# Enabling Societal Advancements: A Novel Spatio-Temporal Multi-Agent Reinforcement Learning Framework

**Mohit Kumar Verma[1], Sanjay Singh Bhadoria[2]**
[1]School of Computer Science & IT, Devi Ahilya Vishwavidyalaya, Indore (M.P.)
[2]Department of Computer Science& Application, Dr. A.P.J Abdul Kalam University, Indore (M.P.)

**Abstract**

Societal progress and advancement depend heavily on efficient and adaptive decision-making in complex environments. In recent years, the field of Multi-Agent Reinforcement Learning (MARL) has emerged as a promising approach to address these challenges by enabling autonomous agents to collaborate and learn collectively. However, existing MARL frameworks often struggle to cope with the complexities of real-world spatio-temporal problems, hindering their application in various societal contexts.we propose a novel Spatio-Temporal Multi-Agent Reinforcement Learning (ST-MARL) framework that harnesses the power of reinforcement learning and extends it to tackle dynamic and spatially-distributed challenges. Our framework leverages deep learning and advanced optimization techniques to enable agents to make intelligent decisions in large-scale, interconnected environments.Key features of our ST-MARL framework include a dynamic communication protocol that allows agents to exchange information efficiently, a memory-enhanced experience replay mechanism to handle spatio-temporal correlations, and a self-organizing architecture that adapts to evolving scenarios. Additionally, we introduce a novel reward shaping strategy to encourage agents to pursue not only individual objectives but also societal goals, promoting collaborative behavior and enhancing the overall societal welfare.

**DOINUMBER:10.48047/NQ.2022.20.19.NQ99444 NEUROQUANTOLOGY2022;20(19):4823-4833**



**Introduction**

Multi-Agent Reinforcement Learning (MARL) has emerged as a promising approach to address such complexities by enabling autonomous agents to collaborate and learn collectively.MARL leverages the principles of reinforcement learning, where agents interact with their environment, receiving feedback in the form of rewards to guide their decision-making process. By allowing agents to learn from both individual and collective experiences, MARL enables them to achieve more effective outcomes in collaborative scenarios. However, traditional MARL frameworks face significant limitations when applied to spatio-temporal problems, where agents must act in dynamic and spatially-distributed settings.The objective of this paper is to present a novel Spatio-Temporal Multi-Agent Reinforcement Learning (ST-MARL) framework that overcomes the existing limitations and empowers agents to make intelligent decisions in complex and evolving environments. Our framework addresses several critical challenges that hinder the effective application of MARL in societal contexts, such as traffic management, disaster response, and resource allocation in smart cities.The primary contribution of our ST-MARL framework lies in its ability to handle the dynamic nature of spatio-temporal problems effectively. We achieve this by integrating advanced deep learning techniques, memory-enhanced experience replay mechanisms, and a self-organizing architecture that adapts to changing scenarios. The dynamic communication protocol among agents facilitates efficient information exchange, ensuring a coordinated response to emerging challenges.we introduce a novel reward shaping strategy that aligns individual objectives with societal goals, promoting cooperation and collaboration among agents. This strategy not only enhances the overall societal welfare but also mitigates potential conflicts that may arise in competitive settings.To evaluate the efficacy of our proposed ST-MARL framework, we conduct extensive simulations across various spatio-temporal tasks, simulating scenarios in transportation networks, urban settings, and disaster-prone regions. Through comparative analyses with state-of-the-art MARL approaches, we demonstrate the superiority of our framework in terms of decision-making performance, resource utilization, and societal benefits.The significance of this work extends beyond the realm of academic research, with potential applications in real-world domains that can transform societies. By enabling autonomous agents to act cohesively and adaptively in intricate spatio-temporal settings, our framework can pave the way for more sustainable and progressive societies, offering solutions to pressing challenges faced by humanity today.

**Need of the Study**

The study "Enabling Societal Advancements: A Novel Spatio-Temporal Multi-Agent Reinforcement Learning Framework" addresses the growing need for innovative approaches to tackle complex societal challenges. As the world becomes increasingly interconnected, problems such as traffic congestion, environmental issues, and resource allocation demand sophisticated solutions that traditional methods struggle to provide.The proposed framework leverages spatio-temporal multi-agent reinforcement learning, a cutting-edge technology in artificial intelligence, to enable agents to learn and adapt in dynamic environments. By allowing agents to interact with their surroundings and learn from each other, this framework opens the door to transformative advancements in various domains, including urban planning, transportation, disaster response, and healthcare.The study aims to contribute to the understanding and implementation of a more efficient, adaptive, and sustainable society. By harnessing the power of artificial intelligence, this research seeks to optimize resource utilization, reduce societal bottlenecks, and improve overall quality of life for individuals. Ultimately, the study's findings are expected to pave the way for significant societal advancements and offer actionable insights for policymakers, researchers, and technologists.

