An Epidemic Model of Estimating Metro Station Vulnerabilities Towards Delay Propagation

Molkenboer, Laetitia; Schulte, Frederik; Zhu, Yongqiu; Cats, Oded

DOI
10.1109/MT-ITS68460.2025.11223584

Publication date
2025

Document Version
Final published version

Published in
2025 9th International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2025

An epidemic model of estimating metro station vulnerabilities towards delay propagation

1st Laetitia Molkenboer
Delft University of Technology
Delft, The Netherlands

2nd Frederik Schulte
Department of Transport Engineering & Logistics
Delft University of Technology
Delft, The Netherlands
f.schulte@tudelft.nl

3rd Yongqiu Zhu
Department of Transport & Planning
Delft University of Technology
Delft, The Netherlands
y.zhu-5@tudelft.nl

4th Oded Cats
Department of Transport & Planning
Delft University of Technology
Delft, The Netherlands
o.cats@tudelft.nl

2025 9th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS) | 979-8-3315-8063-6/25/$31.00 ©2025 IEEE | DOI: 10.1109/MT-ITS68460.2025.11223584

***Abstract*—Metro networks face operational challenges due to increasing ridership and system growth, particularly in managing delay propagation. Epidemiology models have recently been an interesting method in transportation research for studying delays. This study, therefore, aims to investigate if the Susceptible-infectious-susceptible (SIS) model is suitable to help model delay propagation in a metro network through its ability to reproduce the vulnerability of metro stations for specific instances. Using data from the Washington Metro Network, two groups of delay propagation instances were selected and used for model training and testing using a differential evolution algorithm. The results indicate that the vulnerability values as calculated from the real-life data do not follow the expected trend. Still, our model can capture this variation with good vulnerability estimation accuracy for both groups. Also, the predicted vulnerability values for the first group are more accurate than for the second group. However, limitations such as underestimation and overestimation of station vulnerabilities, and sensitivity to training data were observed. These challenges stemmed from the dynamics between specific parameters and the lack of additional factors.**

***Index Terms*—Delay propagation, Vulnerability, SIS model, Metro stations**

I. Introduction

The growth of cities due to urbanization has led to increased ridership and the growth of metro systems in distance (km of infrastructure) and number of stops [1]. As a result, the challenge of delays and their propagation to other network parts has become more crucial. Therefore, how and why delays propagate in metro networks must be studied, so that network operators can take the appropriate measures to prevent and mitigate them to ensure they do not negatively affect travelers' experiences.

One perspective in the literature is assessing the role of node and link vulnerability in propagating delays. Vulnerability is often described in the literature with the words susceptibility and serviceability [2]. Combining these keywords, vulnerability is the exposure of a public transport network to disruptions (susceptibility) and at the same time the ability of the PT network to cope with these disruptions (serviceability) [3]. One of the reasons for considering vulnerability in addition to network topology is that the effects of disruptions were found to be heterogeneous across metro stations and dependent on its location in the network as well as other station-level characteristics, meaning that only using topology analysis would fail to acknowledge other factors playing a role [4]. Furthermore, at the functional level (flow distribution) initial failure propagates faster than at the structural level (network topology) [5] because they have different sources of vulnerabilities [6], and, hence, only considering the structural level would give a limited picture of the problem.

Although advancements have been made in public transportation research with the introduction of vulnerability, those studies do not always include delay propagation and the role metro station vulnerability plays in delay propagation. Therefore, this study aims to fill the knowledge gap of how to model delay propagation in a metro network through the model's ability to reproduce the vulnerability of metro stations for specific instances.

For several transport modes, epidemic models have been used to study delay propagation as the spreading of diseases and propagation of delays show similarities. This method has proven promising in the railways [7], [8]. The studies that have been done for metro systems focused solely on congestion propagation and the passenger perspective, neglecting the operator perspective [9]–[11]. Therefore, the extent to which epidemic models can be used to model delay propagation from the operator's perspective is an interesting research direction.

II. Methodology

This section outlines the study's methodology. Section II-A covers instance selection, followed by vulnerability calculations in Section II-B. The model is detailed in Section II-C, with training and testing metrics in Section II-D.

A. Instance selection criteria

Instances from the data must be found that show delay propagation happened and its effects are felt at multiple stations. The found instances are grouped based on the primary delay station and direction.

