



UNIVERSIDADE  
ESTADUAL DE LONDRINA

---

JOSÉ DE JESUS BRAGA MARTEL JÚNIOR

**MODELAGEM STARIMA E LSTM NA PREVISÃO DO  
TRÁFEGO URBANO VEICULAR:  
UMA ANÁLISE COMPARADA**

---

Londrina  
2025

JOSÉ DE JESUS BRAGA MARTEL JÚNIOR

**MODELAGEM STARIMA E LSTM NA PREVISÃO DO  
TRÁFEGO URBANO VEICULAR:  
UMA ANÁLISE COMPARADA**

Dissertação apresentada ao Programa de Mestrado em Engenharia Elétrica da Universidade Estadual de Londrina para obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Taufik Abraão.

Londrina  
2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

M376m Martel Júnior, José de Jesus Braga.  
Modelagem STARIMA e LSTM na previsão do tráfego urbano veicular : uma análise comparada / José de Jesus Braga Martel Júnior. - Londrina, 2025.  
99 f. : il.

Orientador: Taufik Abrão.  
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Estadual de Londrina, Centro de Tecnologia e Urbanismo, Programa de Pós-Graduação em Engenharia Elétrica, 2025.  
Inclui bibliografia.

1. STARIMA - Tese. 2. LSTM - Tese. 3. Controle de tráfego e congestionamento - Tese. 4. modelos de tráfego urbano veicular - Tese. I. Abrão, Taufik. II. Universidade Estadual de Londrina. Centro de Tecnologia e Urbanismo. Programa de Pós-Graduação em Engenharia Elétrica. III. Título.

CDU 62

JOSÉ DE JESUS BRAGA MARTEL JÚNIOR

**MODELAGEM STARIMA E LSTM NA PREVISÃO DO  
TRÁFEGO URBANO VEICULAR:  
UMA ANÁLISE COMPARADA**

Dissertação apresentada ao Programa de Mestrado em Engenharia Elétrica da Universidade Estadual de Londrina para obtenção do título de Mestre em Engenharia Elétrica.

**BANCA EXAMINADORA**

---

Orientador: Prof. Dr. Taufik Abraão  
Universidade Estadual de Londrina - UEL

---

Prof. Dr. Bruno Bogaz Zarpelão  
Universidade Estadual de Londrina - UEL

---

Profa. Dra. Maria Bernadete Morais de França  
Universidade Estadual de Londrina - UEL

Londrina, 25 de julho de 2025.

*Este trabalho é dedicado aos meus amados filhos:  
Tiago, Alexandre, Júlia, Bruno e Breno.  
E a minha amada esposa Juliana.*

## **AGRADECIMENTOS**

Agradeço aos colegas de laboratório pelo apoio nas atividades diárias, pelas discussões e pela colaboração constante. Ao Professor Dr. Taufik, pela orientação dedicada, pelas críticas construtivas e pela confiança ao longo desta pesquisa. Ao meu amigo Dr. Rafael A. Pedriali, pelo apoio e contribuições nesta jornada. Aos professores do departamento, pelos ensinamentos e contribuições que ampliaram este trabalho. À minha família, pelo amor, paciência e incentivo, que tornaram esta conquista possível.

*Não encontre defeitos, encontre soluções.  
Qualquer um sabe queixar-se. (Henry Ford)*

## RESUMO

MARTEL JÚNIOR, José de Jesus B. **Modelagem STARIMA e LSTM na previsão do tráfego urbano veicular**: uma análise comparada. 2025. 99 f. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Estadual de Londrina, Londrina, 2025.

O crescimento das cidades e da frota de veículos tem tornado o tráfego urbano um desafio complexo para a gestão e o planejamento, exigindo soluções inovadoras para otimizar o fluxo veicular e promover a sustentabilidade. A modelagem preditiva emerge como uma ferramenta importante, utilizando dados históricos (como velocidade, fluxo e densidade) e variáveis contextuais (como eventos, condições climáticas e obras em andamento) para prever o comportamento futuro do tráfego. Essa abordagem facilita ajustes em tempo real, otimiza os tempos de semáforos e rotas de transporte público, e possibilita alertar motoristas sobre possíveis atrasos, permitindo o uso de rotas alternativas. No planejamento de longo prazo, auxilia na previsão da demanda por diferentes modais de transporte e na identificação de áreas que necessitam de melhorias na infraestrutura, promovendo soluções como a ampliação do transporte público e a criação de ciclovias. A relevância dessa proatividade é evidente, especialmente para o desenvolvimento sustentável, adaptando-se ao crescimento populacional e às mudanças nos padrões de mobilidade. A análise preditiva transforma dados em informações acionáveis, tornando os espaços urbanos mais produtivos e otimizados, beneficiando tanto a gestão reativa quanto o planejamento estratégico. Diante deste tema, esta dissertação propõe explorar e implementar técnicas de modelagem preditiva, com foco em dois modelos centrais: o modelo espaço-tempo autorregressivo integrado à média móvel STARIMA (*Space-Time Autoregressive Integrated Moving Average*) e o modelo baseado em redes neurais recorrentes LSTM (*Long Short-Term Memory*). O STARIMA se destaca por sua capacidade de capturar dependências espaço-temporais presentes no tráfego, sendo particularmente eficaz em previsões de curto prazo, considerando variações de velocidade e topologia da rede viária. Por sua vez, o modelo LSTM oferece uma abordagem não linear, com maior capacidade de generalização e aprendizado de padrões complexos e de longo prazo, especialmente útil em contextos com maior variabilidade e dinâmica temporal. Ambos os modelos são avaliados sob a perspectiva do monitoramento, previsão e controle de tráfego urbano (TMPC – *traffic monitoring, prediction and control*) em ambientes de cidades inteligentes. Nesta dissertação será realizada uma comparação dos modelos STARIMA e LSTM para a previsão do tráfego urbano veicular para curto e longo prazo usando conjuntos de dados reais do trânsito de São Paulo e Distrito Federal.

**Palavras-chave:** cidades inteligentes; modelos de tráfego urbano veicular; controle de tráfego e congestionamento; sustentabilidade; STARIMA; LSTM.

## ABSTRACT

MARTEL JÚNIOR, José de Jesus B. **STARIMA AND LSTM modeling in urban vehicle traffic prediction: a comparative analysis.** 2025. 99 p. Graduate Program in Electrical Engineering - State University of Londrina, Londrina, 2025.

The growth of cities and vehicle fleets has made urban traffic a complex challenge for management and planning, requiring innovative solutions to optimize vehicular flow and promote sustainability. Predictive modeling emerges as an important tool, utilizing historical data (such as speed, flow, and density) and contextual variables (such as events, weather conditions, and ongoing construction) to forecast future traffic behavior. This approach facilitates real-time adjustments, optimizes traffic light timing and public transportation routes, and enables drivers to be alerted about potential delays, allowing the use of alternative routes. In long-term planning, it aids in forecasting demand for different transportation modes and identifying areas that need infrastructure improvements, promoting solutions such as expanding public transportation and creating bike lanes. The relevance of this proactive approach is evident, especially for sustainable development, adapting to population growth and changes in mobility patterns. Predictive analysis transforms data into actionable information, making urban spaces more productive and optimized, benefiting both reactive management and strategic planning. Given this theme, this dissertation proposes to explore and implement predictive modeling techniques, focusing on two central models: the Space-Time Autoregressive Integrated Moving Average model STARIMA (Space-Time Autoregressive Integrated Moving Average) and the model based on recurrent neural networks LSTM (Long Short-Term Memory). STARIMA stands out for its ability to capture spatio-temporal dependencies present in traffic, being particularly effective in short-term forecasts, considering variations in speed and road network topology. In turn, the LSTM model offers a non-linear approach, with greater generalization capacity and learning of complex and long-term patterns, especially useful in contexts with greater variability and temporal dynamics. Both models are evaluated from the perspective of traffic monitoring, prediction, and control in urban traffic (TMPC – traffic monitoring, prediction and control) in smart city environments. In this dissertation, a comparison of the STARIMA and LSTM models will be conducted for the prediction of urban vehicular traffic in the short and long term using real traffic datasets from São Paulo and the Federal District.

**Key-words:** smart cities; urban vehicular traffic models; traffic control and congestion; sustainability; STARIMA; LSTM.

## LISTA DE FIGURAS

<b>Figura 1</b> – Camadas do Sistema de Transporte Inteligente (adaptado de [1]).	19
<b>Figura 2</b> – Arquitetura do Sistema de Transporte Inteligente (adaptado de [1]).	20
<b>Figura 3</b> – Visão geral da estrutura ATCS (adaptado de [1]).	20
<b>Figura 4</b> – Estrutura básica de RNN. Fonte: [2].	33
<b>Figura 5</b> – Arquitetura de um bloco de memória de LSTM. Fonte: [2].	35
<b>Figura 6</b> – Exemplo de uma aproximação de um semáforo de tempo fixo [Fonte: autor].	37
<b>Figura 7</b> – Chegada de usuários em um Frame $F$ com $n$ intervalos de tempo considerando distribuição Poisson $\times$ Beta e $M = 100$ entidades.	38
<b>Figura 8</b> – (a) Função Beta como $Beta(\alpha, \beta)$ . (b) Probabilidade de ativação do usuário, $p(t, \beta)$ . (c) Probabilidade de ativação do usuário, $p(t, \alpha)$ .	38
<b>Figura 9</b> – Estrutura em árvore típica da rede rodoviária para fluxo de tráfego.	42
<b>Figura 10</b> – Evolução de $z(t)$ para $L^{(\ell)}[z(t)]$ com $N = 6$ sites. Referente à Equação (3.8).	44
<b>Figura 11</b> – Estrutura em árvore típica da rede rodoviária para fluxo de tráfego. Os pontos representam locais de medição e setas direção do fluxo. $N = 13$ sites. Sistema referente à Equação (3.8).	44
<b>Figura 12</b> – Gráfico do resultado da simulação de $L^{(\ell)}[z(t)]$ com $N = 13$ sites e diferentes $z_{in}$ 's. Referente à Equação (3.8).	46
<b>Figura 13</b> – (Conjunto de dados 1) São Paulo: Sensores de radar posicionados ao longo da Avenida 23 de Maio, parte do Corredor Norte-Sul no centro da cidade. Esta via expressa urbana é uma via arterial importante, com alto volume de tráfego e granularidade temporal.	47
<b>Figura 14</b> – (Conjunto de Dados 2) Brasília: Dados de sensores coletados nos trechos rodoviários DF-027 (EPJK), DF-035 (EPCV) e DF-	

	463 localizados no setor sul do Distrito Federal. Essas vias são oficialmente classificadas como duplicadas e pavimentadas, o que permite alto fluxo de veículos.....	47
<b>Figura 15</b>	– Divisão de treinamento/teste para o Conjunto de Dados 1 (São Paulo). Os dados compreendem contagens de tráfego baseadas em radar de alta frequência, amostradas em intervalos de menos de um minuto. O conjunto de treinamento abrange aproximadamente 70% da sequência total, com os 30% restantes alocados para testes. ....	48
<b>Figura 16</b>	– Divisão de treinamento/teste para o conjunto de dados <b>Dataset 2</b> (Brasília). Os dados consistem em contagens de veículos agregadas a cada 15 minutos ao longo de quatro dias (fevereiro de 2025). Os três primeiros dias formam o conjunto de treinamento, enquanto o último dia é reservado para testes.....	49
<b>Figura 17</b>	– Em a) Comparação visual entre os dados originais (em azul) e os dados semissintéticos gerados via Poisson (em vermelho) para o Radar 10433. E a sobreposição das séries temporais. A seguir, b) a comparação das suas respectivas distribuições de probabilidade.....	53
<b>Figura 18</b>	– Dados de Velocidade: Tráfego Geral e Volume: Pontos de Contagem Volumétrica.....	55
<b>Figura 19</b>	– Rota 04G - Av. Zaki Narchi, Av. Santos Dumont, Av. Tiradentes, Av. Prestes Maia, Fonte: [3] .....	55
<b>Figura 20</b>	– Comparação entre a geração de dados via processo de Poisson e a média original observada em três experimentos distintos, nos períodos da manhã (07:00–09:00, à esquerda) e da tarde (17:00–19:00, à direita). Cada linha representa a evolução temporal dos valores para cada experimento, permitindo observar a proximidade entre a simulação (Poisson) e os dados reais (Média Original) durante os dois intervalos analisados.....	56
<b>Figura 21</b>	– Obtenção dos parâmetros $p$ e $q$ no modelo ARIMA(4,0,3) para amostras seguindo Distribuição Gaussiana. MatLab R2024a.....	57

<b>Figura 22</b> – Estrutura em árvore típica da rede rodoviária para fluxo de tráfego. Os pontos representam locais de medição e setas direção do fluxo. $N = 6$ sites. Sistema referente à Equação (3.11).....	61
<b>Figura 23</b> – Estrutura em árvore típica da rede rodoviária para fluxo de tráfego. Os pontos representam locais de medições e setas direção do fluxo ( $N = 13$ sites). Sistema referente à Equação (3.8). Para nós primários (S1-S5, S9 e S11), tráfego é gerado aleatoriamente, com distribuição Poisson de parâmetro $\lambda$ [veículos/unid tempo] .....	62
<b>Figura 24</b> – Séries temporais geradas por uma distribuição Poisson para cinco sites (S1 a S5) e duas séries adicionais, $\mathbf{z}^{(1)}$ e $\mathbf{z}^{(2)}$ no MatLab R2024a.....	63
<b>Figura 25</b> – Obtenção dos parâmetros do modelo STARIMA(2,1,2). MatLab R2024a.....	64
<b>Figura 26</b> – Séries temporais (3k amostras) geradas por distribuições de Poisson com diferentes $\lambda$ 's para os cinco sites de entrada (S1 a S5) e três sites de tráfego confluentes intermediários S6 e S7 ( $\mathbf{z}^{(1)}$ ) e S8 ( $\mathbf{z}^{(2)}$ ), $\mathbf{z}^{(1)}$ e $\mathbf{z}^{(2)}$ . MatLab R2024a.....	64
<b>Figura 27</b> – Obtenção dos parâmetros do modelo STARMA(2,0,2). MatLab R2024a.....	65
<b>Figura 28</b> – Obtenção dos parâmetros do modelo STARIMA(2,1,2). MatLab R2024a.....	66
<b>Figura 29</b> – Série Temporal Geradas por Distribuição de Poisson.....	67
<b>Figura 30</b> – Predição da Série Temporal.....	68
<b>Figura 31</b> – NMSE e EC para diferentes ordens .....	69
<b>Figura 32</b> – Predição da Série Temporal.....	69
<b>Figura 33</b> – NMSE e EC para diferentes ordens .....	70
<b>Figura 34</b> – Av 23 de Maio - SP (Fonte: Google Maps).....	73
<b>Figura 35</b> – Radares na Av 23 de Maio - SP (Fonte: [4] ) .....	73
<b>Figura 36</b> – Distribuição dos radares na Av 23 de Maio por sentido (AS) - (Fonte:Autor).....	74
<b>Figura 37</b> – Fluxograma da construção do modelo .....	74

<b>Figura 38</b> – (a) Série Temporal do dados do CET radar L0; (b) Autocorrelação e Autocorrelação Parcial .....	76
<b>Figura 39</b> – Rota 04G - Av. Zaki Narchi, Av. Santos Dumont, Av. Tiradentes, Av. Prestes Maia, Fonte: [3] .....	77
<b>Figura 40</b> – Dados expandidos (cada 3 min.) vs. Taxa Média Original Horária .....	78
<b>Figura 41</b> – Dados Treino e Teste .....	79
<b>Figura 42</b> – Função STACF e STPACF .....	79
<b>Figura 43</b> – Treino, Previsão <i>Step-by-Step</i> e Dados Originais .....	80
<b>Figura 44</b> – Treino, Observações Alinhadas e Previsões .....	80
<b>Figura 45</b> – Dados Expandidos vs. Taxa Média Original Horária - Período da manhã .....	81
<b>Figura 46</b> – Dados de Treino e Teste - Período da manhã .....	81
<b>Figura 47</b> – STACF / STPACF Treino e Teste - Período da manhã .....	82
<b>Figura 48</b> – Treino, Previsão e Dados Originais - Período da manhã .....	83
<b>Figura 49</b> – Treino, Observações e Previsões - Período da manhã .....	84
<b>Figura 50</b> – Resultados em LSTM para 3 radares .....	85
<b>Figura 51</b> – Resultados em LSTM para 3 radares - Período da manhã .....	85
<b>Figura 52</b> – Resultados em LSTM para 3 radares - Período da tarde .....	86
<b>Figura 53</b> – Resultados em LSTM para 4 radares .....	87
<b>Figura 54</b> – Comparação entre valores reais e previstos do número de veículos para os radares 1, 2 e 3. 88	
<b>Figura 55</b> – Comparação entre os valores reais e previstos do número de veículos para os radares 1, 2 e 3. ....	88
<b>Figura 56</b> – Comparação entre dados reais e previsões do número de veículos nos radares 1, 2 e 3, entre 08:00 e 09:00. ....	89
<b>Figura 57</b> – Comparação entre dados reais e previsões do número de veículos nos radares 1, 2 e 3, entre 18:00 e 19:00. ....	89
<b>Figura 58</b> – Previsões dos modelos a) STARIMA e b) LSTM no conjunto de dados 2 nos três segmentos de estrada com radares. ....	91
<b>Figura 59</b> – Previsões dos modelos STARIMA e LSTM nos segmentos do radar de Brasília, comparando dados de treinamento de 2 e 4 dias. Os painéis a) e b) mostram os resultados do STARIMA, enquanto os painéis c) e d) mostram os resultados do LSTM. ....	92