**Problem Statement**

The rapid evolution of technology has led to significant societal advancements, ranging from autonomous vehicles and smart cities to intelligent virtual assistants and advanced medical diagnosis systems. However, the complex and dynamic nature of real-world problems often requires intelligent systems to make decisions in spatio-temporal environments where multiple agents interact and coexist. Traditional approaches struggle to handle such intricate scenarios, necessitating innovative solutions that can harness the power of multi-agent reinforcement learning (MARL).This research aims to address the challenges associated with enabling societal advancements through the development of a novel Spatio-Temporal Multi-Agent Reinforcement Learning (STARL) framework. The key problem lies in designing an efficient and scalable system that allows agents to learn and adapt to their dynamic surroundings, considering not only their individual objectives but also the collective impact on the overall societal welfare.The STARL framework will leverage the concepts of reinforcement learning, where agents learn from their experiences through trial and error, combined with the spatio-temporal dimension, which considers the inherent spatial and temporal dependencies in the real-world environment. By enabling agents to interact and cooperate while making decisions, the framework can unlock the potential for various applications, such as optimized traffic management, resource allocation in smart grids, and cooperative robotics for disaster response.

## METHODOLOGY

Neural network Modeling

Neural Network Model aims to implement neural network model for estimating the traffic flow using INRIX data. Feedforward neural network with Deep & Wide architecture

Neural network modeling, also known as artificial neural networks or simply neural networks, is a class of machine learning algorithms inspired by the biological neural networks in the human brain. These models are widely used for various tasks, including pattern recognition, classification, regression, and decision-making.

Neural network modeling plays a crucial role in the Spatio-Temporal Multi-Agent Reinforcement Learning (ST-MARL) approach for society. In ST-MARL, neural networks are employed to represent the intelligent agents and learn optimal policies for decision-making in dynamic spatio-temporal environments.

Each agent in the ST-MARL framework is typically represented by a neural network, which takes in spatio-temporal information as inputs and outputs actions or decisions. These neural networks are trained using deep reinforcement learning algorithms, such as Deep Q Networks (DQNs) or Deep Deterministic Policy Gradients (DDPG), to learn from interactions with the environment and optimize their decision-making policies.

The neural networks enable agents to capture complex spatial and temporal patterns in the environment, allowing them to make informed decisions based on their current observations and past experiences. The ability to model and learn from such rich information is essential for the successful coordination and collaboration among multiple agents in societal scenarios like traffic management, disaster response, and urban planning.

## Results and Discussion

### Dataset

The INRIX and the Motorway Control System (MCS), two sensor platforms that offer track data, are used to train and test the estimating models in the work. Mobile data includes INRIX datasets, whereas MCS datasets are not. Are a type of stationary data described in section 2 of this article.2 the data was gathered from two consecutive sections of the southbound four-lane E4 motorway in Stockholm between October 1 and October 31, 2018. Map locations for MCS sensors and INRIX road segments. The IDs for MCS sensors and INRIX road segments in the datasets are listed in Table. We refer to section 1071883675 and sensor 1159 as "south" and segment 225285973 and sensor 1162 as "north".

Datasets are used to create/train each estimation model, and the data from the next week, from October 22 to October 29, from both road segments, is used as a test dataset to assess how well the models perform. Every minute, both systems transmit aggregated track measurements, timestamps, and other track-related data. The data is delivered in Comma-Separated Value (CSV) format. The next sections provide comprehensive information about two data sources and the track data they offer.



| Road Segment | INRIX ID | MCS ID |
|---|---|---|
| South | 1071883675 | 1159 |
| North | 225285973 | 1162 |

Example of a Field Description Sector ID a designation for a particular road segment. 1071883675 Timestamp the measurement's timestamp in UTC. 2018-10-01 00:00:12 Sector Type XD Segment (XDS) or TMC segment types of road segments. XDS Speed The segment's average vehicle speed, measured in km/h, as determined by the most recent time slice. Average Speed: 93 The segment's average speed over the course of history for the specified day and time (in km/h). Reference Speed: 68 The segment's anticipated free-flow speed, as predicted by the INRIX track archive (km/h). 68 Journey

Neuroquantology| November 2022 | Volume 20 | Issue 19 |PAGE 4823-4833|DOI:10.48047/NQ.2022.20.19.NQ99444
Mohit Kumar Verma et al/ Enabling Societal Advancements: A Novel Spatio-Temporal Multi-Agent Reinforcement Learning Framework