The time window w of the instance determines which train movements are included and is defined by the arrival time of two trains. The start time is the arrival time of the previous train at the primary delay station. The end time is the moment when the last train affected by the delay at the primary delay station arrives at the station where all affected trains from all instances experience a delay. This definition means that not all stations on the affected lines are considered. The reason is that some stations experience delays only for certain instances. This results in significant variability in the delay impact at those stations, posing challenges in effectively capturing these variations and training the model to produce accurate results.

B. Calculation of station vulnerabilities using the data

Train movements determine the vulnerability of a station during the time window w in the direction of the primary delay. These movements share the same infrastructure section k at the station, as delays primarily affect nearby metro stations on the same line [12]. If a station has multiple distinct tracks and platforms, k limits the analysis to relevant lines. Since only movements in the delay's direction are considered, cross-platform delay propagation is excluded.

To determine if a train t is delayed, the arrival time difference at station i between two consecutive trains of the same line, t and $t-1$, is calculated and represented as $b_{t,t-1}$. This difference should be equal to the scheduled headway $m_{t,t-1}$ between trains t and $t-1$. However, due to potential variability in service, deviations in headway between two trains could occur without impacting overall network performance. A delay threshold parameter, denoted as δ, is introduced to account for these variations. If $b_{t,t-1}$ exceeds the sum of the scheduled headway $m_{t,t-1}$ and the delay threshold δ, the train t is considered delayed. In such cases, the variable a_{tiw}, which indicates whether train t was delayed during time window w at station i, is set to 1. These considerations result in (1).

$$a_{tiw} = \begin{cases} 1 & \text{if } b_{t,t-1} \geq m_{t,t-1} + \delta, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

The vulnerability of station i is the ratio of delayed trains to all observed trains T_{ikw} stopping at a station i using infrastructure k. Mathematically, the vulnerability of station i translates to (2).

$$v_{ik,data} = \frac{\sum_{t=1}^{T_{ikw}} a_{tiw}}{|T_{ikw}|} \tag{2}$$

C. SIS mathematical model for a metro network

The translation of the heterogeneous SIS model to a metro network is explained in sections II-C1 through II-C4.

1) Network Graph: An undirected P-space graph was created to represent the stations and the links between them. A weight is assigned to each link in the network. The propagation strength of a delay diminishes with increasing distance [9]. Hence, the further away two stations are from each other, the less chance a delay starting at station j has to reach station i. Therefore, The link weight is the inverse of the travel time t_{ij} between nodes i and j as shown in (3).

$$w_{ij} = \frac{1}{t_{ij}} \tag{3}$$

Then, for three types of edges, the weight is increased to reflect the temporarily increased chance of infection due to the primary delay. These adjustments are based on either the P-space or L-space path between two nodes. The L-space is a network graph where all stations directly adjacent to each other and connected by a service are connected by an edge. In the P-space, each station is connected to the stations that can be reached without a transfer [13]. Which edges to adjust is visualized in Fig. 1 and listed below:

1) Edges between the primary delay node and the nodes in the direction of the delay (P-space);
2) Edges between the nodes in the direction of the delay (L-space) ;
3) Edges that are part of the L-space path between two nodes and the path includes the primary delay node.

The edges of types 1, 2, and 3 are part of set E. First, the weights of type 1 and 2 edges were changed in the P-space using (4).

$$w_{ij} = w_{ij} + \frac{w_{ij} * \alpha * a_e}{e^{h(p,j)*\gamma}} \tag{4}$$

The higher the primary delay duration α, the more the weights are increased as a more severe delay causes more trouble [14]. a_e represents the number of lines that run on an edge e. The higher the value for a_e, the higher the delay impact because the complexity of operations increases with multiple train operation routes [15]. $h(p,j)$ defines how many nodes are between the primary delay node p and the currently considered node j in the L-space and, hence, captures the space component of delay propagations. γ is a to-be-trained parameter that captures how quickly the propagation effect diminishes with time.

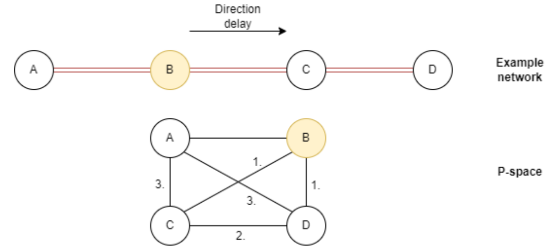

Fig. 1. Example network and P-space showing which edge types weight would be adjusted. The yellow node is the infected node in the example.