## LISTA DE TABELAS

<b>Tabela 1</b>	– Modelos de predição de tráfego .....	22
<b>Tabela 2</b>	– Vizinhos de primeira e segunda ordem para o sítio 0 .....	27
<b>Tabela 3</b>	– Comparação dos coeficientes do modelo original com os ajustados e respectivos NMSE .....	57
<b>Tabela 4</b>	– Comparação dos coeficientes adotados no modelo original com os coeficientes estimados e respectivos NMSE relativos à série temporal da Figura 25 .....	62
<b>Tabela 5</b>	– Comparação dos coeficientes adotados no modelo original com os coeficientes estimados e respectivos NMSE relativos à série temporal da Figura 28 .....	67
<b>Tabela 6</b>	– Comparação de desempenho entre STARIMA e LSTM nos períodos da manhã e da tarde .....	90
<b>Tabela 7</b>	– Comparação de desempenho entre os modelos STARIMA e LSTM nas três séries de radares de Brasília .....	91
<b>Tabela 8</b>	– Comparação de desempenho entre os modelos STARIMA e LSTM nas três séries de radares de Brasília; Resultados de treinamento de 2 dias vs. treinamento de 4 dias .....	93

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	14
<b>1.1</b>	<b>Objetivos</b> .....	14
<b>1.2</b>	<b>Metodologia e contribuições</b> .....	15
<b>1.3</b>	<b>Estrutura</b> .....	15
<b>2</b>	<b>TRABALHOS RELACIONADOS</b> .....	17
<b>2.1</b>	<b>Congestionamento, monitoramento, predição e controle de tráfego em cidades inteligentes</b> .....	17
<b>2.2</b>	<b>Modelagem, monitoramento e gerenciamento de tráfego urbano inteligente</b> .....	18
2.2.1	Sistemas de transportes inteligentes em ambientes de cidade inteligentes .....	18
2.2.2	Sistema de controle de tráfego com estrutura adaptativa (ATCS) .....	19
2.2.3	Modelos de previsão de tráfego para cidades inteligentes.....	21
2.2.4	Predição do tempo de viagem.....	21
<b>2.3</b>	<b>Formulação do problema de predição do tráfego de veículos</b> .....	22
<b>2.4</b>	<b>Séries temporais: base para a modelagem do tráfego urbano</b> .....	23
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b> .....	24
<b>3.1</b>	<b>Modelos paramétricos baseados em ARMA espaço-temporais</b> .....	24
3.1.1	Modelo ARIMA.....	25
<b>3.2</b>	<b>STARIMA: multi-espacial ARIMA</b> .....	26
3.2.1	Vizinhança.....	27
3.2.2	Equações espaço-temporais para o modelo paramétrico STARIMA .....	27
3.2.3	Autocovariância e autocorrelação .....	29
<b>3.3</b>	<b>Variações e melhorias no STARIMA</b> .....	29
3.3.1	$\Delta$ -STARIMA: STARIMA com atrasos variáveis no tempo.....	29

3.3.2	STARIMA dinâmico.....	30
3.3.3	STARIMA aprimorado .....	30
<b>3.4</b>	<b>STARIMA para previsão de fluxo de tráfego de curto prazo .....</b>	<b>30</b>
3.4.1	Efeito da segmentação de séries temporais no modelo STARIMA .....	31
<b>3.5</b>	<b>Redes neurais recorrentes (RNN) e redes com memória de curto e longo prazos (LSTM).....</b>	<b>32</b>
3.5.1	Portas de entrada.....	34
3.5.2	Portas de esquecimento.....	35
3.5.3	Células de memória .....	35
3.5.4	Portas de saída .....	36
3.5.5	Saída da célula .....	36
3.5.6	Sumário das características do modelo aprendizagem de máquina LSTM.....	36
<b>3.6</b>	<b>Processos de Poisson .....</b>	<b>36</b>
3.6.1	Emulando o tráfego de veículos baseado em distribuições de Poisson e Beta .....	37
<b>4</b>	<b>MATERIAIS E MÉTODOS .....</b>	<b>40</b>
<b>4.1</b>	<b>Metodologia para o ajuste de parâmetros do modelo paramétrico.....</b>	<b>40</b>
4.1.1	Número de parâmetros ajustáveis do modelo .....	40
4.1.2	Procedimento de ajuste dos parâmetros.....	40
<b>4.2</b>	<b>Associação da malha viária às equações espaço-temporais STARIMA .....</b>	<b>41</b>
4.2.1	Resultados das simulações para a equação (3.8) .....	45
<b>4.3</b>	<b>Coleta de dados em vias reais .....</b>	<b>46</b>
4.3.1	Treinamento e teste em modelos paramétricos e de aprendizagem .....	48
<b>4.4</b>	<b>Métricas de desempenho .....</b>	<b>49</b>
4.4.1	Erro quadrático médio normalizado (NMSE).....	49
4.4.2	Raiz do erro quadrático médio (RMSE).....	49
4.4.3	Erro médio absoluto (MAE) .....	50
4.4.4	Erro absoluto percentual médio (MAPE).....	50

4.4.5	Coeficiente de eficiência .....	50
<b>5</b>	<b>RESULTADOS E DISCUSSÃO .....</b>	<b>52</b>
<b>5.1</b>	<b>Geração de dados semissintéticos .....</b>	<b>52</b>
5.1.1	Exemplo de validação .....	53
5.1.2	Estudo de caso: adensamento de dados da CET .....	54
<b>5.2</b>	<b>Conjunto de dados hipotéticos gerados a partir de modelos ARMA (p,q) e extração dos parâmetros com ARIMA (p,D,q) .....</b>	<b>56</b>
5.2.1	Determinação da ordem AR(p), MA(q), de integração (D) no modelo STARIMA.....	58
<b>5.3</b>	<b>Rede rodoviária típica via simulação Matlab modelo STARIMA.....</b>	<b>60</b>
<b>5.4</b>	<b>Emulando o tráfego urbano na Av. 23 de Maio - São Paulo, BR.....</b>	<b>70</b>
5.4.1	Métodos .....	71
5.4.1.1	Método – volume.....	71
5.4.1.2	Método - velocidade e tempo de retardamento .....	72
5.4.2	Descrição da base experimental e métodos de identificação aplicados.....	72
5.4.3	Identificação das ordens p, d e q .....	74
<b>5.5</b>	<b>Monitoração da fluidez do sistema viário da Av. Tiradentes - São Paulo, BR .....</b>	<b>77</b>
<b>5.6</b>	<b>Validação do modelo LSTM.....</b>	<b>87</b>
5.6.1	Treinamento .....	87
5.6.2	Resultados numéricos.....	88
<b>5.7</b>	<b>Predição de tráfego para o conjunto de dados 2 (Brasília) .....</b>	<b>91</b>
<b>6</b>	<b>CONCLUSÕES E PERSPECTIVAS.....</b>	<b>94</b>
<b>6.1</b>	<b>Trabalhos futuros.....</b>	<b>95</b>
6.1.1	Modelos híbridos e ensembles.....	95
6.1.2	Integração de fontes de dados heterogêneas .....	95
6.1.3	Aplicações em tempo real e computação de borda .....	95
6.1.4	Expansão para outros desafios de mobilidade urbana.....	95
6.1.5	Interpretabilidade e explicabilidade (XAI).....	95

6.1.6	Otimização de parâmetros e arquiteturas de modelos .....	96
6.1.7	Análise de cenários e simulações preditivas .....	96
6.1.8	Impacto de eventos extremos e anomalias .....	96
<b>REFERÊNCIAS</b> .....		<b>97</b>

# 1 INTRODUÇÃO

Na engenharia, a modelagem preditiva é uma prática que envolve a construção e o uso de modelos matemáticos e computacionais para prever o comportamento futuro do tráfego de veículos. Esses modelos são desenvolvidos a partir de dados históricos e variáveis relevantes, permitindo prever o fluxo do trânsito [5–8].

O tráfego urbano, impulsionado pelo crescimento das cidades e da frota de veículos, apresenta desafios complexos para a gestão e o planejamento urbano. A modelagem preditiva surge como uma ferramenta para otimizar o fluxo de veículos, contribuindo para a eficiência energética, a sustentabilidade e a melhoria da qualidade de vida nas cidades inteligentes [6, 8].

Na gestão de cidades inteligentes, a modelagem preditiva se destaca como uma importante ferramenta na otimização e controle de veículos. Este tipo de modelagem permite a análise de dados históricos de tráfego, como velocidade, fluxo e densidade de veículos, juntamente com informações contextuais, como eventos, condições climáticas e obras em andamento. Os modelos preditivos podem antecipar congestionamentos, incidentes e padrões de deslocamento. Essa capacidade de previsão possibilita que gestores de tráfego tomem decisões proativas, ajustando os tempos dos semáforos em tempo real, otimizando rotas de transportes públicos e alertando os motoristas sobre possíveis atrasos, permitindo que rotas alternativas sejam usadas [5–8].

Além disso, a análise preditiva ajuda as cidades a filtrar e traduzir dados em informações acionáveis, que tornam os espaços urbanos mais produtivos e otimizados. Não só a gestão reativa do tráfego é beneficiada; a modelagem preditiva contribui significativamente para o planejamento urbano de longo prazo. Ao prever a demanda futura por diferentes modais de transporte e identificar áreas que necessitam de melhorias na infraestrutura, as cidades podem investir em soluções de mobilidade mais eficientes, como transportes públicos, ampliando ciclovias e outras medidas para mitigar congestionamentos [5–8].

Sendo assim, essa abordagem proativa é importante para o desenvolvimento sustentável das cidades, permitindo que elas se adaptem ao crescimento populacional e às mudanças nos padrões de deslocamento. A modelagem matemática da dinâmica veicular é relevante para o entendimento da dinâmica e para a análise do desempenho dos sistemas de controle implementados, corroborando a importância da modelagem preditiva no controle eficaz de veículos [9].

Diante desse cenário, esta dissertação se propõe a comparar o desempenho do modelo STARIMA (*Space-Time Autoregressive Integrated Moving Average*) e LSTM (*Long Short-Term Memory*) para previsões do tráfego urbano veicular de curto e longo prazo [5–8].

## 1.1 Objetivos

O presente estudo tem como finalidade a elaboração de previsões em horizonte de curto prazo referentes ao volume de tráfego em uma rua ou via, com o intuito de subsidiar processos decisórios voltados à gestão do tráfego urbano em cidades metropolitanas. Particularmente, busca-se avaliar a acurácia e a aplicabilidade do modelo STARIMA como ferramenta preditiva de curto e longo prazo. Para a condução das análises, serão utilizados dados empíricos provenientes de radares instalados na região metropolitana de São Paulo (Av. 23 Maio) ao longo do ano de 2018. O modelo será calibrado a partir de uma amostra desses dados, correspondente a um intervalo temporal a ser determinado. O modelo realizará estimativas do fluxo viário futuro, as quais serão confrontadas com os dados observados no mesmo período, conforme registrado pelos

radares. Inicialmente, a análise contemplará dados coletados em diversos pontos do corredor Norte-Sul, abrangendo as avenidas 23 de Maio e 9 de Julho.

No entanto, fica o questionamento, o modelo STARIMA é acurado e aplicável para a previsão de curto e longo prazo do volume de tráfego urbano em vias de cidades metropolitanas?

Especificamente, pode-se elencar os seguintes procedimentos para se atingir os objetivos e metas propostos e mitigar o principal questionamento sugerido:

1. Simular o tráfego de veículos;
2. Realizar o ajuste nos parâmetros  $\phi$  e  $\theta$  associados ao modelo paramétrico STARIMA;
3. Propor melhorias e adaptações ao modelo STARIMA para otimizar a sua precisão e aplicabilidade na previsão do tráfego urbano de curto ou longo prazo;
4. Comparar o desempenho e a complexidade da metodologia STARIMA versus métodos não paramétricos, por exemplo, redes neurais LSTM na predição do tráfego de curto e longo prazo.

## 1.2 Metodologia e contribuições

A metodologia científica adotada para a realização da pesquisa consiste nos procedimentos de atualização do conhecimento, envolvendo revisão bibliográfica continuada, aquisição dos conceitos necessários ao desenvolvimento do plano de pesquisa proposto, além do estudo e domínio de técnicas e procedimentos adequados para se atingir soluções promissoras de problemas do congestionamento do tráfego de veículos nas cidades metropolitanas e possíveis adaptações em novos problemas em Cidades Inteligentes. Utilizou-se o simulador matemático Matlab.

As principais contribuições deste trabalho são elencadas a seguir.

- Avaliação comparativa dos modelos STARIMA e LSTM: O estudo apresenta uma comparação empírica entre o modelo STARIMA e a rede neural LSTM. Enquanto o STARIMA captura efetivamente dependências espaço-temporais por meio de uma abordagem paramétrica estruturada, o LSTM oferece uma capacidade de modelagem flexível e não linear que pode aprender padrões temporais complexos a partir dos dados;
- Uso de conjuntos de dados do mundo real: Os experimentos são realizados usando dois conjuntos de dados do mundo real, permitindo uma avaliação mais realista e prática do desempenho de previsão dos modelos em cenários reais de monitoramento de tráfego;
- Avaliação da precisão da previsão e viabilidade prática: Os modelos são avaliados em termos de análises de erro, proporcionando uma comparação quantificável do desempenho preditivo;
- Discussão sobre a transferibilidade e implementação do modelo.

## 1.3 Estrutura

O presente trabalho está organizado em seis capítulos. O Capítulo 1 apresenta a introdução, explicitando os objetivos, a metodologia empregada e a estrutura do estudo. No Capítulo 2 são discutidos os trabalhos relacionados, abordando pesquisas sobre congestionamento, monitoramento, previsão e controle de tráfego em cidades inteligentes, bem como a formulação do problema e o papel das séries temporais. O Capítulo 3 trata da fundamentação teórica, contemplando modelos paramétricos, o modelo STARIMA e suas variações,

redes neurais LSTM e processos estocásticos como Poisson e Beta. No Capítulo 4 é descrita a metodologia adotada, destacando o ajuste de parâmetros e as métricas de desempenho utilizadas. O Capítulo 5 apresenta os resultados e discussões, incluindo a geração de dados, a aplicação dos modelos ARIMA, STARIMA e LSTM, além da simulação de cenários reais de tráfego. Por fim, o Capítulo 6 traz as conclusões e perspectivas, indicando possíveis direções para trabalhos futuros e destacando as contribuições do estudo para a área de mobilidade urbana inteligente.

## 2 TRABALHOS RELACIONADOS

A previsão de tráfego é um campo de estudo necessário para o desenvolvimento de sistemas de transporte inteligentes (*ITS - Intelligent Transportation System*) e cidades inteligentes. O objetivo é antecipar as condições de tráfego futuro para otimizar o fluxo, reduzir o congestionamento e melhorar a eficiência do transporte. Métodos tradicionais de previsão de tráfego incluem modelos estatísticos, como séries temporais, enquanto abordagens mais recentes utilizam técnicas de aprendizado de máquinas e inteligência artificial.

Os sistemas destinados à estimativa e à previsão de tráfego podem ser organizados conforme o período de antecipação considerado. Aqueles voltados ao curto prazo abrangem intervalos que variam de menos de um minuto até algumas horas à frente do tempo atual de processamento, enquanto os de longo prazo podem alcançar projeções para vários dias. A previsão em curto prazo, realizada próxima ao tempo real, revela-se uma estratégia viável para a gestão integrada do tráfego, sendo compatível com diversas tecnologias de ITS. Essa abordagem pode ser aplicada em Centros de Controle Operacional para auxiliar na tomada de decisões rápidas, permitindo a comparação entre diferentes alternativas operacionais diante de eventos em curso [5–8].

Os modelos estatísticos de previsão de tráfego como *ARIMA* e suas variações (*SARIMA*, *STARIMA*) têm sido amplamente utilizados devido à sua capacidade de capturar padrões temporais e espaciais no tráfego. Para demonstrar a importância de considerar a dependência espaço-temporal, em [10] aplicaram modelos *STARIMA* para prever o fluxo de tráfego em redes urbanas. As técnicas de aprendizado de máquina como redes neurais (RNNs, LSTMs), máquinas de vetores de suporte (SVMs) e árvores de decisão têm ganhado popularidade devido à sua capacidade de lidar com dados não lineares e complexos, como exemplo, o trabalho em [11] realizara uma revisão comparativa de modelos de previsão de tráfego, destacando a eficácia das redes neurais em cenários de tráfego dinâmicos.

Outros modelos híbridos consistem da combinação de modelos estatísticos e de aprendizado de máquina tem se mostrado promissora para melhorar a precisão da previsão.