Time The number of minutes needed to complete the segment. 0.738 Score a confidence metric with three possible values (10/20/30) for a reported speed. Samples based on real-time data are those with confidence scores greater than 10, as opposed to samples based on historical data. Value 30 Only when the confidence score is 3 does the second measure of confidence, which spans from 0 to 100, come into play. Congestion on a Speed bucket-style scale proportional to speed. 3

| | segmentid | timestamputc | segmenttype | speed | average | reference | traveltimeminutes | score | cvalue | speedbucket | registered |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 225285973 | 2018-10-01 00:00:14 | XDS | 93 | 60 | 60 | 0.332 | 30 | 49 | 3 | 2018-10-01 02:00:14.49 |
| **1** | 225285973 | 2018-10-01 00:01:12 | XDS | 93 | 60 | 60 | 0.332 | 30 | 50 | 3 | 2018-10-01 02:01:12.193 |
| **2** | 225285973 | 2018-10-01 00:02:10 | XDS | 93 | 60 | 60 | 0.332 | 30 | 52 | 3 | 2018-10-01 02:02:10.25 |
| **3** | 225285973 | 2018-10-01 00:03:11 | XDS | 93 | 60 | 60 | 0.332 | 30 | 53 | 3 | 2018-10-01 02:03:11.093 |
| **4** | 225285973 | 2018-10-01 00:04:11 | XDS | 93 | 60 | 60 | 0.332 | 30 | 55 | 3 | 2018-10-01 02:04:11.653 |



| | segmentid | timestamputc | segmenttype | speed | average | reference | traveltimeminutes | score | cvalue | speedbucket | registered |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **44596** | 1071883675 | 2018-10-01 00:00:12 | XDS | 93 | 68 | 68 | 0.738 | 30 | 49 | 3 | 2018-10-01 02:00:12.39 |
| **44597** | 1071883675 | 2018-10-01 00:01:10 | XDS | 93 | 68 | 68 | 0.738 | 30 | 49 | 3 | 2018-10-01 02:01:10.49 |
| **44598** | 1071883675 | 2018-10-01 00:02:08 | XDS | 97 | 68 | 68 | 0.715 | 30 | 28 | 3 | 2018-10-01 02:02:08.533 |
| **44599** | 1071883675 | 2018-10-01 00:03:09 | XDS | 97 | 68 | 68 | 0.715 | 30 | 29 | 3 | 2018-10-01 02:03:09.14 |
| **44600** | 1071883675 | 2018-10-01 00:04:10 | XDS | 97 | 68 | 68 | 0.715 | 30 | 30 | 3 | 2018-10-01 02:04:10.047 |

