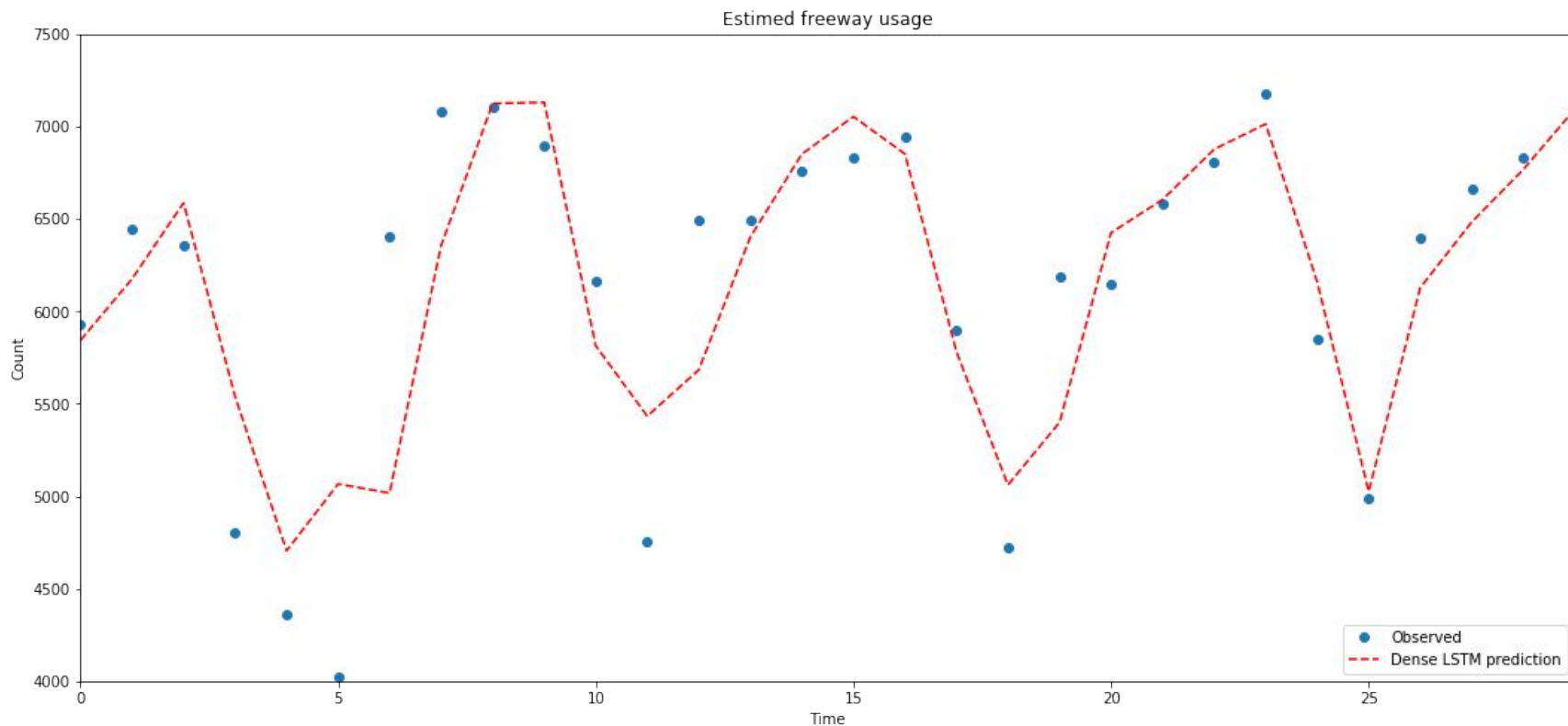
(5/6) Results Cont.