Finally, the type 3 edge weights are adjusted in the P-space. First, the L-space path is determined. Then, it is calculated which percentage of the edges in the path is before and after the primary delayed station. The initial link weight is then multiplied by the two percentages. The link weight portion reflecting the edges after is adjusted using (4). That value is then added to the unaffected portion of the link weight. When all the necessary edge weights are changed, all the edge weights are normalized using (5), where N represents the number of stations.

$$w_{ij} = \frac{w_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}} \tag{5}$$

2) Infection Rate: The infection rate of the model represents the susceptibility part of the vulnerability definition. All trains arriving at station i can infect the station if they are delayed. Also, transfer stations receive many more trains increasing their chance of infection [15]. Furthermore, stations linked to transfer stations could be infected by delays from trains of the other lines if they use the same infrastructure section k. The infection rate equation is presented in (6).

$$i_{ik} = \sum_{j}^{J}(w_{ji} + (w_{ji} * \sum_{s\epsilon S, s\neq j,i} w_{sj} * g_j)) + z_{ik} \tag{6}$$

The infection rate i_{ik} for section k at station i is determined by summing the edge weights between station i and stations j, which are the stations directly connected to station i in the P-space, meaning that no transfer is needed. If station j is a transfer station (g_j = 1), stations s also have to be considered given that the line connecting stations s and station j uses the same infrastructure section k as the line connecting station i to j. Also, the parameter z_{ik} captures any unobservable factors influencing the infection rate and is also trained.

3) Recovery Rate: The recovery rate highlights the ability of a station to cope with a disturbance. The recovery rate of a station i is dependent on the number of tracks and the traffic density at the station. More tracks increase the station's flexibility because traffic operators can appoint the trains to more tracks. Also, higher service frequency lines are affected more than lines with lower service frequency [3], [14]. Stations with a higher traffic density are less flexible because more traffic has to be considered and, hence, have a lower recovery rate. The number of tracks u_{ki} and the traffic density d_{ki} might differ per section k due to different serviced lines. The recovery rate equation is presented in (7).

$$r_{ik} = c_{ik} + \left(\frac{u_{ki}}{d_{ki}}\right)^{\theta} \tag{7}$$

In (7) d_{ki} is the number of trains stopping at a station i per hour per infrastructure section k using the scheduled headway of line l.

θ is introduced in (7) to see how heterogeneous the recovery rates of stations are. Also, the parameter c_{ik} captures any unobservable factors influencing the recovery rate. Both θ and c_{ik} are trained and ≥ 0.

4) Vulnerability equation: Equation 8 shows how the vulnerability of station i is calculated for a specific section k. This equation reflects the definition of vulnerability given earlier in section I. As a result of a vulnerability value per k, a station could have multiple vulnerability values if it has several infrastructure sections that are all part of the training.

$$0 = -r_{ik} * v_{ik,model} + (1 - v_{ik,model}) * i_{ik} \tag{8}$$

D. Model training and testing metrics

The training of the four parameters presented earlier is done over a group of similar instances, which are randomly split into two sub-groups: 1. training instances and 2. testing instances. The first sub-group contains 80% of the instances and the second sub-group 20%. The mean squared error is computed for each instance used in the training. This calculation is based on the vulnerabilities calculated from the data and determined by the model, using the inner part of the formula shown in (9), where N is the number of all stations of all affected lines. The objective function used in the training is to minimize the sum of the MSE of all instances F, which is the outer part of (9).

$$\min \sum_{f=1}^{F} \text{MSE}_f = \sum_{f=1}^{F} \frac{\sum_{i=1}^{N}(v_{ik,data} - v_{ik,model})^2}{N} \tag{9}$$

Additionally, the mean vulnerabilities of the model and data are compared and the differences in vulnerability between the data and model are determined for each station. All of this is done per k if a station i has multiple infrastructure sections and they are part of the training.

III. Application

A. Case Study: Washington Metro network

This section discusses the case study data used and its processing. The Washington Metropolitan Area Transit Authority (WMATA) provided the data about the Washington metro network.