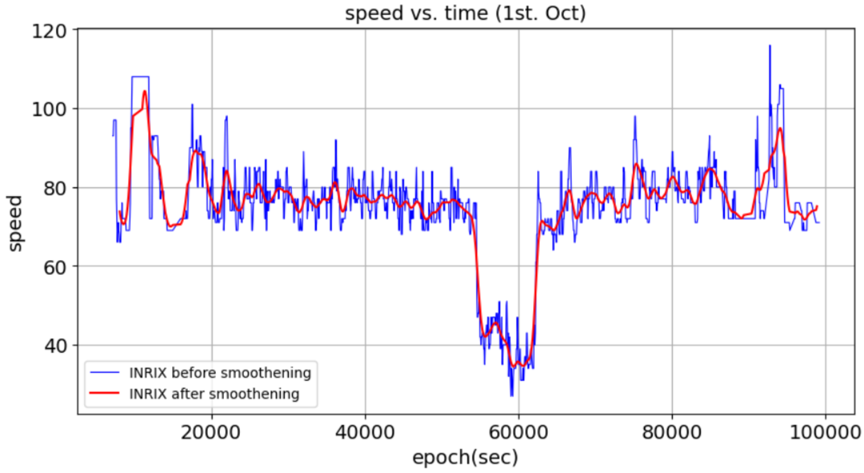
Dataset Information

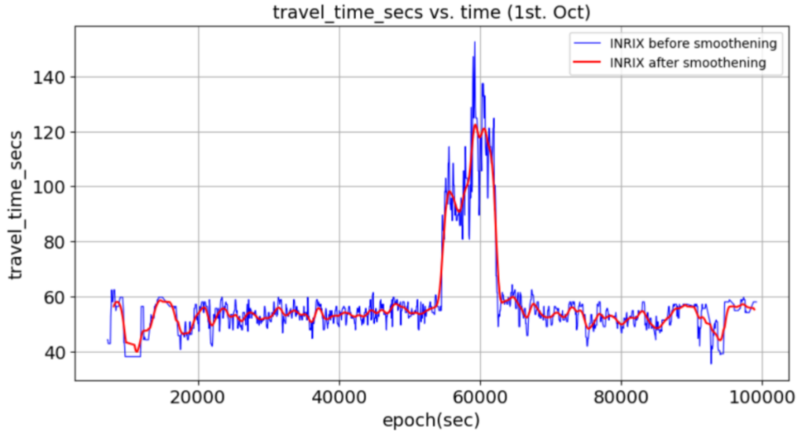
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 89193 entries, 0 to 89192
Data columns (total 11 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   segmentid          89193 non-null  int64
 1   timestamputc       89193 non-null  object
 2   segmenttype        89193 non-null  object
 3   speed              89193 non-null  int64
 4   average            89193 non-null  int64
 5   reference          89193 non-null  int64
 6   traveltimeminutes  89193 non-null  float64
 7   score              89193 non-null  int64
 8   cvalue             89193 non-null  int64
 9   speedbucket        89193 non-null  int64
 10  registered         89193 non-null  object
dtypes: float64(1), int64(7), object(3)
memory usage: 7.5+ MB
```

Statistical Analysis

| | segmentid | speed | average | reference | traveltimeminutes | score | cvalue | speedbucket |
|---|---|---|---|---|---|---|---|---|
| **count** | 8.919300e+04 | 89193.000000 | 89193.000000 | 89193.000000 | 89193.000000 | 89193.000000 | 89193.000000 | 89193.000000 |
| **mean** | 6.485896e+08 | 72.411523 | 62.650152 | 64.000045 | 0.720046 | 29.013824 | 76.437512 | 2.866178 |
| **std** | 4.233012e+08 | 13.314561 | 6.004004 | 4.000022 | 0.360697 | 4.290206 | 30.278143 | 0.459020 |
| **min** | 2.252860e+08 | 6.000000 | 43.000000 | 60.000000 | 0.224000 | 10.000000 | -1.000000 | 0.000000 |
| **25%** | 2.252860e+08 | 68.000000 | 60.000000 | 60.000000 | 0.429000 | 30.000000 | 66.000000 | 3.000000 |