### 2.1 Congestionamento, Monitoramento, Predição e Controle de Tráfego em Cidades Inteligentes

Sistemas de monitoramento e gerenciamento de tráfego inteligente em cidades inteligentes é um tópico de pesquisa ativo devido às tecnologias emergentes, tais como a Internet de Coisas (IoT) e Inteligência Artificial (IA) e as tecnologias habilitadoras no contexto de sistemas 5G. Métodos de previsão de tráfego existentes são dedicados principalmente ao gerenciamento de tráfego rodoviário e urbano, e estudos limitados focados em estradas coletoras e campi fechados. Em [12] os autores propõem um modelo de sistema baseado em IoT para coletar, processar e armazenar dados de tráfego para tal cenário. O objetivo foi fornecer atualizações de tráfego em tempo real sobre congestionamento de tráfego e incidentes de trânsito incomuns através de unidades de mensagens na estrada e, assim, melhorar a mobilidade. Mensagens de alerta ajudam os cidadãos a economizar tempo, especialmente nos horários de pico. Também, o sistema transmite as atualizações de tráfego das autoridades administrativas. Um protótipo foi implementado para avaliar a viabilidade do modelo, e os resultados dos experimentos mostram boa exatidão no veículo de testes e um baixo erro relativo na estimativa de ocupação de estradas. O estudo faz parte de um programa financiado em Omã, investigando o feedback em tempo real para sinais de trânsito adaptativos

## 2.2 Modelagem, Monitoramento e Gerenciamento de Tráfego Urbano Inteligente

Com os avanços tecnológicos e o contexto das Cidades Inteligentes, é possível desenvolver sistemas de monitoramento e gerenciamento de tráfego capazes de analisar o tráfego e o fluxo de transporte, bem como tomar decisões em tempo real. Esses sistemas ajudam os usuários das vias a fazerem escolhas mais informadas sobre suas viagens, melhorando a qualidade do trânsito, reduzindo emissões e resolvendo problemas de congestionamento. A previsão de congestionamentos de tráfego visa fornecer informações sobre as condições de trafegabilidade das vias com antecedência, aumentando assim a eficácia do planejamento urbano.

Os sistemas ITS amplamente implementados na Europa e nos EUA, utilizam sensores infravermelhos, sistemas embarcados com microcontroladores e módulos Wi-Fi integrado a um aplicativo Android para prever e notificar sobre congestionamentos ao longo de rotas específicas, Estes aplicativos monitoram e alertam os usuários em tempo real sobre as condições de tráfego, incluindo rotas alternativas [13]. Ademais, os sistemas ITS têm melhorado significativamente a eficiência do monitoramento e a disponibilidade de dados sobre as condições das vias em tempo real. Em uma Cidade Inteligente, as estradas são equipadas com infraestrutura para análise de fluxo de tráfego, permitindo uma conectividade mais rápida. Cidades em todo o mundo enfrentam problemas significativos de congestionamento, resultando em desperdício de combustível, aumento de estresse, atrasos em entregas e riscos financeiros [13].

Em países em desenvolvimento, como a Índia, o congestionamento é uma questão crítica. As rodovias nacionais, que cobrem uma vasta extensão do território, enfrentam desafios devido ao alto número de veículos, impulsionado pelo desenvolvimento econômico e urbano [13].

Construir uma infraestrutura inteligente usando IoT e modelos de previsão pode minimizar esses desafios. Estudos propuseram soluções, como a análise de tráfego usando o algoritmo KNN, que seleciona dados pertinentes para prever congestionamentos em rotas específicas, proporcionando notificações preditivas aos usuários. Essa solução demonstrou ser precisa na previsão de fluxos de tráfego, eficaz e necessária em Cidades Inteligentes conectadas, onde sensores coletam dados de tráfego em tempo integral, fornecendo previsões de congestionamento e rotas alternativas para otimizar o fluxo urbano [13].

Com o crescimento populacional, o número de veículos nas estradas aumenta diariamente, agravando os problemas de congestionamento. Sistemas de previsão automática e controle de sinais de trânsito, como o STMS habilitado para IoT, propõem soluções inovadoras. Nesses sistemas, veículos de emergência com etiquetas RFID têm prioridade nos semáforos, agilizando seu deslocamento [14].

Técnicas tradicionais, baseadas em cronogramas fixos para sinais de trânsito, são ineficazes diante de padrões de tráfego dinâmicos. Sistemas de controle de tráfego adaptativos oferecem uma solução mais responsiva, ajustando-se às condições de tráfego em tempo real e melhorando a eficácia na tomada de decisões [14].

### 2.2.1 Sistemas de Transportes Inteligentes em Ambientes de Cidade Inteligentes

A Arquitetura do ITS oferece uma estrutura para qualquer forma de iniciativa relacionada à tecnologia no setor de transporte. A arquitetura do sistema é uma descrição ampla das instalações, módulos ITS, interconexões e mapeamento do fluxo de informações para os diferentes sistemas e subsistemas que abrangem o projeto ITS. Com uma infraestrutura devidamente projetada, proprietários e interessados podem definir tanto os serviços prestados pelos usuários finais, quanto os dados para esses serviços e a interconexão entre os vários subsistemas [1]. A preparação e implementação do ITS, por exemplo, já ocorreu na Índia. Onde existem implementados vários itens de infraestrutura para o controle de tráfego, incluindo: tecnologia de transporte, controle de tráfego, cobrança de tarifas e sistemas de cobrança de pedágio [1].

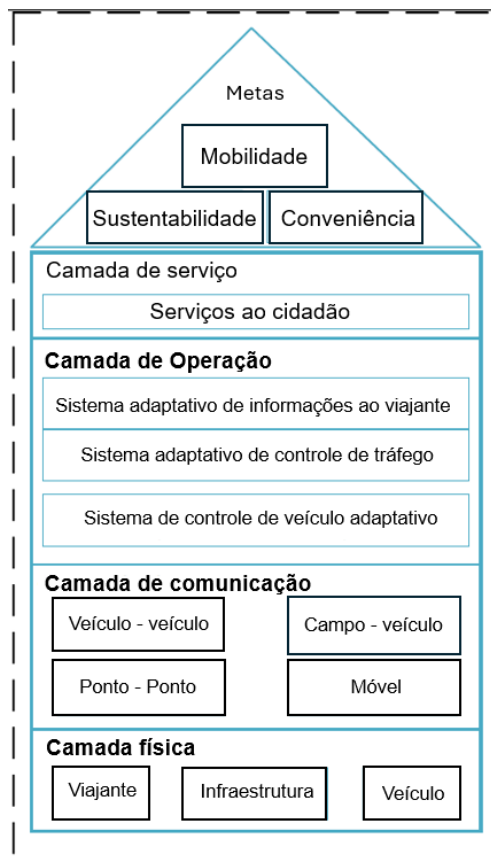


Figura 1 – Camadas do Sistema de Transporte Inteligente (adaptado de [1]).

A arquitetura ITS envolve muitos serviços de usuários, como serviços de informação ao viajante, serviços de gerenciamento de tráfego, serviços de transporte público, serviço de pagamento eletrônico de pedágio, operações de veículos comerciais, serviços de gerenciamento de emergência, segurança de veículos e sistema de controle e serviços de armazenamento de informações, como podemos ver na Figura 1. Nas cidades metropolitanas (por exemplo, na Índia) o sistema de controle de tráfego existente é ineficiente devido à aleatoriedade no padrão e densidade de tráfego durante o dia. Por isso, embora o nível de tráfego seja muito menor nas cidades menores, os veículos devem esperar muito tempo. Assim, o problema de congestionamento de trânsito pode ser reduzido se o temporizador do sinal de tráfego configurado puder ser manipulado com uma densidade de tráfego variável ao invés de constante [1]. Na Figura 2 apresenta-se a Arquitetura do Sistema de Transporte Inteligente.

### 2.2.2 Sistema de Controle de Tráfego com Estrutura Adaptativa (ATCS)

O algoritmo ATCS altera continuamente os tempos dos semáforos com base na demanda de tráfego nas interseções e nas chegadas projetadas das interseções vizinhas. Embora a estratégia resulte em aumento do tempo de viagem, empurrando os veículos de forma constante pelas luzes verdes, reduzindo o congestionamento, o sistema de controle adaptativo ATCS permite obter um fluxo mais suave [1], (ver Figura 3).

Utilizando processamento de imagens, o sistema proposto em [1] adapta o temporizador do semáforo de acordo com a densidade de tráfego aleatório. Este modelo usa câmera de alta resolução para detectar mudanças nos padrões de tráfego ao redor do sinal de tráfego e manipular o temporizador do sinal de acordo, enviando sinais para o sistema de controle do temporizador.

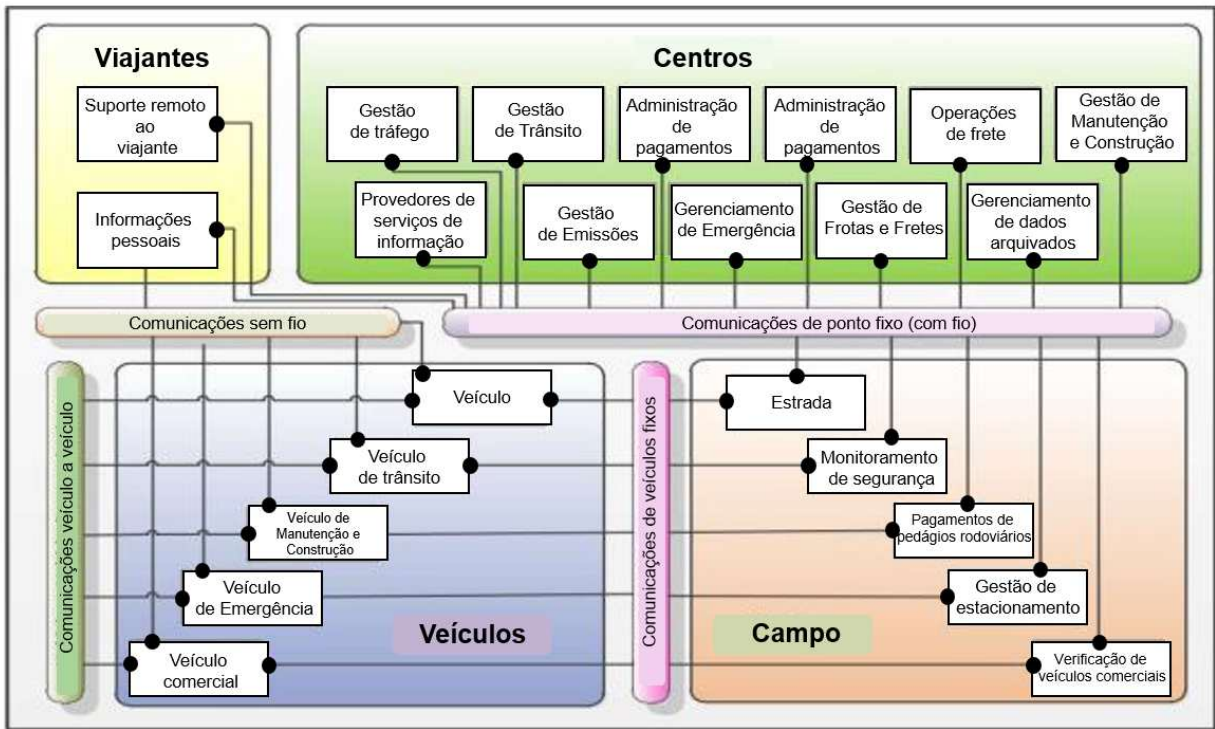


Figura 2 – Arquitetura do Sistema de Transporte Inteligente (adaptado de [1])

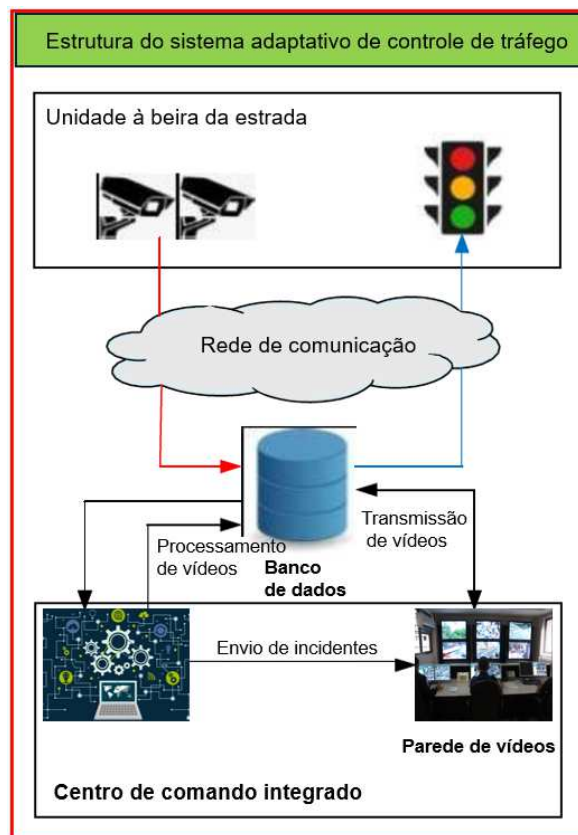


Figura 3 – Visão geral da estrutura ATCS (adaptado de [1])

O congestionamento do tráfego depende da regulação do fluxo e, por conseguinte, do temporizador do semáforo. O resultado disso, é que os veículos tem que enfrentar uma pausa irregular durante as viagens nas áreas metropolitanas. Em [1], propõe adicionar uma rede de sensoriamento que fornece entrada para a rede existente para que ela possa responder aos padrões em evolução de densidade de tráfego e fornecer ao controlador os sinais necessários em tempo real, otimizando assim os atrasos no trânsito de veículos em horários irregulares do dia.

### 2.2.3 Modelos de Previsão de Tráfego para Cidades Inteligentes

A previsão precisa das condições de tráfego em cidades inteligentes foi facilitada pelo rápido desenvolvimento do aprendizado de máquina e pelo surgimento de novas fontes de dados. Isto pode ajudar a otimizar a concepção e a gestão dos serviços de transporte numa futura cidade automatizada. Uma discussão detalhada dos métodos de previsão de tráfego para essas cidades inteligentes é fornecida em [15], fornecendo também uma visão geral das fontes de dados e modelos de previsão existentes.

A abordagem empírica de previsão de tráfego frequentemente usada na literatura, bem como modelos analítico-paramétricos são apresentados na Tabela 1. Destaca-se as vantagens e desvantagens dos modelos, que podem ser categorizados em três categorias principais: modelos ingênuos, paramétricos e não paramétricos. Uma comparação analítica completa dos modelos existentes não é fácil, porque o desempenho de um determinado modelo de previsão é altamente dependente do cenário exato, das fontes de dados disponíveis e dos parâmetros ambientais. Portanto, alguns dos modelos superariam outros num determinado cenário; mas num cenário diferente, o seu desempenho deteriorar-se-ia. Portanto, o modelo apropriado deve ser escolhido examinando previamente o cenário dado e compreendendo o efeito de vários parâmetros.

Na abordagem proposta em [16] os autores exploram uma solução adaptativa para os tempos de abertura e fechamento dos semáforos, os quais se adaptam ao fluxo de tráfego em tempo real. No entanto, não expõem detalhes e melhorias da tecnologia associada a antenas inteligentes. O trabalho parte do pressuposto de que a densidade de celulares corresponde à densidade de veículos nas rodovias. O Notificador Integrado ao TDC (Traffic-light Duration Controller) alerta o motorista sobre a transição do vermelho para o verde. Melhoria do tempo médio de espera (AWT - Average Waiting Time) em 59,3%. Economia de 3,42 galões de combustíveis e reduz a emissão de CO<sub>2</sub> em 67 libras por ano para cada passageiro. No entanto, essas soluções são complicadas por que envolvem comunicação direta entre veículos e semáforos. Além de serem relativamente dispendiosas uma vez que exigem componentes dedicados, como bobinas de indução, câmeras, sensores, fibras óticas e computadores, que também possuem alto custo de instalação/manutenção [16].

### 2.2.4 Predição do Tempo de Viagem

Um dos principais desafios nos sistemas de transporte inteligente é prever várias condições de tráfego, como fluxo de tráfego, velocidade e tempo de viagem; entre estes, o tempo de viagem é considerado o mais importante, pois tem implicações diretas para os viajantes individuais e para as autoridades de rede que gerenciam as redes rodoviárias. Fornecer informações sobre o tempo de viagem a todos os usuários facilita decisões mais informadas sobre escolhas de rota, reduzindo assim o tempo de viagem total na rede rodoviária. Da mesma forma, o conhecimento dos tempos de viagem e dos volumes de tráfego pode ajudar significativamente na provisão de infraestrutura (como pontos de acesso na beira da estrada) para fornecer serviços de comunicação por meio de redes ad hoc veiculares.