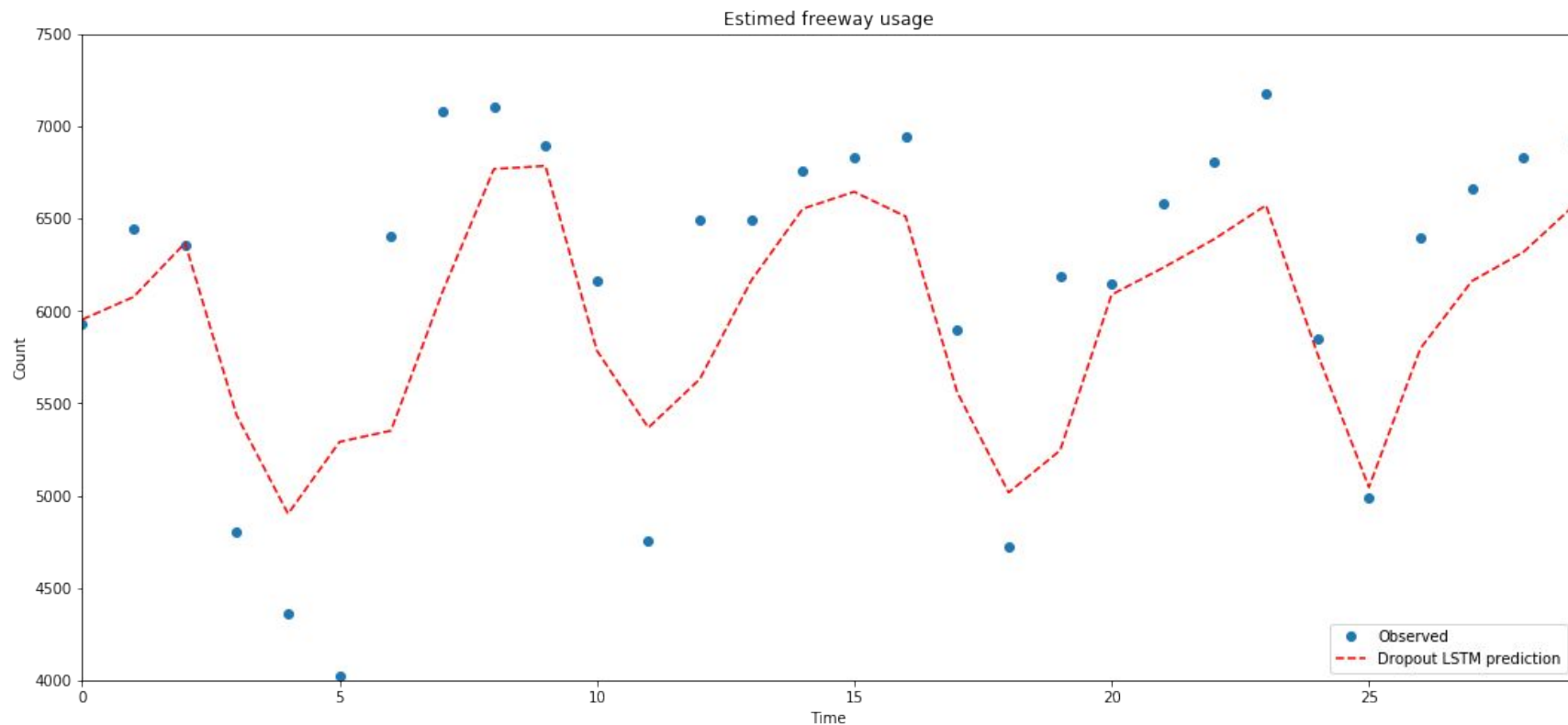
(5/6) Results Cont.



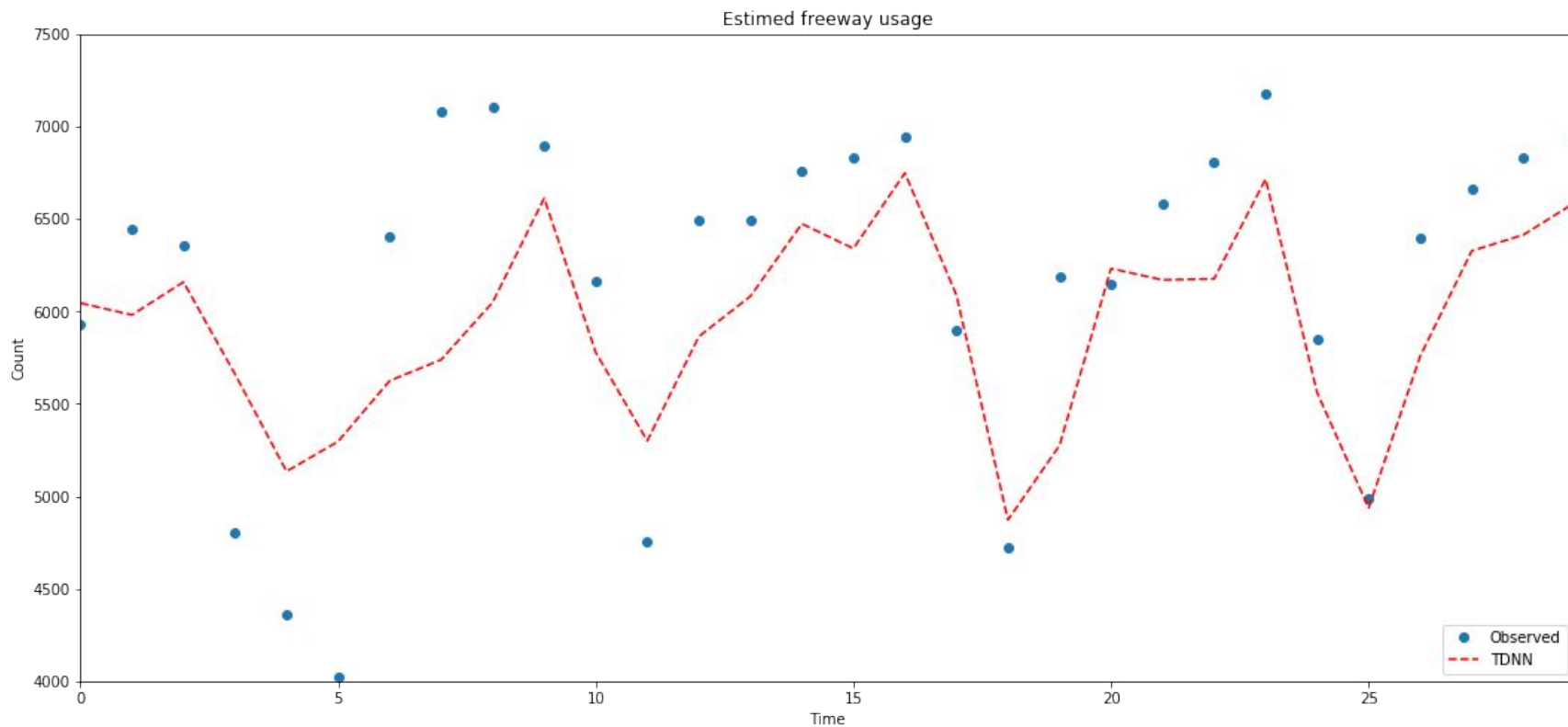
(5/6) Results Cont.