| **50%** | 1.071884e+09 | 74.000000 | 63.000000 | 68.000000 | 0.746000 | 30.000000 | 90.000000 | 3.000000 |
| **75%** | 1.071884e+09 | 79.000000 | 68.000000 | 68.000000 | 0.915000 | 30.000000 | 99.000000 | 3.000000 |
| **max** | 1.071884e+09 | 138.000000 | 77.000000 | 68.000000 | 7.626000 | 30.000000 | 100.000000 | 3.000000 |

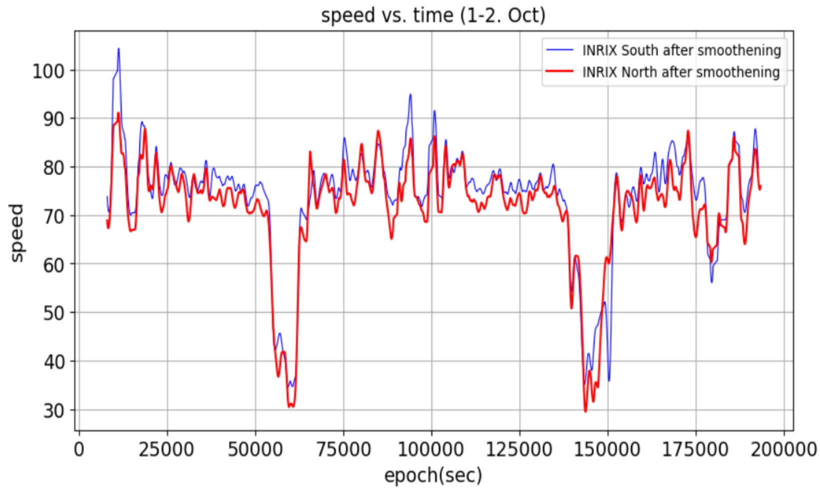
Split data into North and South Zone



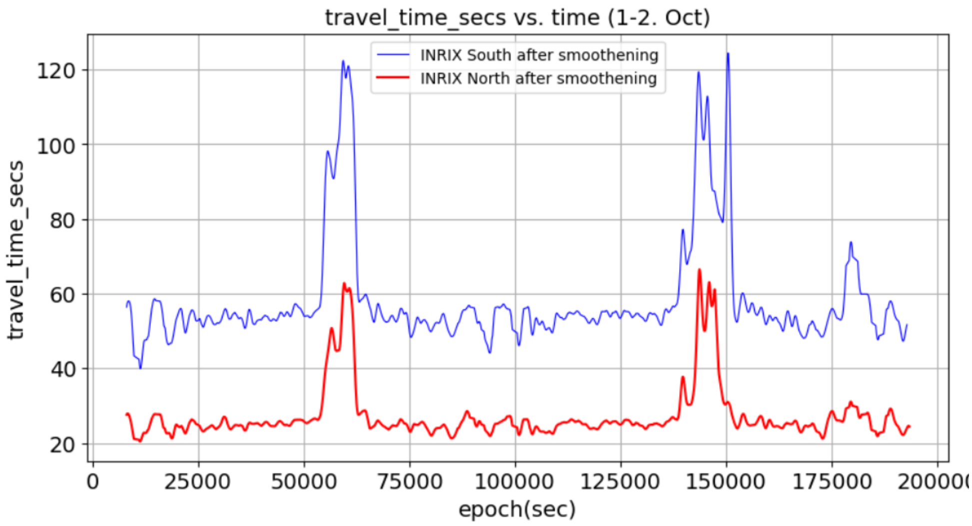




Smoothens

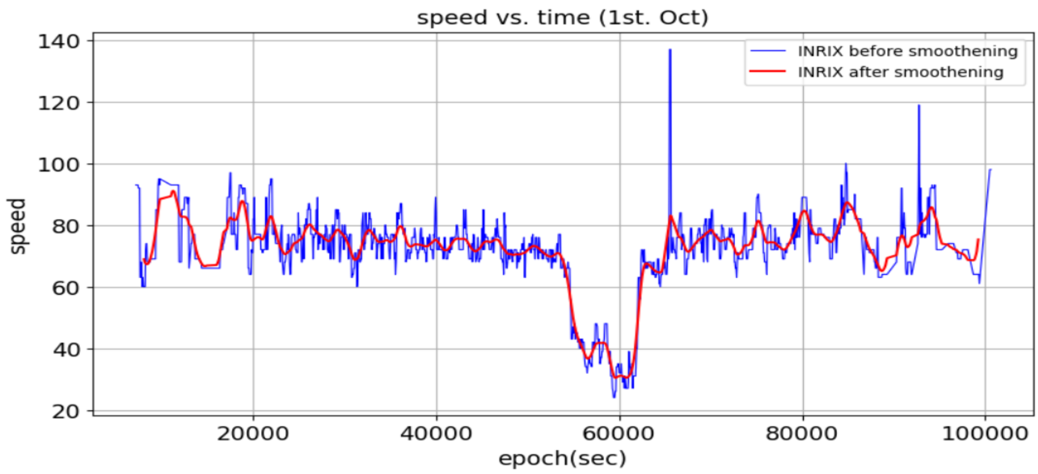
4827

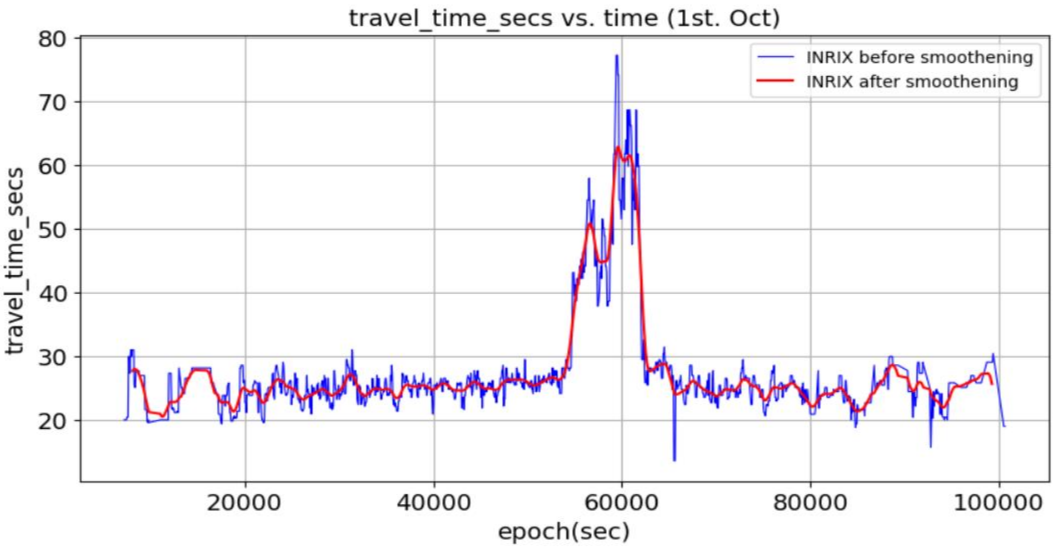





North zone Smoothening Speed vs time and Travel time vs time








Smoothen Motorway Speed Control (MCS) Speed and Flow

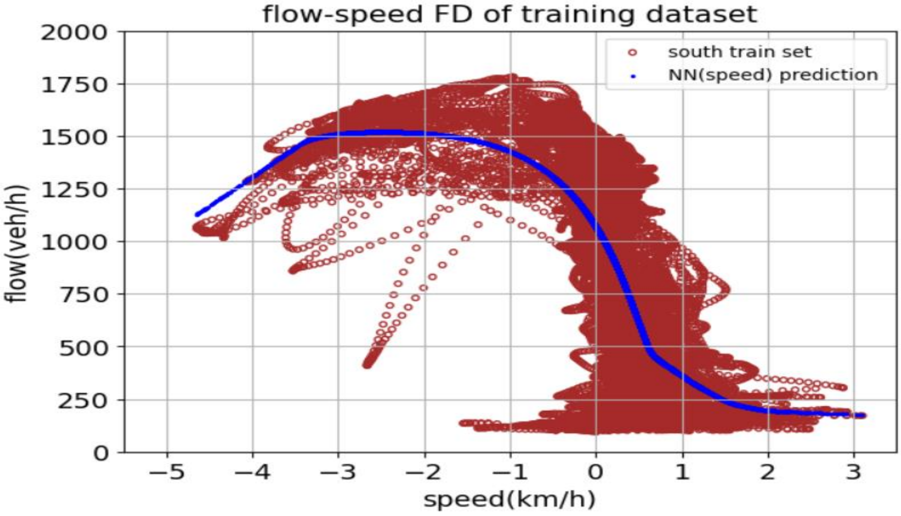
| | fk_id | date | speed | speed_std_dev | flow | Epoch_mcs |
|---|---|---|---|---|---|---|
| **10** | 1159 | 2018-10-01 00:10:00 | NaN | 18.422532 | NaN | 600.0 |
| **11** | 1159 | 2018-10-01 00:11:00 | NaN | 9.000000 | NaN | 660.0 |
| **12** | 1159 | 2018-10-01 00:12:00 | NaN | 51.006960 | NaN | 720.0 |
| **13** | 1159 | 2018-10-01 00:13:00 | NaN | 49.947840 | NaN | 780.0 |
| **14** | 1159 | 2018-10-01 00:14:00 | NaN | 23.191380 | NaN | 840.0 |
| **15** | 1159 | 2018-10-01 00:15:00 | 94.083548 | 17.879580 | 176.933333 | 900.0 |
| **16** | 1159 | 2018-10-01 00:16:00 | 94.064026 | 22.449960 | 176.666667 | 960.0 |
| **17** | 1159 | 2018-10-01 00:17:00 | 93.980823 | 42.395760 | 176.311111 | 1020.0 |
| **18** | 1159 | 2018-10-01 00:18:00 | 93.825538 | 25.200000 | 175.644444 | 1080.0 |