Prever o tempo de viagem é difícil, pois as condições de tráfego podem variar amplamente nos domínios espacial e temporal. Diferentes modelos de previsão de tempo de viagem têm sido propostos dependendo da complexidade e do tamanho da rede rodoviária e do tamanho da janela de previsão. Em uma extremidade da

Tabela 1 – Modelos de predição de tráfego

Modelo	Tipo	Dados	Integração de Dados	Propriedades Espaciais	Não Linear	Não Estacionário
ITT <sup>b</sup>	Simplista	Escalar				
Média Histórica	Simplista	Escalar				
ARIMA	Paramétrico	Escalar				✓ <sup>a</sup>
SARIMA	Paramétrico	Escalar				✓ <sup>a</sup>
<b>STARIMA</b>	Paramétrico	Escalar		✓		✓ <sup>a</sup>
KARIMA	Paramétrico	Escalar				✓ <sup>a</sup>
ARIMAX	Paramétrico	Escalar	✓	✓		✓ <sup>a</sup>
VARMA	Paramétrico	Vetor		✓		
Kalman Filter	Paramétrico	Escalar				✓
Bayesian Networks	Não Paramétrico	Escalar, Vetor, Matriz		✓	✓	✓
K-Nearest Neighbors	Não Paramétrico	Vector	✓	✓	✓	✓
Feed Forward Neural Network	Não Paramétrico	Escalar	✓	✓	✓	✓
Time Delayed Neural Network	Não Paramétrico	Escalar			✓	✓
Recurrent Neural Network	Não Paramétrico	Escalar			✓	✓
LSTM-RNN <sup>b</sup>	Não Paramétrico	Escalar			✓	✓
Gated Recurrent Unit Neural Network	Não Paramétrico	Escalar			✓	✓
Convolutional Neural Network	Não Paramétrico	Escalar, Vetor, Matriz		✓	✓	✓
Combination of CNN and FFNN	Não Paramétrico	Escalar, Vetor, Matriz	✓	✓	✓	✓
Combination of CNN and LSTM	Não Paramétrico	Escalar, Vetor, Matriz	✓	✓	✓	✓

<sup>a</sup> Requer etapa de diferenciação; tal etapa é necessária para lidar com dados não estacionários.

<sup>b</sup> ITT – Tempos de Viagem Instantâneos.

<sup>c</sup> LSTM-RNN – Long-short Term Recurrent Neural Network.

escala, redes complexas e horizontes de previsão maiores precisam ser considerados para fins de planejamento, enquanto redes mais simples e previsões imediatas de tempo de viagem são adequadas para fins de orientação em tempo real [17]. Uma variedade de técnicas pode ser utilizada para medir tempos de viagem em redes rodoviárias, incluindo detectores de laço, câmeras de processamento de imagem instaladas em interseções, identificação automática de veículos, carros flutuantes, rastreamento de telefones celulares e localização baseada em GPS. Detectores de laço convencionais são a forma mais comumente usada de medição; dados de fluxo, ocupação e tempo de viagem destes detectores para redes de autoestradas estão amplamente disponíveis. Esses dados têm sido usados como fonte em muitos modelos de previsão de tempo e filtragem de Kalman têm sido utilizados para prever fluxo de tráfego em horizontes de tempo muito curtos, mas as previsões de fluxo nem sempre resultam em previsões precisas de tempo de viagem. Portanto, esforços recentes de pesquisa utilizando redes neurais ou modelos estatísticos focam principalmente na previsão direta do tempo de viagem usando dados históricos e atuais de níveis de fluxo e ocupação.

## 2.3 Formulação do Problema de Predição do Tráfego de Veículos

O problema de previsão de tráfego é formulado como uma série temporal no tempo  $t$  representada por  $y_t$ . Nosso objetivo é prever os valores futuros da série

$$\mathbf{y}_{t_0:T} = [y_{t_0}, y_{t_0+1}, \dots, y_{t_0+T}], \quad (2.1)$$

com base somente em seus valores históricos

$$\mathbf{x}_{t_0-c:t_0-1} = [x_{t_0-c}, \dots, x_{t_0-2}, x_{t_0-1}], \quad (2.2)$$

onde  $t_0$  denota o ponto de partida a partir do qual os valores futuros  $y_t, t = t_0, \dots, T$  devem ser previstos.

Diferencia-se o intervalo de tempo histórico  $[t_0 - c, t_0 - 1]$  e o intervalo de previsão  $[t_0, T]$  como o comprimento do contexto e o comprimento da predição, respectivamente. Nossa abordagem se concentra em gerar previsões pontuais para cada passo de tempo no comprimento de previsão, visando obter previsões precisas e confiáveis.

## 2.4 Séries Temporais: Base para a Modelagem do Tráfego Urbano

Uma sequência de observações de uma variável ao longo do tempo é uma série temporal. Ou seja, é um conjunto sucessivo de observações realizadas em intervalos uniformes durante um certo período. Os diversos métodos de previsão existentes, compartilham a característica de que o comportamento do passado observado interfere no comportamento do futuro estimado. Que previsões levam em consideração somente comportamentos prévios, que não são precisos. Portanto, à medida que o horizonte de previsão se estende, existe a possibilidade de que o erro associado aumente. Adicionalmente, a eficácia das técnicas de previsão depende da acessibilidade a dados, bem como das limitações de tempo e recursos disponíveis [4].

As análises de séries temporais baseiam-se em três elementos fundamentais: tendência, sazonalidade e ciclo. A tendência descreve o movimento de longo prazo da série, indicando se há um aumento, uma diminuição ou estabilidade, assim como a rapidez dessas mudanças. A sazonalidade refere-se a variações periódicas que acontecem em intervalos regulares. Por outro lado, os ciclos representam flutuações suaves e repetitivas que se sobrepõem à tendência. A distinção crucial entre ciclos e sazonalidade reside no fato de que a sazonalidade tem um padrão previsível e ocorre em períodos específicos, enquanto os ciclos podem apresentar variações irregulares [4].

Quando se investigam séries temporais, dois são os propósitos principais. O primeiro é a análise e modelagem da série temporal, isto é, a descrição do fenômeno em estudo, a verificação de características significativas e a investigação de possíveis conexões com outras séries temporais. O segundo propósito visa a projeção de valores futuros com base nos dados históricos coletados. As técnicas de previsão variam desde abordagens simples e diretas até métodos mais elaborados e fundamentados. Entre os modelos estatísticos para projeções futuras, [4] menciona:

1. **Modelos Univariados:** baseiam-se em apenas uma série histórica. Alguns deles são:
  - *Decomposição por componentes não observáveis*, modelo mais adotado até 1960;
  - *Modelos automáticos* que integram *modelos de regressão e médias móveis*, com ajuste sazonal e suavização exponencial;
  - *Modelos univariados* [4]. Trata-se de uma classe geral de modelos lineares, mais conhecidos como ARIMA;
2. **Modelos de Função de Transferência:** a série temporal de interesse é explicada não só pelo seu passado histórico, como também por outras séries temporais não correlacionadas entre si.
3. **Modelos Multivariados:** modelam simultaneamente duas ou mais séries temporais sem qualquer exigência em relação à direção da causalidade entre elas [4]."

# 3 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os fundamentos teóricos necessários para o desenvolvimento do trabalho, com ênfase em modelos voltados à previsão de séries temporais e espaço-temporais. São abordadas as bases dos modelos ARIMA e STARIMA, suas variações e melhorias, bem como estratégias para determinar sua ordem e aplicabilidade. Também são introduzidos os conceitos de Redes Neurais Recorrentes, especialmente o modelo LSTM, com destaque para seus mecanismos internos. Por fim, são discutidos os Processos de Poisson e Beta, utilizados na modelagem probabilística do tráfego veicular.

## 3.1 Modelos Paramétricos Baseados em ARMA Espaço-Temporais

Um dos desafios presentes nos sistemas de estimação e previsão de tráfego é a inclusão do aspecto espacial da rede de tráfego nos sistemas de previsão de curto prazo produzindo assim sistemas que tratem simultaneamente os aspectos espaço-temporais.

Incorporar a característica espaço-temporal (ST) em um modelo de Previsão de Tráfego de Curto Prazo (*Short Time Traffic Forecast - STTF*) pode propiciar um maior entendimento da evolução do tráfego na rede, uma vez que possibilita estimar como a mudança de padrões de tráfego, em determinado local, se propaga para o resto da rede viária [18].

Para a inclusão do fator temporal, os modelos de estimação e previsão de curto prazo são organizados e resumidos por intervalos, que podem variar de poucos segundos até alguns minutos, construindo-se séries temporais das características do tráfego. Dentre os modelos utilizados em análises temporais tem-se os modelos Autorregressivos Integrados de Média Móvel (*Autorregressive Integrated Moving Average - ARIMA*) que é um modelo univariado e o tratamento através de Modelos de Regressão para Dados em Painel Longitudinal ou com medidas repetidas [18].

No modelo STARIMA, a incorporação do aspecto espacial na análise de dados é realizada por meio de matrizes de peso  $\mathbf{W}$ , ou pela inclusão de variáveis explicativas em análises multivariadas. Uma revisão científica sobre modelos espaço-temporais ST-STTF identificou 130 publicações entre 1984 e 2016, com um aumento significativo após 2008 [18]. Desses estudos, 45% utilizaram modelos de Aprendizado de Máquina, 37% modelos estatísticos e 18% modelos mistos. Embora os modelos de Aprendizado de Máquina tenham se destacado, não há garantia de que sejam superiores aos modelos estatísticos avançados, como o STARIMA. Outra revisão, com 211 publicações de 1984 a 2018, mostrou um aumento no uso de métodos estatísticos, enquanto os baseados em Redes Neurais diminuíram. [18] destaca que o tema ainda é aberto à pesquisa, optando por comparar o modelo STARIMA, que é univariado, com uma abordagem de Dados em Painel longitudinais e hierarquizados, que é multivariada e se aproxima dos modelos de aprendizado de máquina.

Modelos populares usam uma variedade de métodos estatísticos, incluindo regressões paramétricas lineares/não lineares e não paramétricas, ARIMA, Modelo Autorregressivo Integrado às Médias Móveis Espaço-Temporal (STARIMA) e ATHENA [17]. Entre esses modelos, o STARIMA é especialmente notável por sua aplicabilidade em domínios espaço-temporais e pelo relativamente pequeno número de parâmetros necessários para estimar o modelo.

No contexto da previsão de tráfego urbano e intercity, modelos STARIMA podem ser usados quando há séries temporais espaciais coletadas em um local específico em intervalos de tempo constantes. Eles podem ser usados para modelar os efeitos de choques que são introduzidos em um determinado ponto espacial no sistema. Por exemplo, um acidente em um elo que afeta as condições de tráfego nos elos

vizinhos. Modelos STARIMA podem modelar uma rede inteira com apenas um modelo, enquanto o ARIMA precisa de um modelo separado para cada elo [10].

Embora os modelos STARIMA tenham sido usados para prever uma infinidade de dados multivariados, incluindo taxas de criminalidade, preços de batata, carga de servidores e preços de farinha, eles também foram usados para prever tempo de viagem e fluxo de tráfego até recentemente. Trabalhos relataram uma precisão impressionante em previsões de curto e médio prazo. Em todos os casos, os modelos foram estimados para redes urbanas de grade de Manhattan estritamente pequenas com sinais de trânsito, ou redes de autoestradas. De acordo com [19], a natureza multivariada e a capacidade de capturar o comportamento espaço-temporal tornam os modelos STARIMA a escolha preferida para redes de tráfego urbano e muito promissores para planejamento e alocação de recursos [17].

### 3.1.1 Modelo ARIMA

O processo de média móvel integrada autorregressiva (ARIMA) gera séries não estacionárias que são integradas de ordem  $D$ , denotadas por  $I(D)$ . Um processo  $I(D)$  não-estacionário é aquele que pode ser tornado estacionário tomando-se  $D$  diferenças. Tais processos são frequentemente chamados de processos de diferença estacionária ou raiz unitária.

Uma série temporal que pode ser modelada como um processo ARMA( $p, q$ ) estacionário após ser diferenciada  $D$  vezes é denotada por ARIMA( $p, D, q$ ). A forma do modelo ARIMA( $p, D, q$ ), é descrito pela equação:

$$\Delta^D y_t = c + \phi_1 \Delta^D y_{t-1} + \dots + \phi_p \Delta^D y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \quad (3.1)$$

onde  $\Delta^D y_t$  é a série temporal diferenciada de  $D$ -ésima ordem;  $\varepsilon_t$  representa o processo de inovação não correlacionado com média zero, por exemplo ruído Gaussiano de média zero e variância  $\sigma_\varepsilon^2$ , os símbolos  $\phi_i$  e  $\theta_i$  são os coeficientes dos componentes autoregressivos e de médias móveis do modelo ARIMA( $p, D, q$ ) referentes ao  $i$ -ésimo site.

Na notação do operador lag (defasagem),  $L^\ell y_t = y_{t-\ell}$ , pode-se escrever o modelo ARIMA( $p, D, q$ ) como

$$\phi^*(L)y_t = \phi(L)(1-L)^D y_t = c + \theta(L)\varepsilon_t, \quad (3.2)$$

sendo  $\phi^*(L)$  um operador AR polinomial instável com exatamente  $D$  raízes unitárias. Fatorando esse polinômio como  $\phi(L)(1-L)$ , onde  $\phi(L) = (1 - \phi_1 L - \dots - \phi_p L^p)$  é um polinômio do operador de defasagem AR de grau  $p$  estável (com todas as raízes fora do círculo unitário).

Da mesma forma,  $\theta(L) = (1 + \theta_1 L + \dots + \theta_q L^q)$  é um polinômio de operador de defasagem MA de grau  $q$  inversível (com todas as raízes fora do círculo unitário).

Os sinais dos coeficientes no polinômio do operador de defasagem AR,  $\phi(L)$ , são opostos ao lado direito da Equação (3.1).

**Obs.1:** Para especificar e interpretar coeficientes AR a partir do Toolbox Econometrics do MatLab, refira à Equação (3.1).

**Obs.2** Para uma série temporal estacionária ( $D = 0$ ), a forma da Equação (3.1) reduz-se a:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \varepsilon_t. \quad (3.3)$$

Na seção 4.1.2 apresenta-se a extração dos Parâmetros AR, MA ou ARMA ( $p, q$ ) a partir do modelo ARIMA( $p, D, q$ ) ou ARIMA( $p, 1, q$ ) de predição e síntese, disponível em MatLab.

### 3.2 STARIMA: Multi-espacial ARIMA

Séries espaço-temporais multivariadas podem ser aproximadas por modelos lineares autorregressivos e de média móvel integrados (STARIMA - *Space-Time Auto-Regressive Integrated Moving Average*). O modelo de série multivariada pode ser adequado se o processo em estudo apresentar dependência sistemática entre as observações em cada local e vizinhança. Este fenômeno é conhecido como correlação espacial.

Eles são referidos como modelos de tráfego urbano ST. Pfeifer e Deutsch apresentaram pela primeira vez o STARIMA; ele pode ser implementado com modificações simples de modelos autorregressivos vetoriais. Os modelos STARMA são caracterizados pela dependência linear espaço-temporal defasada [10]. As observações  $\mathbf{z}_i(t)$  da variável aleatória  $\mathbf{Z}_i(t)$ , considerada uma suposição, refletem as condições de tráfego onde há disponibilidade para  $N$  locais espaciais ( $i = 1, \dots, N$ ) e  $T, 2T, 3T, \dots$  períodos de tempo [10].

O modelo STARMA apresenta a estrutura autorregressiva que expressa as observações no tempo  $t$  obtidas na localização espacial  $i$ -ésima, uma combinação linear do passado, onde foram tomadas observações da localização  $i$  e medições de localizações vizinhas. Se as relações para todos os locais no sistema forem válidas, então o processo exibe estacionaridade espacial, que é adequada para essas formas e modelos espaço-tempo [10].

Sendo  $L^{(\ell)}$  o operador de defasagem espacial de ordem espacial  $\ell$ , seja tal que

$$L^{(\ell)} [z_i(t)] = \sum_{j=1}^N w_{ij}^{(\ell)} z_j(t) \quad (3.4)$$

$$\text{com } L^{(0)} [z_i(t)] = z_i(t), \quad i = 1, \dots, N, \quad (3.5)$$

sendo:

- $L^{(\ell)}$  : operador de defasagem espacial de ordem  $\ell$ ,
- $\ell$  :  $\ell$  (lag) é o número de intervalos de tempo tratados pelo modelo,
- $z_i(t)$  :  $i$ -ésima observação espacial da variável aleatória,
- $N$  : localizações espaciais,
- $i$  :  $i$ -ésima localização espacial (sites).