At least 10 similar instances had to be found to form a group to ensure enough data to train with. Instances where the primary delay led to recovery strategies such as short-turning were not considered. Moreover, all stations of the considered lines had to be open. Two groups of instances were formed. The first group includes delays starting at King St-Old Town station, affecting Yellow and Blue trains toward Greenbelt and Downtown Largo. It covers stations between King St-Old Town and Rosslyn, and King St-Old Town and L'Enfant Plaza (Fig. 2).

The second group concerns primary delays of Blue, Silver, and Orange trains in the direction of Franconia-Springfield, Wiehle-Reston East, and Vienna respectively. The considered stations are those between Stadium-Armory and Rosslyn and displayed in Fig. 3.

For each of the instances, the vulnerability was calculated using the data. The delay threshold to determine whether a

Fig. 2. This figure shows the stations considered for group 1.

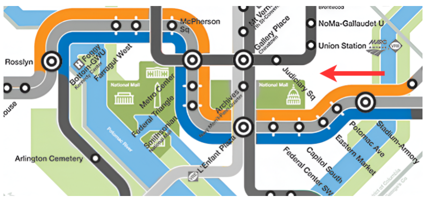

Fig. 3. This figure shows the stations considered for group 2.

train was delayed was set to two minutes. The vulnerabilities of the not-considered stations were forced to be 0.

B. Implementation

A few steps were taken to implement the model. P-space and L-space files from previously done research were used to create the necessary graphs in Python [16]. The travel times were also obtained from these data files and set as edge attributes. Then, the groups of instances were used to train the model. For this training, the differential evolution algorithm of the SciPy Python package was used [17]. The bounds on the parameters were as follows:

- γ: Diminishing effect of delay propagation in 4 [0, 2];
- z_{ik}: Corrects for factors not considered in 6 [0, 0.5];
- θ: Heterogeneity of station recovery rates in 7 [0.5, 2];
- c_{ik}: Corrects for factors not considered in 7 [0, 0.5].

The bounds were established through test runs aimed at characterizing the model's behavior. The training process consistently resulted in $\theta = 0$, a result attributable to the model's formulation rather than inherent patterns in the data.

IV. Results

The training and testing results for the two formed groups are presented in sections IV-A and IV-B. Then, the sensitivity analysis results are given in subsection IV-C.

A. Results Group 1: King St-Old Town station

The vulnerability values calculated from the data are displayed in Fig. 4. The vulnerabilities were expected to decrease with increasing distance from the primary delay station. A delayed train should catch up on its delay as it uses buffer time. However, the vulnerabilities across stations show other trends. There are a few reasons for this behavior. Firstly, delayed

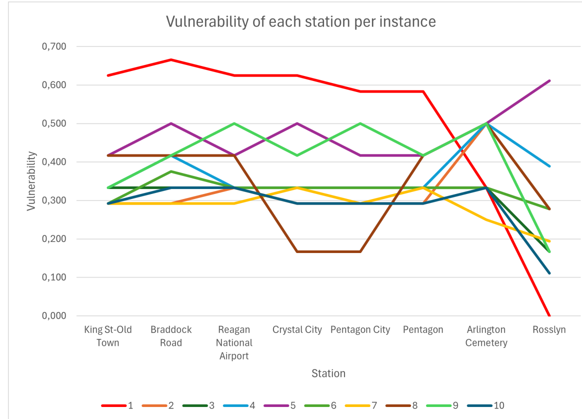

Fig. 4. Vulnerability of each group 1 station per instance.

trains unaffected by the primary delay are also included in the vulnerability calculations. Another explanation is that the arrival time is around the two-minute threshold, making it delayed at one station, but not at the next. Lastly, the number of trains used in the vulnerability calculations differs across stations due to the time window.

Furthermore, more factors seem to influence the propagation effects. For instances 3 and 8, the primary delay is a Blue line train, which is five minutes delayed at 11:20 am on a weekday and travels with a headway of 12 minutes. The effects of these instances are expected to be similar. However, the vulnerability trends for these two instances are different. Hence, other factors must explain this difference in delay propagation effects.

The training MSE value is 0.022. The trained parameters are presented in Table I. The low γ value indicates that the edge weights were minimally adjusted. For θ the lower bound was found as the best value, meaning the recovery rates were made more homogeneous. The large range of values used for the z_{ik} and c_{ik} suggest that the training algorithm chose to compensate with those parameters, instead of using γ and θ. The trained parameters were then tested, whose results are shown in Table II.