| **19** | 1159 | 2018-10-01 00:19:00 | 93.649874 | 24.059916 | 174.666667 | 1140.0 |

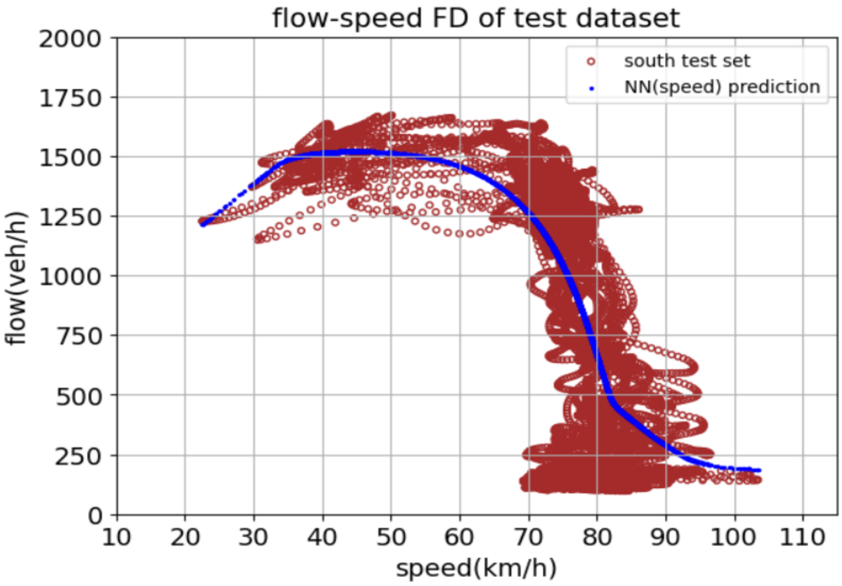
Removed Null values

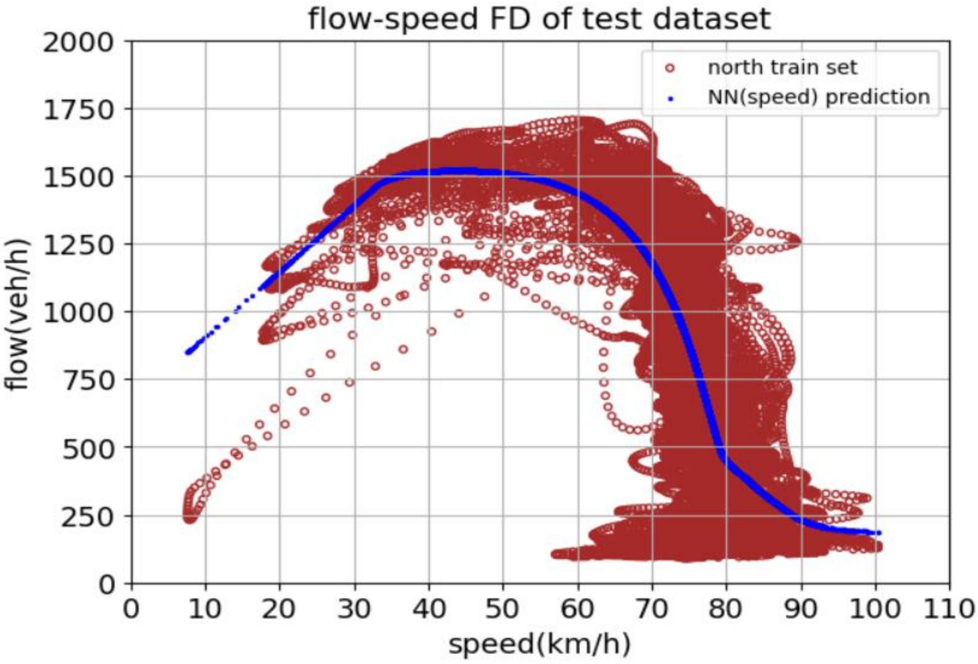
| | fk_id | date | speed | speed_std_dev | flow | Epoch_mcs |
|---|---|---|---|---|---|---|
| **0** | 1159 | 2018-10-01 00:15:00 | 94.083548 | 17.879580 | 176.933333 | 900.0 |
| **1** | 1159 | 2018-10-01 00:16:00 | 94.064026 | 22.449960 | 176.666667 | 960.0 |
| **2** | 1159 | 2018-10-01 00:17:00 | 93.980823 | 42.395760 | 176.311111 | 1020.0 |
| **3** | 1159 | 2018-10-01 00:18:00 | 93.825538 | 25.200000 | 175.644444 | 1080.0 |
| **4** | 1159 | 2018-10-01 00:19:00 | 93.649874 | 24.059916 | 174.666667 | 1140.0 |
| **...** | ... | ... | ... | ... | ... | ... |
| **44251** | 1159 | 2018-10-31 23:36:00 | 86.414550 | 13.227228 | 259.444444 | 2676960.0 |
| **44252** | 1159 | 2018-10-31 23:37:00 | 86.497375 | 26.939916 | 258.988889 | 2677020.0 |
| **44253** | 1159 | 2018-10-31 23:38:00 | 86.590658 | 11.879388 | 259.233333 | 2677080.0 |
| **44254** | 1159 | 2018-10-31 23:39:00 | 86.700006 | 33.614712 | 260.188889 | 2677140.0 |
| **44255** | 1159 | 2018-10-31 23:40:00 | 86.811252 | 13.708404 | 260.822222 | 2677200.0 |

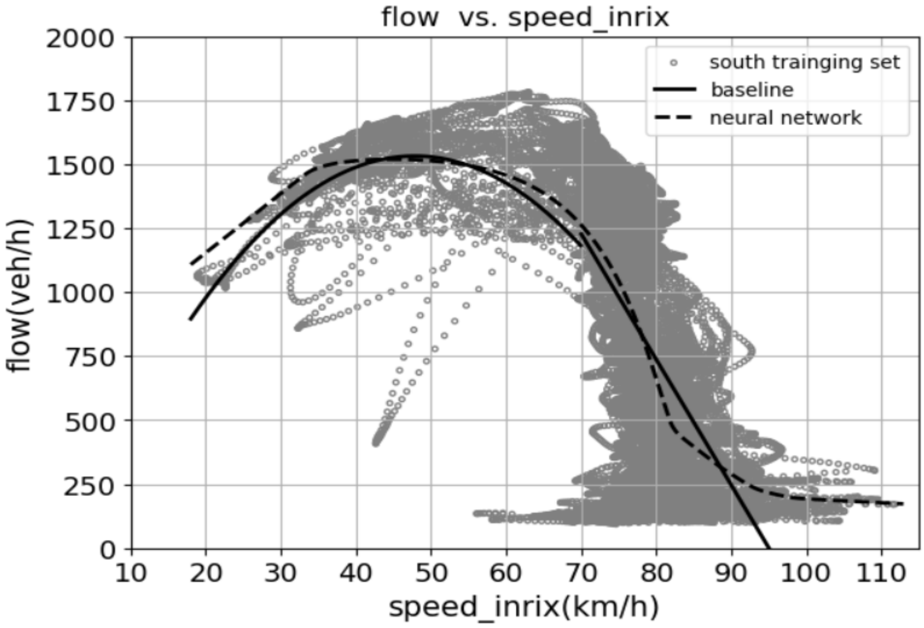
44256 rows × 6 columns

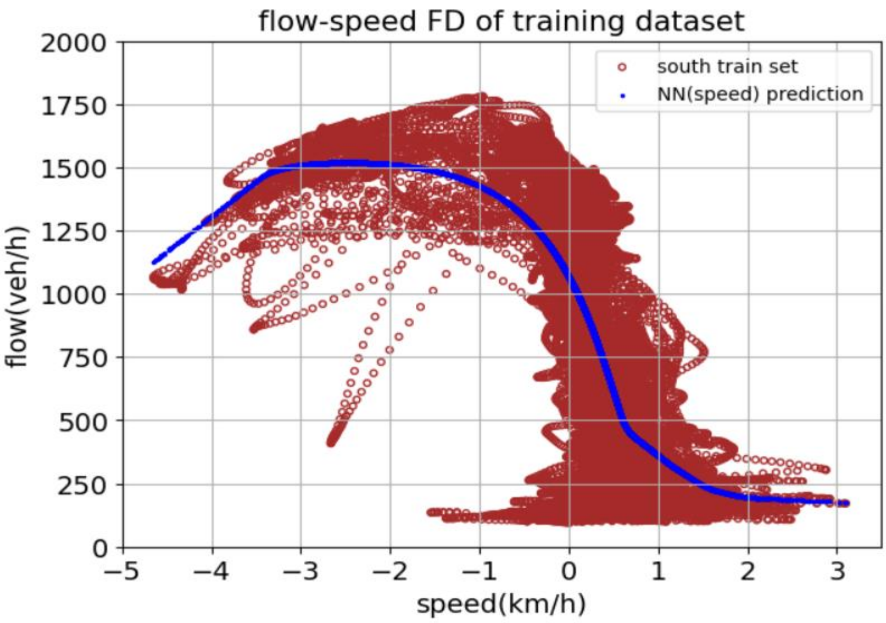
Plot original speed and smoothened speed for MCS South' Prediction




flow-speed FD of test dataset
north train set
NN(speed) prediction
flow(veh/h)
speed(km/h)
0
250
500
750
1000
1250
1500
1750
2000
0
10
20
30
40
50
60
70
80
90
100
110


flow vs. speed_inrix
south trainging set