Observe que  $z_i(t)$  é um escalar; é o elemento na posição  $i$  do vetor  $\mathbf{z}(t)$ , i.e.,  $\mathbf{z}(t) = [z_1(t), z_2(t), \dots, z_N(t)]$ . Ainda, os elementos  $w_{ij}^{(\ell)}$  são os pesos normalizados, com a seguinte restrição:

$$\sum_{j=1}^N w_{ij}^{(\ell)} = 1, \quad \text{para } i = 1, \dots, N, \quad (3.6)$$

assim, todos  $i$  e  $w_{ij}^{(\ell)}$  são diferentes de zero apenas se os sites  $i$  e  $j$  forem vizinhos de  $\ell$ -ésima ordem. Portanto, a representação matricial do conjunto de pesos  $w_{ij}^{(\ell)}$  é  $\mathbf{W}^{(\ell)}$ , um  $N \times N$  matriz quadrada com cada linha somando um. E  $\mathbf{z}(t)$  é um vetor coluna de dimensão  $\mathbb{R}^{N \times 1}$ , a coluna  $N \times 1$  vetor  $\mathbf{z}(t)$ , que contém as observações espaciais  $z_i(t)$ ,  $i = 1, \dots, N$  no instante  $t = 0, 1, 2, \dots, M$ , com as seguintes relações válidas [10]:

$$L^{(0)} [\mathbf{z}(t)] = \mathbf{W}^{(0)} \mathbf{z}(t) = \mathbf{I}_N \mathbf{z}(t) \quad (3.7)$$

e

$$L^{(\ell)} [\mathbf{z}(t)] = \mathbf{W}^{(\ell)} \mathbf{z}(t) \quad \text{para } \ell > 0, \quad (3.8)$$

sendo:

- $L^{(\ell)}$  : operador de defasagem espacial de ordem  $\ell$ ,
- $\mathbf{z}(t)$  :  $[z_1 \ z_2 \ \dots \ z_N]^T$ , as observações espaciais de um v.a. no instante  $t = 0, 1, 2, \dots, M$ ,
- $\mathbf{W}$  : a representação matricial do conjunto de pesos  $w_{ij}^{(\ell)}$ ,
- $\mathbf{I}$  : matriz de identidade.

### 3.2.1 Vizinhança

A descrição dos pesos no formato  $w_{ij}^{(\ell)}$  deve refletir o layout da rede rodoviária. Usando  $w_{ij}^{(\ell)}$ , pode-se integrar as distâncias entre locais de medição  $i$  e  $j$ , obstáculos naturais, ou até mesmo a facilidade de acesso do site  $i$  ao site  $j$ . Vizinhos de primeira ordem são aqueles que estão mais próximos do local de interesse. Os vizinhos de segunda ordem devem estar situados mais longe que os vizinhos de primeira ordem, mas mais próximos que os vizinhos de terceira ordem [10], conforme ilustrado na Tabela 2.

Tabela 2 – Vizinhos de primeira e segunda ordem para o sítio 0

$$\begin{array}{ccccc}
 & & & & 2 \\
 & & & & 2 \\
 & & & 1 & 2 \\
 2 & 1 & \mathbf{0} & 1 & 2 \\
 & & & 2 & 1 & 2 \\
 & & & & & 2
 \end{array}$$

### 3.2.2 Equações Espaço-temporais para o Modelo Paramétrico STARIMA

Assim como nas séries temporais univariadas,  $z_i(t)$  é expresso como uma combinação linear de observações e erros passados. Aqui, entretanto, em vez de permitir a dependência de  $z_i(t)$  apenas em observações passadas e erros na localização  $i$ , a dependência é permitida com localizações vizinhas de várias ordens espaciais  $\ell$ . Em particular, o modelo STARMA para a  $i$ -ésima localização (site de observação) pode ser descrito na forma algébrica como:

$$z_i(t) = \sum_{k=1}^p \sum_{\ell=0}^{\lambda_k} \phi_{k\ell} L^{(\ell)} [z_i(t-k)] - \sum_{k=1}^q \sum_{\ell=0}^{m_k} \theta_{k\ell} L^{(\ell)} [\epsilon_i(t-k)] + \epsilon_i(t), \quad (3.9)$$

sendo:

- $\ell$  : atraso espacial
- $k$  : atraso temporal
- $p$  : ordem autorregressiva,
- $q$  : ordem da média móvel,
- $\lambda_k$  : ordem espacial do  $k$ -ésimo termo autorregressivo,
- $m_k$  : ordem espacial do  $k$ -ésimo termo da média móvel,
- $\phi_{k\ell}$  : parâmetro autorregressivo na defasagem temporal  $k$  e defasagem espacial  $\ell$ ,
- $\theta_{k\ell}$  : parâmetro de média móvel na defasagem temporal  $k$ , e defasagem espacial  $\ell$ ,
- $\epsilon_i(t)$  : amostra para o erro aleatório Gaussiano de média  $\mathbb{E}[\epsilon_i(t)] = 0$  e covariância.

$$\mathbb{E}[\epsilon_i(t)\epsilon_j(t+s)] = \begin{cases} \sigma^2 & \text{se } s = 0 \\ \zeta(s) & \text{se } s \neq 0, \end{cases} \quad (3.10)$$

a qual pode ser re-escrito na forma vetorial como STARMA vetorial [10], [17]:

$$\mathbf{z}(t) = \sum_{k=1}^p \sum_{\ell=0}^{\lambda_k} \phi_{k\ell} \mathbf{W}^{(\ell)} [\mathbf{z}(t-k)] - \sum_{k=1}^q \sum_{\ell=0}^{m_k} \theta_{k\ell} \mathbf{W}^{(\ell)} [\boldsymbol{\epsilon}(t-k)] + \boldsymbol{\epsilon}(t), \quad (3.11)$$

sendo:

- $\mathbf{z}(t)$  : vetor de observação de dimensão  $N \times 1$  nos  $N$  sites e instante  $t$ ,  
 $\mathbf{W}^{(\ell)}$  : matriz de pesos  $N \times N$  para ordem espacial  $\ell$ ,  
 $\boldsymbol{\epsilon}(t)$  : vetor de erro aleatório com distribuição Normal no tempo  $t$  com média zero e variância  $\sigma^2$ :

$$\mathbb{E}[\boldsymbol{\epsilon}(t)] = 0 \quad \text{e} \quad \mathbb{E}[\boldsymbol{\epsilon}(t)\boldsymbol{\epsilon}(t+s)'] = \begin{cases} \sigma^2 \mathbf{I}_N & \text{se } s = 0 \\ -\zeta(s) & \text{se } s \neq 0, \end{cases} \quad (3.12)$$

O modelo STARMA da Equação (3.11) tenta descrever e prever inter-relações espaciais entre diferentes regiões com um conjunto de  $N$  séries temporais observáveis que representam padrões em  $N$  pontos espaciais de observação (sites). O modelo STARMA é caracterizado por dependência linear defasada tanto no espaço quanto no tempo [17]. Este modelo é usado para ajustar e descrever variáveis aleatórias  $z_i(t)$  observadas em  $N$  combinações lineares de observações passadas  $z_i(t-k)$  e erros  $\boldsymbol{\epsilon}(t-k)$  que podem estar defasadas tanto no espaço quanto no tempo. O mecanismo de combinação linear ponderada é obtido pela ordenação hierárquica dos vizinhos de cada região com a sequência de uma matriz de pesos  $N \times N$ . Esta matriz de pesos pode representar qualquer relação espacial entre observações em diferentes regiões, dependendo da natureza da observação.

Finalmente, o modelo paramétrico STARIMA é expresso na Equação (3.13):

$$\nabla^d \mathbf{z}(t) = \sum_{k=1}^p \sum_{\ell=0}^{\lambda_k} \phi_{k\ell} \mathbf{W}^{(\ell)} \nabla^d [\mathbf{z}(t-k)] - \sum_{k=1}^q \sum_{\ell=0}^{m_k} \theta_{k\ell} \mathbf{W}^{(\ell)} [\boldsymbol{\epsilon}(t-k)] + \boldsymbol{\epsilon}(t), \quad (3.13)$$

agregando-se o operador diferença sazonal  $\nabla_s^D$  ou o operador diferença não sazonal  $\nabla^d$  [17], onde  $D$  e  $d$  são, respectivamente, o número de diferenças sazonais e não sazonais necessárias,  $\nabla_s^D$  e  $\nabla^d$  são os operadores de diferença sazonal e não sazonal para o  $k$ -ésimo termo da média móvel, tal que,  $\nabla_s^D = (\mathbf{I} - \mathbf{B}^s)^D$  e  $\nabla^d = (\mathbf{I} - \mathbf{B})^d$  com defasagem sazonal  $s$  representam as matrizes diferenças de ordem  $d$  e  $D$  (sazonal), respectivamente.

A matriz  $N \times N$  para o operador diferença  $\nabla$  de ordem  $d$  pode ser calculada a partir da definição do vetor diferença  $N \times 1$  de ordem  $d$ , dada por:

$$\begin{cases} \underline{d=1} : & \nabla \mathbf{z}(t) = \mathbf{z}(t) - \mathbf{z}(t-1), \\ \underline{d=2} : & \nabla^2 \mathbf{z}(t) = \nabla \{\nabla \mathbf{z}(t)\} = \mathbf{z}(t) - 2\mathbf{z}(t-1) + \mathbf{z}(t-2). \end{cases} \quad (3.14)$$

Assim, ao invés de modelar  $\mathbf{z}(t)$ , pode-se modelar o processo diferença de primeira ordem  $\nabla \mathbf{z}(t)$ , ou de segunda ordem,  $\nabla^2 \mathbf{z}(t)$ .

Quando não há componente sazonal (bastante improvável no fluxo de tráfego) e  $d = 0$  o modelo STARIMA colapsa no modelo STARMA mais fácil de interpretar, que tem a forma em (3.11). Note-se que em (3.11),  $p$  é a ordem autorregressiva (AR),  $q$  é a ordem da média móvel (MA),  $\lambda_k$  é a ordem espacial do termo AR  $k$ -ésimo,  $m_k$  é a ordem espacial do  $k$ -ésimo termo da média móvel,  $\phi_{k\ell}$  é o parâmetro autorregressivo (a ser estimado) no atraso temporal  $k$  e no atraso espacial  $\ell$ ;  $\theta_{k\ell}$ , é o parâmetro da média móvel (a ser estimado) no atraso temporal  $k$  e no atraso espacial  $\ell$ ;  $\mathbf{W}^{(\ell)}$  é a matriz  $N \times N$  de pesos para a ordem espacial  $\ell$ ; os elementos  $w_{i,j}^\ell$  é diferente de zero se houver uma relação espacial entre a localização  $i$  e  $j$ ; e  $\boldsymbol{\epsilon}(t)$  é um vetor de erro aleatório no tempo  $t$  com um valor esperado e variância dado em (3.12).

Para garantir que o modelo STARMA represente com precisão um processo estacionário, onde a estrutura de covariância permanece inalterada ao longo do tempo, pré-requisitos específicos devem ser cumpridos. Esses pré-requisitos são chamados de condições de estacionariedade e estipulam que qualquer solução  $\mathbf{x}_U$  [10] comporte a seguinte igualdade:

$$\det \left[ \mathbf{x}_U^p \mathbf{I} - \sum_{k=1}^p \sum_{\ell=0}^{\lambda_k} \phi_{k\ell} \mathbf{W}^{(\ell)} \mathbf{x}_U^{p-k} \right] = 0. \quad (3.15)$$

A variável em estudo é a velocidade relativa, definida como o volume de tráfego dividido pela ocupação das vias. Esta variável é menos estável que as outras duas, mas reflete claramente as condições de tráfego. Quando multiplicado por uma constante relacionada ao comprimento médio do veículo, pode fornecer uma aproximação para a velocidade exata [9],

$$VR = \frac{VT}{OR}, \quad (3.16)$$

onde  $VR$  representa a velocidade relativa,  $VT$  o volume de tráfego e  $OR$  a ocupação rodoviária.

### 3.2.3 Autocovariância e Autocorrelação

As funções usadas na análise de processos espaço-temporais, especificamente funções de autocovariância e autocorrelação, auxiliam no entendimento das dependências tanto no espaço quanto no tempo.

A função de Auto-Covariância Espaço-Temporal  $\Lambda_{\ell k}(s)$  mede a covariância estatística entre duas observações do mesmo processo espaço-temporal, separadas por uma defasagem  $s$ , dada por:

$$\Lambda_{\ell k}(s) = \mathbb{E} \left\{ \frac{[\mathbf{W}^{(\ell)} \nabla^d [\mathbf{z}(t)]]^T [\mathbf{W}^{(\ell)} \nabla^d [\mathbf{z}(t+s)]]}{N} \right\}. \quad (3.17)$$

onde,  $\mathbf{W}^{(\ell)}$  é um vetor de pesos aplicado às diferenças de ordem  $d$ , i.e.,  $\nabla^d [\mathbf{z}(t)]$  de um processo multivariado  $\mathbf{z}(t)$ , observado ao longo de  $N$  amostras.

A função de Autocorrelação Espaço-Temporal  $R_{\ell}(s)$  é dada por:

$$R_{\ell}(s) = \frac{\Lambda_{\ell k}(s)}{\sqrt{[\Lambda_{\ell \ell}(0)\Lambda_{kk}(0)]}}. \quad (3.18)$$

A Função de Autocorrelação Parcial Espaço-Tempo é expressa como:

$$\Lambda_{h0}(s) = \sum_{j=1}^{\infty} \sum_{i=0}^{\infty} \phi_{j\ell} \Lambda_{h\ell}(s-j). \quad (3.19)$$

## 3.3 Variações e Melhorias no STARIMA

### 3.3.1 $\Delta$ -STARIMA: STARIMA com Atrasos Variáveis no Tempo

$\Delta$ -STARIMA é um modelo paramétrico STARIMA com atrasos variáveis no tempo, capaz de melhorar a previsão de tráfego de curto prazo [20, 21]. Com base na observação de que a correlação entre o tráfego observado em dois pontos de medição ou estações de tráfego pode variar no tempo, atribuível à velocidade variável no tempo que, subsequentemente, causa variações no tempo necessário para viajar entre os dois pontos, em [21] desenvolve-se um modelo modificado de média móvel integrada autorregressiva espaço-tempo (STARIMA) com atrasos variáveis no tempo  $\Delta$ -STARIMA para previsão de fluxo de tráfego de curto prazo.

Particularmente, os atrasos temporais no  $\Delta$ -STARIMA mudam com a velocidade variável no tempo em diferentes horas do dia ou, equivalentemente, mudam com o tempo (variável no tempo) necessário para viajar entre dois pontos de medição. Primeiramente, os autores em [21] desenvolvem uma técnica para avaliar o atraso temporal no modelo STARIMA, onde o atraso temporal é formulado como uma função do atraso espacial (distância espacial) e da velocidade média. Em seguida, um algoritmo de classificação não supervisionado baseado no algoritmo ISODATA é projetado para classificar diferentes períodos de tempo do dia de acordo com a variação da velocidade. A classificação ajuda a determinar o intervalo de tempo apropriado para usar no modelo STARIMA. Finalmente, um modelo baseado em STARIMA com intervalos de tempo variáveis é desenvolvido para previsão de tráfego de curto prazo. Resultados experimentais usando

dados de tráfego reais mostram que o modelo baseado em STARIMA desenvolvido com intervalos de tempo variáveis tem precisão superior em comparação com sua contraparte desenvolvida usando a função de correlação cruzada tradicional e sem empregar intervalos de tempo variáveis.

### 3.3.2 STARIMA Dinâmico

Uma abordagem híbrida espaço-temporal para previsão de tráfego de curto prazo, conhecida como modelo dinâmico de média móvel integrada autorregressiva espaço-temporal (STARIMA dinâmico), é proposta em [22]. O modelo STARIMA e o modelo dinâmico de previsão de razão de conversão (DTRP) são usados nesta técnica para melhorar a precisão e a eficiência da previsão em cruzamentos urbanos. Para validar a modelagem dinâmica-STARIMA em um cenário do mundo real, um modelo experimental é construído usando dados reais para estimar o fluxo de tráfego para uma parte da rede metropolitana de Pequim, China. Comparando o modelo dinâmico STARIMA com outras técnicas de previsão, a precisão da previsão é satisfatória, demonstrando o benefício e a viabilidade da abordagem proposta.

### 3.3.3 STARIMA Aprimorado

O modelo STARIMA aprimorado prioriza o controle paramétrico nos dados do sistema independente da aplicação, cuja a previsão confiável principalmente no curto prazo. Como destacado em [23], a correlação entre dados ausentes e dados espaço-temporais disponíveis, bem como a correlação entre a topologia da rede rodoviária e o fluxo de tráfego são estudados a fim de minimizar o preenchimento de pontos de dados ausentes. Os autores propuseram um método de otimização para preenchimento de dados ausentes com base no mapa topológico espaço-temporal. Primeiramente, usam as “seções rodoviárias relevantes” definidas e os “níveis de correlação do fluxo de tráfego” para analisarem a correlação temporal e espacial do fluxo de tráfego da rede rodoviária e gerar um mapa de previsão do fluxo de tráfego. A previsão do fluxo de tráfego é realizada usando um modelo STARIMA. Os resultados da previsão mostram que o algoritmo proposto neste artigo reduz o tempo de previsão ao reduzir o número de pontos de preenchimento. Ao mesmo tempo, como o algoritmo seleciona a seção de tráfego mais relevante para a previsão do fluxo de tráfego, a precisão da previsão é ainda mais aprimorada, e um fluxo de tráfego efetivo e em tempo real é finalmente realizado.

## 3.4 STARIMA para Previsão de Fluxo de Tráfego de Curto Prazo

O modelo STARIMA tem sido amplamente utilizado para previsão de fluxo de tráfego de curto prazo devido à sua capacidade de incorporar as correlações espaciais e temporais presentes em sistemas de transporte. No trabalho [24], os autores propuseram uma abordagem unificada baseada no STARIMA, que demonstra resultados promissores para cenários complexos de tráfego. Discute-se a seguir a descrição e modelagem do STARIMA.