The MSE values of 0.004 for both testing instances mean the model could predict the vulnerability values fairly accurately. The difference between the mean vulnerabilities is also small

TABLE I
TRAINED PARAMETERS FOR GROUP 1: KING ST-OLD TOWN

γ		0.455
θ		0.500
Stations	z_{ik}	c_{ik}
Arlington Cemetery	0.254	0.451
Braddock Road	0.390	0.007
Crystal City	0.264	0.248
King St-Old Town	0.303	0.495
L'Enfant Plaza	0.222	0.127
Pentagon	0.269	0.417
Pentagon City	0.412	0.457
Reagan National Airport	0.217	0.378
Rosslyn	0.150	0.358

TABLE II
TESTING METRICS FOR GROUP 1.

Metric	*Testing instance 1*	*Testing instance 2*
MSE	0.004	0.004
Mean vulnerability data	0.319	0.336
Mean vulnerability model	0.295	0.299

TABLE III
COMPARISON OF VULNERABILITIES AS DETERMINED FROM THE DATA AND BY THE MODEL FOR GROUP 1. DIFF = DATA - MODEL

	Testing instance 1			Testing instance 2		
Stations	*Data*	*Model*	*Diff*	*Data*	*Model*	*Diff*
Arlington Cemetery	0.333	0.399	-0.066	0.333	0.397	-0.064
Braddock Road	0.333	0.265	0.068	0.375	0.277	0.098
Crystal City	0.333	0.288	0.045	0.333	0.294	0.039
King St-Old Town	0.333	0.235	0.098	0.292	0.240	0.052
L'Enfant Plaza	0.375	0.330	0.045	0.417	0.328	0.089
Pentagon	0.333	0.317	0.016	0.333	0.324	0.009
Pentagon City	0.333	0.275	0.058	0.333	0.281	0.052
Reagan Airport	0.333	0.279	0.054	0.333	0.290	0.043
Rosslyn	0.167	0.266	-0.099	0.278	0.264	0.014

for both testing instances. The higher model mean suggests underestimation by the model. To better understand this underestimation the difference per station i was determined and is displayed in Table III. For almost every station the vulnerability is underestimated. Only the stations Arlington Cemetery and Rosslyn were slightly overestimated. Looking back at Table I, the training algorithm chose to compensate the stations with a too low recovery rate instead of stations with a too low infection rate. The recovery rate values will be higher than the infection rates, leading to lower vulnerability values. As the vulnerability values of the testing instances are similar to the mean vulnerability values of the training instances, the testing instances will also be underestimated.

B. Results group 2: Stadium-Armory station

The station vulnerabilities as calculated from the data are visualized in Fig. 5.

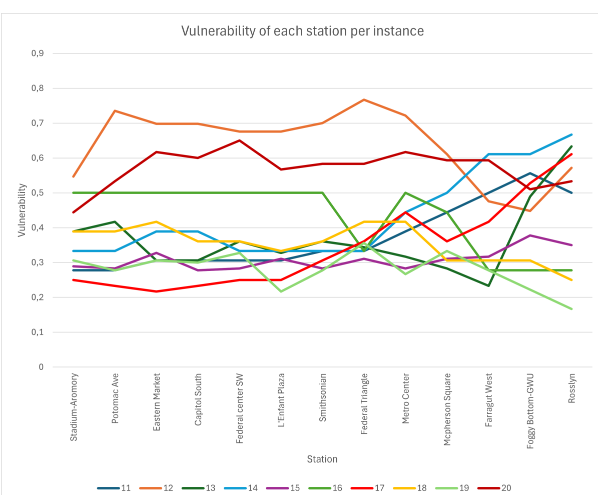

Fig. 5. Vulnerability of each group 2 station per instance.

The vulnerability values of the stations in group 2 have a larger range than group 1. Furthermore, these instances also fluctuate instead of showing the expected decreasing trend. The stations considered for group 2 are in the busiest part of the network. If the traffic density is high, it is more difficult for the trains to catch up on their delay. Also, instances 11, 13, 14, and 17 show almost identical behavior. However, these four instances differ in line, time of primary delay, and scheduled headways. Therefore, the group 2 results reiterate that more factors could help explain the vulnerability trends.

The MSE value of the trained model for group 2 is 0.043. Table IV shows the trained parameter values for group 2.