baseline
neural network
flow(veh/h)
speed_inrix(km/h)
0
250
500
750
1000
1250
1500
1750
2000
10
20
30
40
50
60
70
80
90
100
110




Neural network with 2 features speed and travel

**Conclusion**

The Spatio-Temporal Multi-Agent Reinforcement Learning (ST-MARL) approach for society holds tremendous promise as a transformative paradigm for addressing complex societal challenges. This research has demonstrated the potential of ST-MARL in revolutionizing decision-making and coordination in diverse spatio-temporal environments.The collaborative and adaptive nature of ST-MARL allows multiple agents to learn and adapt their behaviors based on real-time spatio-temporal information, leading to efficient resource allocation, coordination, and improved overall system performance. By integrating reinforcement learning with spatial and temporal considerations, ST-MARL enables agents to navigate dynamic environments and handle real-world complexities effectively.

**References**

Wang, Y., Xu, T., Niu, X., Tan, C., Chen, E., &Xiong, H. (2020). STMARL: A spatio-temporal multi-agent reinforcement learning approach for cooperative traffic light control. IEEE Transactions on Mobile Computing, 21(6), 2228-2242.

Du, X., Wang, J., Chen, S., & Liu, Z. (2021). Multi-agent deep reinforcement learning with spatio-temporal feature fusion for traffic signal control. In Machine Learning and Knowledge Discovery in Databases. Applied Data Science Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part IV 21 (pp. 470-485). Springer International Publishing.