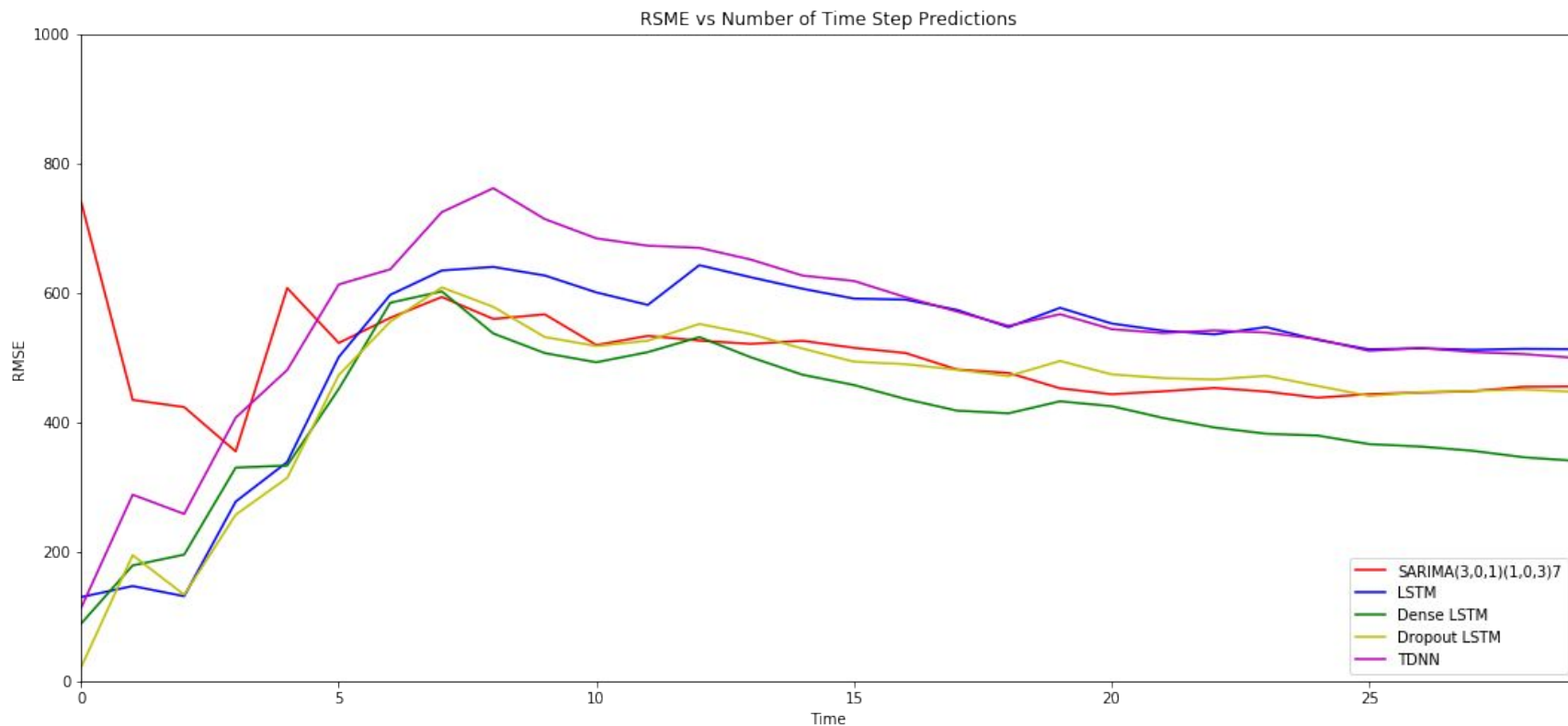
(5/6) Results Cont.



(5/6) Results Cont.



(5/6) Results Cont.



(6/6) Further Studying

- Larger Data Set
- Gated Recurrent Neural Networks

Thank you!