A essência do modelo STARIMA reside na extensão dos modelos ARIMA convencionais, adicionando componentes que consideram dependências espaciais e temporais. O fluxo de tráfego em um local específico em um instante  $t$  é influenciado não apenas por seus próprios valores históricos, mas também pelos fluxos de tráfego em localidades vizinhas em tempos anteriores. Essa interdependência é capturada por:

- **Dependência Temporal:** Utiliza termos autoregressivos e de médias móveis para modelar padrões temporais.
- **Dependência Espacial:** Introduce matrizes de peso espacial para representar a influência das localidades vizinhas.

No trabalho de Duan et al. [24], os autores propõem uma abordagem unificada para o modelo STARIMA, que aprimora a previsão de tráfego de curto prazo ao integrar:

- **Seleção Adaptativa de Pesos Espaciais:** O modelo permite ajustar dinamicamente os pesos na matriz espacial com base nas condições do tráfego em tempo real, garantindo maior precisão ao capturar correlações espaciais dinâmicas.
- **Integração Multiescala:** A metodologia unificada combina informações de granularidades espaciais e temporais diferentes, permitindo ao modelo capturar tanto flutuações rápidas quanto tendências de longo prazo no fluxo de tráfego.
- **Estimativa Robusta de Parâmetros:** Técnicas avançadas de estimação garantem que o modelo se adapte bem a dados ruidosos e situações de tráfego irregulares.

Os experimentos apresentados no artigo demonstram a eficácia do modelo unificado STARIMA em comparação com métodos tradicionais e outras variantes modernas. Em diferentes conjuntos de dados reais de tráfego urbano, os resultados indicam: a) *Melhoria na Precisão de Previsão*, pois o modelo unificado STARIMA apresentou menor erro médio absoluto (MAE) e raiz do erro quadrático médio (RMSE) em comparação com abordagens concorrentes. b) *Maior Resiliência a Padrões de Tráfego Não Estacionários*: A capacidade do modelo de capturar variações dinâmicas espaciais e temporais garantiu desempenho consistente em cenários de tráfego congestionado.

A abordagem unificada baseada no modelo STARIMA proposta por Duan et al. [24] representou um avanço significativo no campo da previsão de fluxo de tráfego de curto prazo. Sua estrutura flexível e robusta o torna ideal para aplicações em sistemas de transporte inteligentes, onde a precisão e a adaptabilidade são fundamentais. A integração de componentes espaço-temporais adaptativos estabelece uma base sólida para futuros desenvolvimentos nessa área.

### 3.4.1 Efeito da Segmentação de Séries Temporais no Modelo STARIMA

O modelo STARIMA é amplamente aplicado na previsão de fluxo de tráfego devido à sua capacidade de capturar dependências espaço-temporais. No trabalho [25], os autores investigaram como a segmentação de séries temporais pode influenciar o desempenho do modelo STARIMA em cenários de tráfego dinâmicos e heterogêneos.

A segmentação de séries temporais é uma técnica que divide os dados em segmentos menores, de modo que cada segmento capture diferentes padrões temporais. Essa abordagem é particularmente útil em sistemas de tráfego, onde as condições podem variar significativamente entre diferentes períodos do dia (ex.: horários de pico e períodos de baixa atividade). A segmentação permite que o modelo STARIMA:

- **Adapte-se a Variações Locais:** Ao tratar segmentos independentes, o modelo pode ajustar seus parâmetros para capturar melhor os padrões específicos de cada segmento.
- **Melhore a Estabilidade do Modelo:** Dados altamente variáveis podem levar à superestimação ou subestimação de parâmetros. A segmentação reduz essa variabilidade.

O modelo STARIMA segmentado é formalizado como uma coleção de modelos STARIMA específicos para cada segmento:

$$\Phi_i(L, S)Y_i(t) = \Theta_i(L, S)\epsilon_i(t), \quad \forall i \in \{1, 2, \dots, n\}, \quad (3.20)$$

onde:

- $i$  representa o segmento específico;
- $Y_i(t)$  é o vetor de fluxos de tráfego no instante  $t$  dentro do segmento  $i$ ;
- $\Phi_i(L, S)$  e  $\Theta_i(L, S)$  são os polinômios espaço-temporais ajustados para o segmento  $i$ ;
- $\epsilon_i(t)$  é o termo de erro branco para o segmento  $i$ .

Salamanis et al. [25] avaliaram o impacto da segmentação de séries temporais no desempenho do modelo STARIMA com base nos seguintes passos:

#### 1. Segmentação dos Dados:

- Utilizaram critérios baseados em horários (ex.: divisão em períodos de pico e fora de pico).
- Compararam segmentações fixas (ex.: divisão por períodos de tempo iguais) e dinâmicas (baseadas em padrões de tráfego detectados).

#### 2. Treinamento de Modelos Locais:

- Para cada segmento, um modelo STARIMA distinto foi treinado.

#### 3. Medição de Desempenho:

- Avaliaram a precisão das previsões utilizando métricas de desempenho.

Os experimentos realizados pelos autores revelaram que a segmentação temporal pode melhorar significativamente o desempenho do modelo STARIMA, especialmente em cenários onde os padrões de tráfego apresentam forte sazonalidade ou variabilidade. As principais descobertas e conclusões são:

- **Melhoria da Exatidão:** a segmentação temporal permitiu capturar padrões específicos de horários de pico, reduzindo os erros de previsão em até 15% em comparação com o modelo STARIMA sem segmentação.
- **Impacto da Escolha de Segmentação:** Segmentações dinâmicas baseadas em características do tráfego apresentaram melhores resultados do que segmentações fixas.
- **Limitações:** o aumento no número de segmentos introduziu maior complexidade computacional e possíveis problemas de *overfitting* em segmentos com poucos dados.

O estudo apresentado em [25] destaca a importância da segmentação de séries temporais na melhoria do desempenho do modelo STARIMA em previsões de tráfego. A abordagem segmentada oferece maior adaptabilidade e exatidão, mas exige um equilíbrio entre granularidade da segmentação e capacidade de generalização. Esses resultados reforçam a relevância de estratégias personalizadas em sistemas de transporte inteligentes para capturar as complexas dinâmicas de tráfego em ambientes urbanos.

### 3.5 Redes Neurais Recorrentes (RNN) e Redes com Memória de Curto e Longo Prazos (LSTM)

Diversos tipos de dados utilizados em aplicações reais, como séries temporais, textos em linguagem natural e sinais de áudio, apresentam uma característica comum: a presença de dependências temporais, ou seja, informações que estão codificadas não apenas no conteúdo individual de cada elemento da sequência, mas também na ordem em que esses elementos ocorrem. O significado e o comportamento dinâmico desses

dados só podem ser plenamente compreendidos considerando-se a sequência de eventos e suas inter-relações ao longo do tempo. Para lidar com esse tipo de estrutura sequencial, as redes neurais recorrentes (Recurrent Neural Networks — RNNs) foram desenvolvidas como uma classe especializada de modelos neurais capazes de manter um estado interno que evolui a cada passo temporal, permitindo à rede capturar e representar informações contextuais passadas que influenciam as previsões futuras [2, 26].

Uma forma de compreender o funcionamento das redes neurais recorrentes é por meio de sua representação desdobrada apresentada na Figura 4, na qual a rede é expandida ao longo da sequência temporal de entrada. Essa abordagem permite visualizar claramente o fluxo de informações entre os estados ao longo do tempo, eliminando os ciclos presentes na arquitetura original e possibilitando a definição adequada das fases de propagação direta e retropropagação.

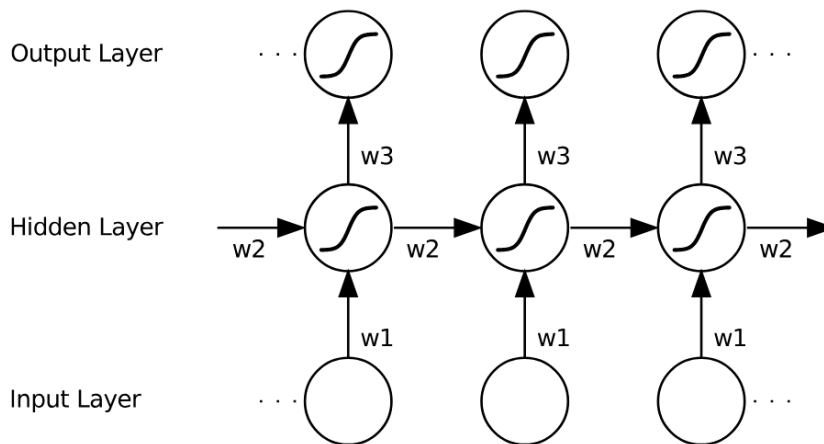


Figura 4 – Estrutura básica de RNN. Fonte: [2]

O processo de propagação direta em uma rede neural recorrente pode ser interpretado como uma extensão da arquitetura de um perceptron multicamada com uma única camada oculta. A principal diferença reside no fato de que, nas RNNs, os neurônios da camada oculta recebem entradas tanto da entrada externa atual quanto das ativações ocultas do passo de tempo anterior. Isso permite que as RNNs mantenham uma memória das ativações anteriores, modelando dependências temporais. A sequência completa de ativações é obtida por meio de um processo recursivo que começa em  $t = 1$  e depende de valores iniciais para os estados ocultos, usualmente definidos como zero, embora outras inicializações também possam ser benéficas para a estabilidade da rede. As saídas, por sua vez, são calculadas de forma semelhante à de um MLP e podem empregar funções como softmax ou sigmoide, dependendo da tarefa de classificação ou regressão [2].

Considerando uma sequência de entrada de comprimento  $T$ ,  $\mathbf{x}$ , apresentada a uma RNN com  $I$  unidades de entrada,  $H$  unidades ocultas e  $K$  unidades de saída, tem-se a seguinte relação [2]:

$$a_h^t = \sum_{i=1}^I w_{ih} x_i^t + \sum_{h'=1}^H w_{h'h} b_{h'}^{t-1}, \quad (3.21)$$

em que  $x_i^t$  é o valor da entrada  $i$  no instante  $t$ ,  $a_j^t$  e  $b_j^t$  são, respectivamente, a entrada e a ativação da unidade  $j$  no instante  $t$ ,  $w_{ih}$  é o peso conectando a entrada  $i$  à unidade oculta  $h$ , e  $w_{h'h}$  é o peso recorrente conectando a unidade oculta  $h'$  à unidade  $h$ .

Em seguida, uma função de ativação não linear e diferenciável,  $\theta_h$  é aplicada para obter a saída [2]:

$$b_h^t = \theta_h(a_h^t), \quad (3.22)$$

embora úteis, redes recorrentes convencionais enfrentam limitações práticas para capturar dependências de longo prazo, devido ao problema conhecido como *vanishing gradient* [2]. Esse fenômeno ocorre quando

os gradientes necessários para o treinamento se tornam extremamente pequenos ou grandes ao longo das iterações, comprometendo o aprendizado de relações distantes no tempo [2]. Diversas abordagens foram propostas para contornar essa limitação; no entanto, a solução mais adotada e eficaz tem sido a arquitetura Long Short-Term Memory (LSTM) [2].

A arquitetura LSTM é uma variação das RNNs projetada para lidar com dependências de longo prazo em sequências temporais. Sua principal inovação está na substituição das unidades tradicionais de soma por blocos de memória, como ilustrado na Figura 5.

Cada bloco de memória é composto por:

- Uma célula de memória, que atua como um armazenamento de longo prazo, permitindo que a rede retenha informações importantes ao longo do tempo. Seu funcionamento é análogo ao de um chip de memória em computadores digitais;
- Três portas multiplicativas (portas de entrada, esquecimento e saída), que são unidades não lineares controladas por funções de ativação, geralmente a função sigmoide (com saída entre 0 e 1). Essas portas regulam o fluxo de informações dentro e fora da célula.

O funcionamento das portas é descrito a seguir:

**Porta de entrada:** regula a entrada de novas informações na célula. Quando seu valor de ativação está próximo de 1, permite que dados externos sejam armazenados. Quando próxima de 0, bloqueia a entrada, mantendo o conteúdo atual da célula.

**Porta de esquecimento:** controla a retenção das informações anteriores armazenadas na célula. Um valor próximo de 1 mantém a memória; um valor próximo de 0 "esquece" a informação, o que ajuda a eliminar conteúdos irrelevantes ou desatualizados.

**Porta de saída:** decide se o conteúdo da célula será transmitido para a próxima camada ou para o próximo instante de tempo, controlando o impacto da memória armazenada no restante da rede.

As funções de ativação adicionais dentro da célula, como tanh, são utilizadas para limitar e normalizar os valores de entrada e saída. Além disso, o LSTM pode empregar conexões peephole (representadas por linhas tracejadas na figura), que permitem que o estado interno da célula influencie diretamente as portas, proporcionando um controle ainda mais refinado.

As conexões internas dentro do bloco geralmente possuem peso fixo (tipicamente 1.0), garantindo estabilidade e simplicidade na propagação do estado da célula. As saídas do bloco, no entanto, são moduladas pela porta de saída, assegurando que apenas informações relevantes sejam transmitidas.

Graças a essa estrutura, a LSTM é capaz de armazenar e acessar informações por longos períodos de tempo, mitigando o conhecido problema do desvanecimento do gradiente que afeta RNNs tradicionais. Isso torna a arquitetura especialmente eficaz em tarefas que exigem memória de longo alcance, como:

A passagem direta da rede LSTM consiste em uma série de operações envolvendo portões e estados de célula, os quais determinam como a informação flui ao longo do tempo. Cada operação de portão é definida por somas ponderadas da entrada atual, do estado oculto anterior e do estado de célula anterior, seguidas por funções de ativação não lineares.

### 3.5.1 Portas de Entrada

A ativação das portas de entrada no tempo  $t$ , denotado por  $a^t$ , pode ser computada através da seguinte equação (3.23) [2]:

$$a^t = \sum_{i=1}^I w_i x_i^t + \sum_{h=1}^H w_h b_h^{t-1} + \sum_{c=1}^C w_c s_c^{t-1}, \quad (3.23)$$

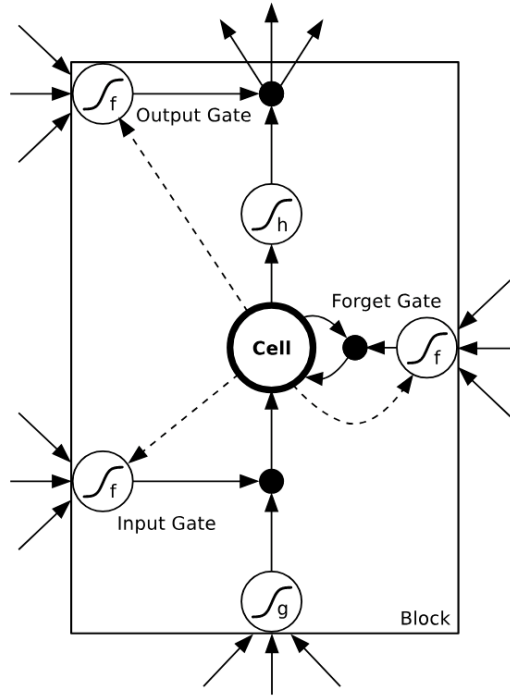


Figura 5 – Arquitetura de um bloco de memória de LSTM. Fonte: [2]

em que  $x_i^t$  representa a  $i$ -ésima componente do vetor de entrada no tempo  $t$ ,  $b_h^{t-1}$  é a saída da unidade oculta  $h$  no passo anterior (tempo  $t - 1$ ) e  $s_c^{t-1}$  é o estado da célula no tempo anterior. Os pesos  $w_i$ ,  $w_h$  e  $w_c$  definem as contribuições da entrada, das unidades ocultas do estado anterior e o estado anterior da célula. O número total de entradas, unidades ocultas e células são denotadas por  $I$ ,  $H$  e  $C$ , respectivamente.

A saída no *gate* de entrada é computado aplicando uma função de ativação não-linear em  $a_i^t$  [2]:

$$b^t = f(a^t). \quad (3.24)$$

### 3.5.2 Portas de Esquecimento

De forma análoga, pode-se definir a saída das portas de esquecimento [2]:

$$a_\phi^t = \sum_{i=1}^I w_{i\phi} x_i^t + \sum_{h=1}^H w_{h\phi} b_h^{t-1} + \sum_{c=1}^C w_{c\phi} s_c^{t-1}. \quad (3.25)$$

Assim como para as portas de entrada, essa equação computa a soma da entrada atual,  $x_i^t$ , saídas das unidades ocultas anteriores  $b_h^{t-1}$ , e os estados das células anteriores,  $s_c^{t-1}$  por meio dos pesos  $w_{i\phi}$ ,  $w_{h\phi}$  e  $w_{c\phi}$ : A saída é, então obtida aplicando uma função de ativação não-linear [2]:

$$b_\phi^t = f(a_\phi^t). \quad (3.26)$$

### 3.5.3 Células de Memória

A ativação candidata da célula de memória é computada com base na entrada atual e no estado oculto anterior [2]:

$$a_c^t = \sum_{i=1}^I w_{cx} x_i^t + \sum_{h=1}^H w_{ch} b_h^{t-1}. \quad (3.27)$$

Diferentemente dos portões, a ativação da célula normalmente não depende diretamente do estado de célula anterior. O novo estado de célula,  $s_c^t$  é então obtido como uma combinação entre o estado anterior

modulado pelo portão de esquecimento e a nova ativação candidata modulada pelo portão de entrada [2]:

$$s_c^t = b_\phi^t s_c^{t-1} + b^t g(a_c^t), \quad (3.28)$$

onde  $g(\cdot)$  é uma função de ativação que introduz não linearidade.