Also for group 2, the trained value for γ is low and for θ is the lower bound. Due to the central location of the group 2 stations in the network, they have high baseline infection rate values (so without considering z_{ik}), but the training algorithm still chose to use high values for z_{ik} for most stations. Similarly, while the baseline recovery rates of these stations are low, the training algorithm does not use the upper bound.

Table V indicates how well the model performed on the testing instances. The MSE values are higher and the distance between the means is larger in comparison to the group 1 results. Group 2 has a larger range of vulnerability values, making it more difficult to find parameter values that fit, which could partially explain the worse performance. Table VI gives more insights into how the model performed for the stations.

The vulnerability values of the group 2 testing instances are overestimated. This overestimation is not unexpected, because the data vulnerability values are lower for these testing instances than the training instances. The training algorithm prioritizes fitting the largest group of similar instances. If they have high vulnerability values, the model adjusts accordingly, leading to overestimation of instances with lower vulnerabili-

TABLE IV
TRAINED PARAMETERS FOR GROUP 2: STADIUM-ARMORY.

γ	0.342	
θ	0.500	
Stations	z_{ik}	c_{ik}
Stadium-Armory	0.219	0.453
Potomac Ave	0.151	0.305
Eastern Market	0.118	0.220
Capitol South	0.151	0.162
Federal Center SW	0.320	0.305
L'Enfant Plaza	0.258	0.212
Smithsonian	0.306	0.195
Federal Triangle	0.348	0.256
Metro Center	0.436	0.242
Mcpherson Square	0.302	0.128
Farragut West	0.320	0.179
Foggy Bottom - GWU	0.440	0.247
Rosslyn	0.379	0.100

TABLE V
TESTING METRICS FOR GROUP 2.

Metric	*Testing instance 1*	*Testing instance 2*
MSE	0.018	0.027
Mean vulnerability data	0.308	0.279
Mean vulnerability model	0.424	0.414

TABLE VI

COMPARISON OF VULNERABILITIES AS DETERMINED FROM THE DATA AND BY THE MODEL FOR GROUP 2. DIFF = DATA - MODEL

	Testing instance 1			Testing instance 2		
Stations	*Data*	*Model*	*Diff*	*Data*	*Model*	*Diff*
Stadium-Armory	0.289	0.319	-0.030	0.389	0.413	-0.024
Potomac Ave	0.283	0.321	-0.038	0.389	0.457	-0.068
Eastern Market	0.328	0.383	-0.055	0.417	0.456	-0.039
Capitol South	0.278	0.279	-0.001	0.361	0.424	-0.063
Federal Center	0.283	0.293	-0.010	0.361	0.435	-0.074
L'Enfant Plaza	0.311	0.450	-0.139	0.333	0.540	-0.207
Smithsonian	0.283	0.428	-0.145	0.361	0.505	-0.144
Federal Triangle	0.311	0.388	-0.077	0.417	0.482	-0.065
Metro Center	0.283	0.502	-0.219	0.417	0.649	-0.232
Mcpherson Sq	0.311	0.441	-0.130	0.306	0.413	-0.107
Farragut West	0.317	0.464	-0.147	0.306	0.453	-0.147
Foggy Bottom	0.378	0.564	-0.186	0.306	0.548	-0.242
Rosslyn	0.350	0.587	-0.237	0.250	0.592	-0.342

ties, which emphasizes the need for more factors.

C. Sensitivity analysis

The results of groups 1 and 2 show similar results even though they are about different parts of the network. Therefore, a sensitivity analysis was performed to investigate the model behavior in more detail. To test how sensitive the model performance is to which instances are used in the model training and testing, the division of training and testing instances was changed for group 1. Table VII shows the results of this sensitivity analysis.

Comparing table II to VII shows that when the other instances were used for testing instead of 2 and 9, the model performed better. The MSE was lower and the difference between the two means was smaller. Therefore, the model is sensitive to which instances are used for training and testing, meaning that model robustness needs to be improved.

V. CONCLUSION

This study proposed a model inspired by the SIS concept to estimate the vulnerabilities of metro stations towards delay propagation for specific instances. The model was constructed and trained for several parameters using data from the Washington Metro network. The model training and testing results for two groups of instances were analyzed and the conditions under which the model produced the results were reflected upon through a sensitivity analysis. The results indicate that the model can reproduce the vulnerabilities of instances accurately for both groups. At the same time, the sensitivity analysis showed the model's sensitivity to outliers, implying the need for model robustness improvement in the future. Furthermore, there is a dependence on the input instances, which means that more factors need to be considered to decrease this dependence and improve how the model reacts to 'odd' instances.