Zhou, T., Kris, M. L., Creighton, D., & Wu, C. (2019). GMIX: Graph-based spatial–temporal multi-agent reinforcement learning for dynamic electric vehicle dispatching system. Transportation Research Part C: Emerging Technologies, 144, 103886.

Novati, G., de Laroussilhe, H. L., &Koumoutsakos, P. (2021). Automating turbulence modelling by multi-agent reinforcement learning. Nature Machine Intelligence, 3(1), 87-96.

Alsalehi, S., Mehdipour, N., Bartocci, E., &Belta, C. (2021, December). Neural network-based control for multi-agent systems from spatio-temporal specifications. In 2021 60th IEEE Conference on Decision and Control (CDC) (pp. 5110-5115). IEEE.

Zhang, W., Liu, H., Wang, F., Xu, T., Xin, H., Dou, D., &Xiong, H. (2021, April). Intelligent electric vehicle charging recommendation based on multi-agent reinforcement learning. In Proceedings of the Web Conference 2021 (pp. 1856-1867).

Zhang, W., Liu, H., Wang, F., Xu, T., Xin, H., Dou, D., &Xiong, H. (2021, April). Intelligent electric vehicle charging recommendation based on multi-agent reinforcement learning.

In Proceedings of the Web Conference 2021 (pp. 1856-1867).

Agarwal, S., Wallner, G., Watson, J., & Beck, F. (2021). Spatio-temporal Analysis of Multi-agent Scheduling Behaviors on Fixed-track Networks. In 2022 IEEE 15th Pacific Visualization Symposium (PacificVis) (pp. 21-30). IEEE.

Qiu, W., Chen, H., &An, B. (2019, August). Dynamic Electronic Toll Collection via Multi-Agent Deep Reinforcement Learning with Edge-Based Graph Convolutional Networks. In IJCAI (pp. 4568-4574).