### 3.5.4 Portas de Saída

A porta de saída controla a quantidade do estado de célula que será exposta como saída. A ativação do portão é dada por [2]:

$$a_o^t = \sum_{i=1}^I w_{ox} x_i^t + \sum_{h=1}^H w_{oh} b^{t-1} h + \sum_{c=1}^C w_{oc} s_c^t. \quad (3.29)$$

Esta equação considera a entrada atual, o estado oculto anterior e o estado de célula atual, com pesos  $w_{ox}$ ,  $w_{oh}$  e  $w_{oc}$ . A saída do portão de saída é obtida pela aplicação da função de ativação [2]:

$$b_o^t = f(a_o^t). \quad (3.30)$$

### 3.5.5 Saída da Célula

Finalmente, o estado oculto (ou saída) da célula no instante  $t$ , denotado por  $b_c^t$ , é computado como a ativação do estado de célula modulado pelo portão de saída [2]:

$$b_c^t = b_o^t h(s_c^t), \quad (3.31)$$

onde a função  $h(\cdot)$ , é uma função de ativação que garante que a saída seja suavizada e permaneça dentro de um intervalo controlado.

### 3.5.6 Sumário das Características do Modelo Aprendizagem de Máquina LSTM

O pré-processamento dos dados quando aplicado ao método LSTM pode reduzir significativamente a complexidade temporal de uma tarefa de aprendizagem [2]. Se for possível transformar um problema com dependências de longo prazo em outro com dependências de curto alcance, arquiteturas complexas como o LSTM podem se tornar desnecessárias. Por exemplo, ao invés de trabalhar com um sinal bruto com alta taxa de amostragem, é possível reduzir a resolução temporal e tornar viável o uso de modelos com alcance contextual limitado, como modelos ocultos de Markov (HMMs).

No entanto, nem sempre esse tipo de transformação é trivial ou viável, especialmente quando não se conhece um bom método de pré-processamento. Nesses casos, ou quando se deseja evitar o esforço de desenvolver soluções específicas para cada tarefa, redes com capacidade de capturar dependências de longo prazo, como o LSTM, continuam sendo fundamentais para o bom desempenho em tarefas sequenciais complexas [2].

## 3.6 Processos de Poisson

Para garantir maior eficiência energética e reduzir o tempo de viagem, o controle do tráfego urbano é melhorado via semáforo re-programável dinamicamente. O controle de semáforo em um único modelo de interseção é projetado com base no sistema de semáforo real. No modelo apresentado, existem quatro abordagens (leste, oeste, norte e sul) com um total de oito fluxos. Cada fluxo é modelado como uma fila M/M/1 (baseado na notação de Kendall) que é constituído com base nos três conceitos da teoria das filas que são: clientes, filas e servidores. O primeiro e o segundo M em M/M/1 significam distribuição “sem memória” de tempo entre chegadas e tempo de serviços, respectivamente. O “1” é a indicação de uma

única interseção do servidor com horário de serviço específico, [27]; Figura 6 esquematiza a intersecção do tráfego urbano.

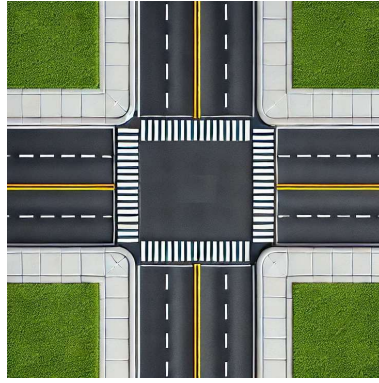


Figura 6 – Exemplo de uma aproximação de um semáforo de tempo fixo [Fonte: autor].

O procedimento Primeiro a Entrar Priemeiro a Sair (FIFO - *First-In-First-Out*) é aplicado à fila de veículos em cada fluxo. Os veiculos neste modelo são conhecidos como clientes, pela teoria das filas, enquanto o tempo de atendimento é o tempo para os veículos saírem da fila e atravessarem a interseção em seu próprio fluxo [27].

A característica do tráfego de chegada é modelada conforme o processo de Poisson [27]. Muitos modelos têm sido usados para descrever o processo de chega de veículos na aproximação, mas a maioria é complexa. Sendo assim, o modelo de Poisson ainda é usado para fins mais práticos. Com o fluxo de Poisson, a probabilidade de chegada de  $\mu$  veículos em uma interseção na via de tráfego de veículos, no intervalo  $\tau$ , pode ser modelado por:

$$P_{\mu}(\tau) = \Pr[\mu \text{ veículos no intervalo } \tau] = \frac{(\lambda\tau)^{\mu}}{\mu!} e^{-\lambda\tau} \quad (3.32)$$

sendo  $\lambda$  a taxa de chegada de veículos ou densidade média de veículos em [# veículos/unid. tempo], por exemplo  $\lambda = \frac{1}{15}$  [veiculos/seg]

Duas distribuições estatísticas mais usadas na emulação do tráfego de veículos e pedestres urbanos são as distribuições Beta e Poisson, revisitadas a seguir.

### 3.6.1 Emulando o Tráfego de Veículos Baseado em Distribuições de Poisson e Beta

Para a distribuição de Poisson:

$$P[M_v = m|\lambda] = \frac{\lambda^m}{m!} e^{-\lambda}, \quad (3.33)$$

onde o número médio de usuários que chegam  $\mathbb{E}[M_v] = \lambda$  é fixo e conhecido.

De acordo com a distribuição Beta, uma probabilidade de ativação do dispositivo em um intervalo de tempo  $t \in [0, T_A]$  é fornecida por [28]:

$$P(t) = \frac{t^{\alpha-1}(T_A - t)^{\beta-1}}{T_A^{\alpha+\beta-1} \text{Beta}(\alpha, \beta)} \quad (3.34)$$

onde  $\text{Beta}(\alpha, \beta) = \int_0^1 t^{\alpha-1}(T_A - t)^{\beta-1} dt$ ,  $\alpha$  e  $\beta$  são parâmetros de forma e  $T_A$  é o tempo de ativação.  $\alpha = 3$ ,  $\beta = 4$  e  $T_A = 10s$  conforme o modelo 3GPP. Figura 8 mostra o comportamento da função Beta para diferentes parâmetros  $\alpha, \beta$ , enquanto a Figura 7 representa  $p(t) \times t \in [0, 2]$  com  $\beta \in [0, 7]$  e  $\alpha = 3$ . A Figura 8 representa o mesmo para  $\alpha \in [0, 5]$  com  $\beta = 4$ .

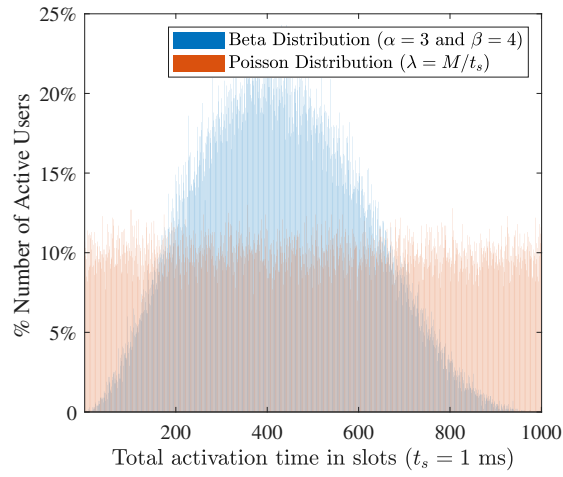
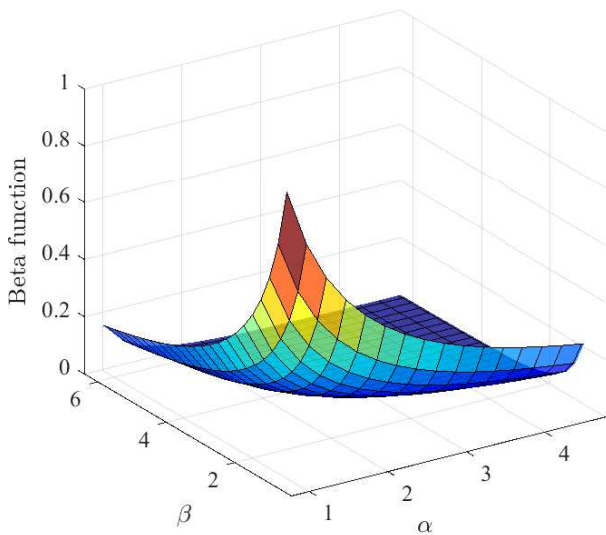
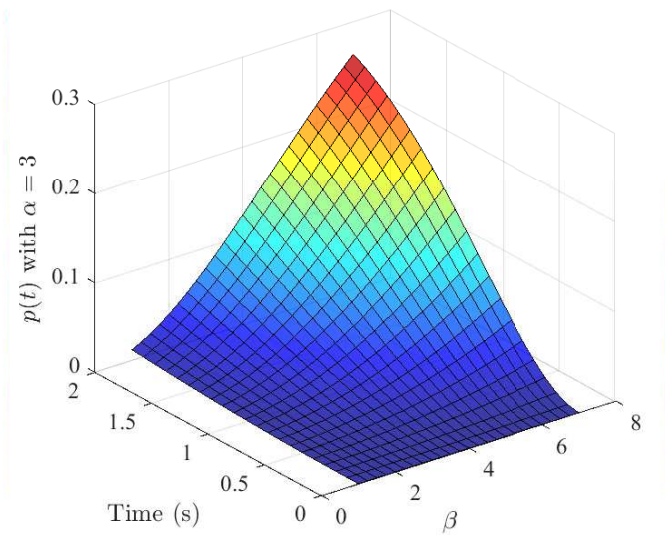


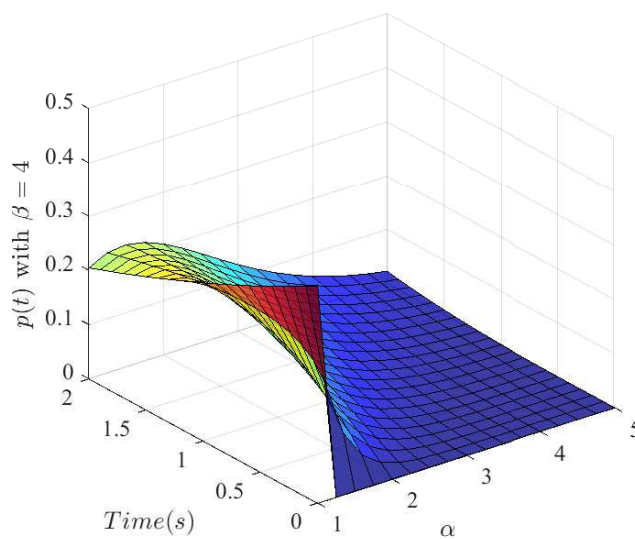
Figura 7 – Chegada de usuários em um Frame  $F$  com  $n$  intervalos de tempo considerando distribuição Poisson  $\times$  Beta e  $M = 100$  entidades.



a) Beta( $\alpha, \beta$ )



b) Ativação do usuário  $p(t)$  para  $\alpha = 3$



c) Ativação do usuário  $p(t)$  para  $\beta = 4$ .

Figura 8 – (a) Função Beta como Beta( $\alpha, \beta$ ). (b) Probabilidade de ativação do usuário,  $p(t, \beta)$ . (c) Probabilidade de ativação do usuário,  $p(t, \alpha)$

Observa-se, na Figura 8, que os parâmetros  $\alpha$  e  $\beta$  influenciam diretamente a forma da distribuição Beta e, conseqüentemente, a probabilidade de ativação dos usuários ao longo do tempo. Em a), nota-se que a função Beta atinge maiores valores para  $\alpha$  e  $\beta$  menores, concentrando a probabilidade nos extremos do intervalo. Já em b), mantendo  $\alpha = 3$ , é possível verificar que o aumento de  $\beta$  desloca o pico da ativação para tempos mais próximos do início da janela de ativação. Por fim, em (c), com  $\beta = 4$ , a variação de  $\alpha$  controla o crescimento inicial da função, modulando a densidade de ativações no início do intervalo. Esses comportamentos destacam a flexibilidade da distribuição Beta na modelagem da chegada de usuários, permitindo ajustes precisos conforme diferentes cenários de tráfego veicular.

# 4 MATERIAIS E MÉTODOS

Inicialmente discute-se o procedimento adotada para o ajuste dos parâmetros de entrada dos modelos paramétricos. Em seguida será descrito as principais métricas utilizadas na análise de desempenho dos modelos de predição das séries temporais associadas ao tráfego urbano.

## 4.1 Metodologia para o Ajuste de Parâmetros do Modelo Paramétrico

### 4.1.1 Número de Parâmetros Ajustáveis do Modelo

O número de parâmetros ajustáveis de um modelo refere-se aos componentes internos que podem ser modificados ou otimizados durante o processo de treinamento, com o objetivo de melhorar a performance preditiva do modelo. Esses parâmetros são cruciais, pois determinam a capacidade do modelo de capturar padrões complexos nos dados e, conseqüentemente, a sua eficácia na realização de previsões precisas. A quantidade e a natureza desses parâmetros variam significativamente conforme o tipo de modelo em questão; por exemplo, redes neurais profundas possuem milhões de parâmetros ajustáveis, enquanto modelos lineares simples podem ter apenas alguns poucos.

### 4.1.2 Procedimento de Ajuste dos Parâmetros

O procedimento de ajuste dos parâmetros no modelo paramétrico também pode variar, desde técnicas de otimização clássicas, como gradiente descendente, até métodos mais avançados como otimização bayesiana ou algoritmos genéticos. O ajuste adequado dos parâmetros é necessário para evitar tanto o *underfitting*, onde o modelo é incapaz de capturar a complexidade dos dados, quanto o *overfitting*, onde o modelo se torna excessivamente ajustado ao conjunto de dados de treinamento, comprometendo sua generalização para novos dados. Portanto, a escolha do número de parâmetros ajustáveis e a estratégia de ajuste são fundamentais para o desenvolvimento de um modelo robusto e eficiente.

Um modelo paramétrico bem ajustado envolve a integração organizada dos seguintes princípios:

- **Definição do Modelo:** Inicialmente, seleciona-se o tipo de modelo a ser utilizado (por exemplo, regressão linear, rede neural, árvore de decisão, etc.). Cada tipo de modelo possui um conjunto específico de parâmetros ajustáveis. Por exemplo, em uma regressão linear, os coeficientes são os parâmetros ajustáveis, enquanto em uma rede neural, os pesos das conexões entre os neurônios são ajustáveis [29].
- **Inicialização dos Parâmetros:** Os parâmetros ajustáveis são inicialmente definidos. Em muitos casos, essa inicialização é feita de forma aleatória ou utilizando alguma heurística pré-definida. A escolha da inicialização pode impactar significativamente o processo de otimização e a convergência do modelo [30].
- **Função de Custo:** Define-se uma função de custo que quantifica o erro entre as previsões do modelo e os valores reais observados. A função de custo guia o processo de ajuste dos parâmetros. Exemplos comuns de funções de custo incluem Erro Quadrático Médio ( *Mean Squared Error* MSE), Cross-Entropy Loss, entre outras [29].
- **Algoritmo de Otimização:** Utiliza-se um algoritmo de otimização para ajustar os parâmetros do modelo de forma a minimizar a função de custo. Algoritmos de otimização comuns incluem

Gradiente Descendente, Gradiente Descendente Estocástico, e algoritmos mais avançados como Adam e RMSprop [30].

- **Validação Cruzada:** Para evitar *overfitting* e garantir que o modelo generalize bem para novos dados, realiza-se a validação cruzada. Nesta etapa, o conjunto de dados é dividido em múltiplas partes, treinando e validando o modelo em diferentes subconjuntos de dados. Isso ajuda a avaliar a performance do modelo de forma mais robusta e ajustar os parâmetros de maneira mais eficaz [29].
- **Ajuste Fino:** Após a validação cruzada, realiza-se um ajuste fino dos parâmetros, muitas vezes utilizando técnicas como *Grid Search* ou *Random Search* para encontrar a combinação ótima de hiperparâmetros que minimizem a função de custo [30].
- **Avaliação Final:** Finalmente, avalia-se a performance do modelo ajustado utilizando um conjunto de testes independente. Esta avaliação fornece uma estimativa da precisão do modelo quando aplicado a dados não vistos anteriormente, garantindo a eficácia dos parâmetros ajustados [29].

O ajuste adequado dos parâmetros é crucial para a construção de modelos preditivos eficazes e robustos.