TABLE VII

TESTING METRICS FOR SENSITIVITY ANALYSIS OF INSTANCE DIVISION FOR TRAINING AND TESTING.

Metric	*Instance 2*	*Instance 9*
MSE	0.010	0.028
Mean vulnerability data	0.350	0.417
Mean vulnerability model	0.304	0.280

Even so, the model allows network operators to identify the vulnerable stations in the network, under which circumstances they are vulnerable, and which factors play a role in this vulnerability. Also, the network operator can be more proactive with targeted infrastructure upgrades to improve station delay recovery ability and to prevent/mitigate delay propagations.

A future research direction is to include additional factors that should help decrease the variation now seen across similar instances. Another direction could be the application of the model to additional parts of the WMATA network or a completely different network to further investigate its effectiveness and identify aspects that need further development.

ACKNOWLEDGMENT

The authors thank WMATA for the data provision.

REFERENCES

[1] UITP, "World metro figures 2021," May 2022. [Online]. Available: https://www.uitp.org.

[2] S. Pan, H. Yan, J. He, and Z. He, "Vulnerability and resilience of transportation systems: A recent literature review," *Physica A: Statistical Mechanics and its Applications*, vol. 581, 2021.

[3] M. Yap and O. Cats, "Predicting disruptions and their passenger delay impacts for public transport stops," *Transportation*, vol. 48, pp. 1703–1731, 2020.

[4] N. Zhang, D. Graham, D. Hörer, and P. Bansal, "A causal inference approach to measure the vulnerability of urban metro systems," *Transportation*, vol. 48, pp. 3269–3300, 2021.

[5] C. Chen, S. Wang, J. Zhang, and X. Gu, "Modeling the vulnerability and resilience of interdependent transportation networks under multiple disruptions," *Journal of Infrastructure Systems*, vol. 29, 2023.

[6] S. Chopra, T. Dillon, M. Bilec, and V. Khanna, "A network-based framework for assessing infrastructure resilience: A case study of the London metro system," *Journal of The Royal Society Interface*, vol. 13, 2016.

[7] M. M. Dekker, A. N. Medvedev, J. Rombouts, G. Siudem, and L. Tupikina, "Modelling railway delay propagation as diffusion-like spreading," *EPJ Data Science*, vol. 11, 2022.

[8] Gurin, D., Prokhorchenko, A., Kravchenko, M. and Shapoval, G., "Development of a method for modelling delay propagation in railway networks using epidemiological SIR models", *Eastern-European Journal of Enterprise Technologies*, vol. 6, 2020.

[9] C. Jia, S. Zheng, H. Qian, B. Cao, and K. Zhang, "Analysis of crowded propagation on the metro network," *Sustainability*, vol. 14, 2022.

[10] Z. Zeng and T. Li, "Analyzing congestion propagation on urban rail transit oversaturated conditions: A framework based on SIR epidemic model," *Urban Rail Transit*, vol. 4, pp. 130–140, 2018.

[11] Wang, X., Yao, E. and Liu, S., "Simulation of metro congestion propagation based on route choice behaviors under emergency-caused delays", *Applied Sciences*, vol. 9, pp 348–359, 2019.

[12] O. Cats and A. Hijner, "Quantifying the cascading effects of passenger delays," *Reliability Engineering and System Safety*, vol. 212, p. 107629, 2021

[13] C. Von Ferber, T. Holovatch, Y. Holovatch, and V. Palchykov, "Public transport networks: Empirical analysis and modeling," *The European Physical Journal B*, vol. 68, pp. 261–275, 2009.

[14] A. D. Marra and F. Corman, "From delay to disruption: Impact of service degradation on public transport networks," *Transportation Research Record*, vol. 2674, pp. 886–897, 2020.

[15] J. Lu, X. Ma, and Y. Xing, "Risk factors affecting the severity of disruptions in metro operation in Shanghai, 2013-2016," *Journal of Transportation Safety & Security*, vol. 13, pp. 69–92, 2021.

[16] O. Cats, S. van Cranenburgh, S. Vijlbrief, P. Krishnakumari, and R. Massobrio, "A curated data set of p-space representations for 51 metro networks worldwide," Oct. 2019. [Online].

[17] SciPy, "Differentialevolution," *SciPy Documentation*. [Online].