## 4.2 Associação da Malha Viária às Equações Espaço-Temporais STARIMA

Nesta seção, o objetivo é apresentar um modelo ilustrativo de como o STARIMA pode ser aplicado ao cenário em análise. Para exemplificar, adotamos a estrutura em árvore da Figura 9, típica de redes rodoviárias utilizadas na descrição do fluxo de tráfego urbano e intermunicipal. Os pontos representam locais de medição/nó medido/observado (site) e as setas indicam a direção do fluxo.  $N = 6$  sites implementado com a Equação (3.11):

$$\mathbf{z}(t) = [z_1(t) \ z_2(t) \ z_3(t) \ z_4(t) \ z_5(t) \ z_6(t)]^T$$

$$\mathbf{W}^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{z}^{(\ell)}(t) = L^{(\ell)}[\mathbf{z}(t)] = L^{(\ell)}[z_1(t), \dots, z_n(t)]^T = \sum_{j=1}^N w_{ij}^{(\ell)} z_j(t) = \mathbf{W}^{(1)} \mathbf{z}(t).$$

A matriz  $\mathbf{W}^{(\ell)}$  segue o critério de formação: Por meio do Lag 1, no site  $z_1$  para o site  $z_3$  é igual a 0.5 e do site  $z_2$  para o site  $z_3$  também é 0.5. Levando em consideração que a via em todos os trechos comporta a mesma quantidade de carros. E validando a somatória dos elementos em uma mesma linha da matriz  $\mathbf{W}^{(\ell)}$  deve ser igual a um, Equação (3.6).

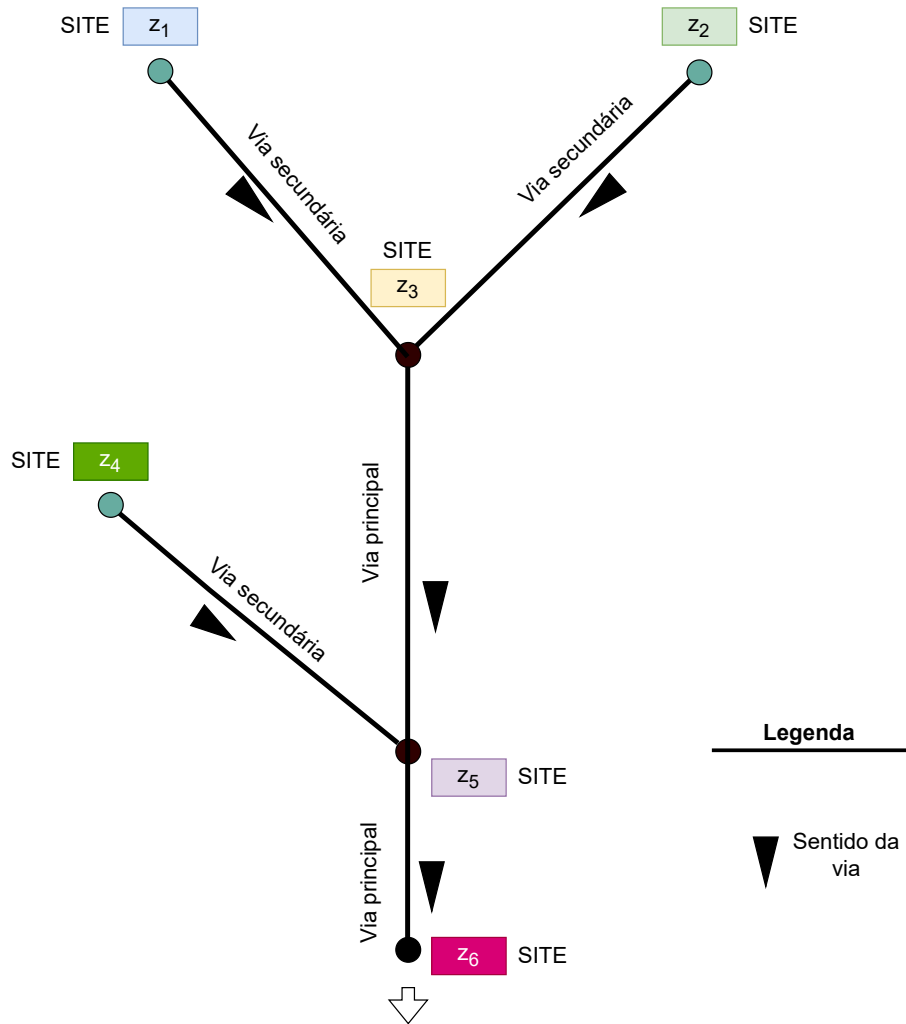


Figura 9 – Estrutura em árvore típica da rede rodoviária para fluxo de tráfego.

**Modelo Ilustrativo 1:**

$$\mathbf{z}(t) = [z_1(t) \ z_2(t), \ z_3(t), \dots, \ z_N(t)]^T$$

$$\mathbf{W}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$

$$L^{(\ell)}[\mathbf{z}(t)] = L^{(\ell)}[z_1(t), \ z_2(t), \ z_3(t), \dots, \ z_n(t)]^T = \sum_{j=1}^N w_{ij}^{(\ell)} z_j(t) = \mathbf{W}^{(2)} \mathbf{z}(t) \quad (4.1)$$

$$\mathbf{z}(t) = [z_1(t) \ z_2(t), \ z_3(t), \dots, \ z_N(t)]^\top$$

$$\mathbf{W}^{(3)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$L^{(\ell)}[\mathbf{z}(t)] = L^{(\ell)}[z_1(t), \ z_2(t), \ z_3(t), \dots, \ z_n(t)]^\top = \sum_{j=1}^N w_{ij}^{(\ell)} z_j(t) = \mathbf{W}^{(3)}\mathbf{z}(t), \quad (4.2)$$

sendo  $L^{(\ell)}[\mathbf{z}(t)]$  uma atualização dos estados / sites de  $\mathbf{z}(t)$ .

Para  $\ell = 4$  (lag espacial igual a 4),

$$\mathbf{z}(t) = [z_1(t) \ z_2(t), \ z_3(t), \dots, \ z_n(t)]^\top$$

$$\mathbf{W}^{(4)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$L^{(\ell)}[\mathbf{z}(t)] = L^{(\ell)}[z_1(t), \ z_2(t), \ z_3(t), \dots, \ z_n(t)]^\top = \sum_{j=1}^N w_{ij}^{(\ell)} z_j(t) = \mathbf{W}^{(4)}\mathbf{z}(t), \quad (4.3)$$

**Modelo Ilustrativo 2 – Evolução dos Lags Espaço-Tempo:** Considere a malha viária hipotética da Figura 9. Parâmetros adotados nestas simulações:

- 6 sites,  $N = 6$ , conforme Figura 9.
- Janela temporal de simulação de 4 segundos (geração e análise de tráfego nos pontos  $\mathbf{z}[n] = (z_1[n] \ z_2[n] \ z_3[n] \ z_4[n] \ z_5[n] \ z_6[n])$ ).
- A taxa de amostragem adotada foi de  $T = 1$  [amostras/seg].
- Valores de tráfego gerado em cada entrada da via, foram determinados neste momento inicial para efeito de testes do código e validação da equação.
- $z_{in} = [3 \ 2 \ 0 \ 0 \ 0 \ 0]$ ;

Na Figura 10, é possível observar o resultado da multiplicação da matriz  $\mathbf{W}^{(\ell)}$  pelo vetor  $\mathbf{z}(t)$ . Na Figura 10.a), durante o período, foi observado a entrada de 3 carros em  $z_1$  e 2 carros em  $z_2$ . Observam-se as entradas nesses locais, com a matriz  $\mathbf{W}^{(0)}$  sendo utilizada na multiplicação por  $\mathbf{z}(t)$ . Em  $t_2$ , 2 carros chegam ao local  $z_3$ , um proveniente de  $z_1$  e o outro de  $z_2$ . Em  $t_3$  chegam um valor esperado (médio) de 1,25 carros em  $z_6$  provenientes dos nós medidos/observados (sites) anteriores. Na Figura 10.b), observa-se que o tráfego foi fluindo ao longo da via semelhante ao exemplo anterior.

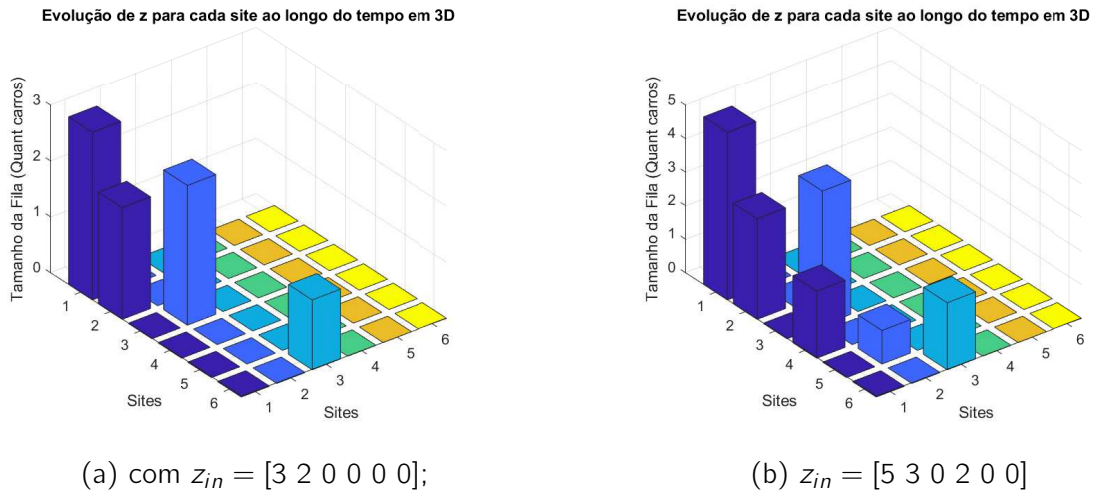


Figura 10 – Evolução de  $z(t)$  para  $L^{(\ell)} [z(t)]$  com  $N = 6$  sites. Referente à Equação (3.8).

**Exemplo 3 – Malha Rodoviária Típica:  $N=13$  sites de Observação:** Na Figura 11 apresentam-se as matrizes  $W^{(\ell)}$  com lags de  $\ell = 1$  a 5 com  $\ell > 0$ , seguindo a estrutura em árvore típica da rede rodoviária de uma cidade de pequeno ou médio porte.

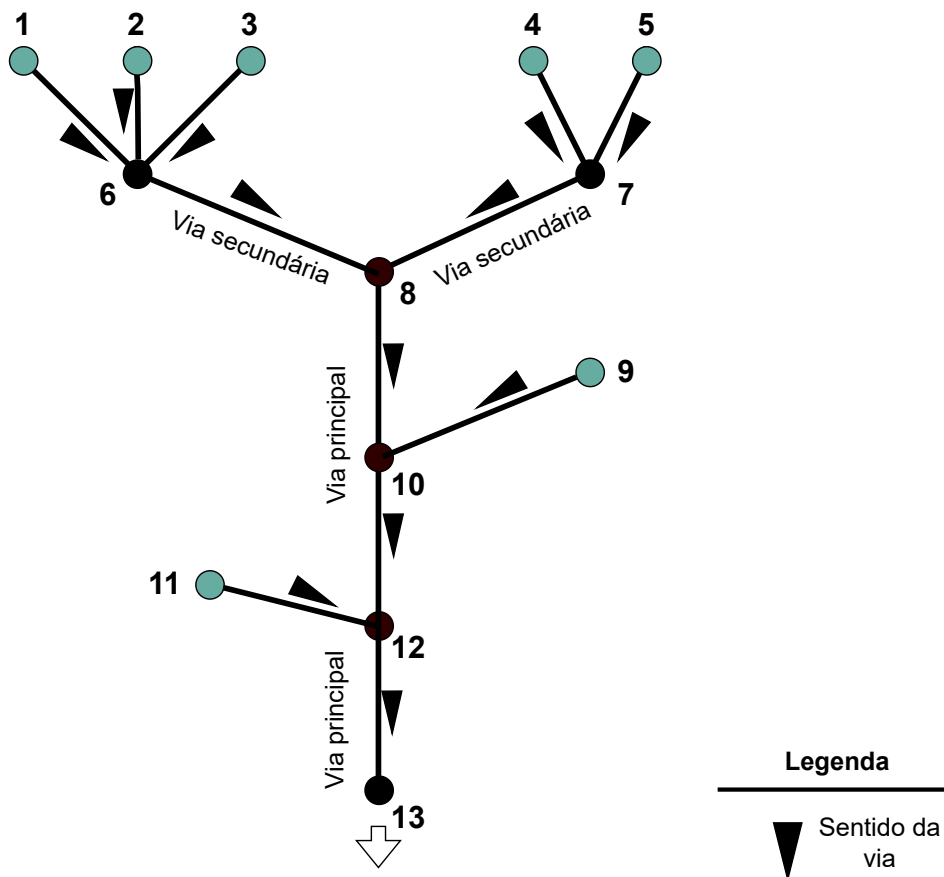


Figura 11 – Estrutura em árvore típica da rede rodoviária para fluxo de tráfego. Os pontos representam locais de medição e setas direção do fluxo.  $N = 13$  sites. Sistema referente à Equação (3.8).



- $z_{in} = [2\ 2\ 2\ 2\ 2\ 0\ 0\ 0\ 1\ 0\ 1\ 0\ 0]$ , Figura 12.a); e
- $z_{in} = [5\ 3\ 1\ 4\ 6\ 0\ 0\ 0\ 2\ 0\ 3\ 0\ 0]$ , Figura 12.b).

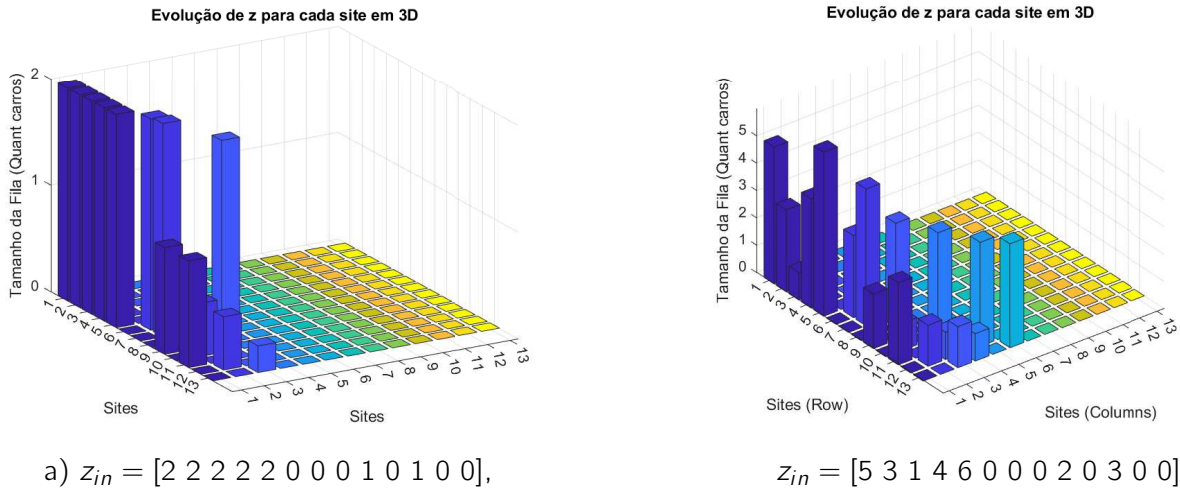


Figura 12 – Gráfico do resultado da simulação de  $L^{(\ell)}[\mathbf{z}(t)]$  com  $N = 13$  sites e diferentes  $z_{in}$ 's. Referente à Equação (3.8).

### 4.3 Coleta de Dados em Vias Reais

A precisão e a capacidade de generalização de modelos de previsão de séries temporais, como AR, MA, ARMA, ARIMA, ARIMA e STARIMA estão intrinsecamente ligadas à qualidade, densidade e continuidade dos dados de entrada. Esses modelos dependem do aprendizado de dependências temporais e espaciais latentes incorporadas às sequências observadas, que só podem ser efetivamente capturadas se a série temporal for suficientemente longa e amostrada com granularidade adequada. Valores ausentes ou conjuntos de dados escassamente amostrados podem degradar a estimativa de parâmetros, reduzindo assim o desempenho da previsão e levando a previsões não confiáveis [31].

Essa dependência de dados é igualmente crítica em modelos baseados em aprendizado profundo, como redes LSTM. Projetadas para modelar dependências temporais não lineares de longo alcance, as LSTMs são particularmente eficazes quando treinadas em taxas de dados densas e de alta amostragem que refletem dinâmicas de granularidade fina, como variações no fluxo de tráfego que ocorrem em curtos intervalos de tempo. Em contraste, dados com baixa amostragem limitam a capacidade do modelo de extrair padrões significativos, especialmente em tarefas de previsão de tráfego de curto prazo, onde alta resolução temporal é essencial.

Para apoiar a avaliação abrangente do modelo, dois conjuntos de dados de tráfego do mundo real complementares foram empregados neste estudo:

**Conjunto de Dados 1:** Um conjunto de dados de tráfego de alta resolução coletados por sensores de radar ao longo do Corredor Norte-Sul em São Paulo, Brasil, particularmente na Avenida 23 de Maio, durante o ano de 2018 [4]. Este conjunto de dados oferece uma taxa de amostragem densa e excelente continuidade, tornando-o adequado para comparar o desempenho de modelos de previsão clássicos e de aprendizado profundo sem a necessidade de pré-processamento ou reamostragem.

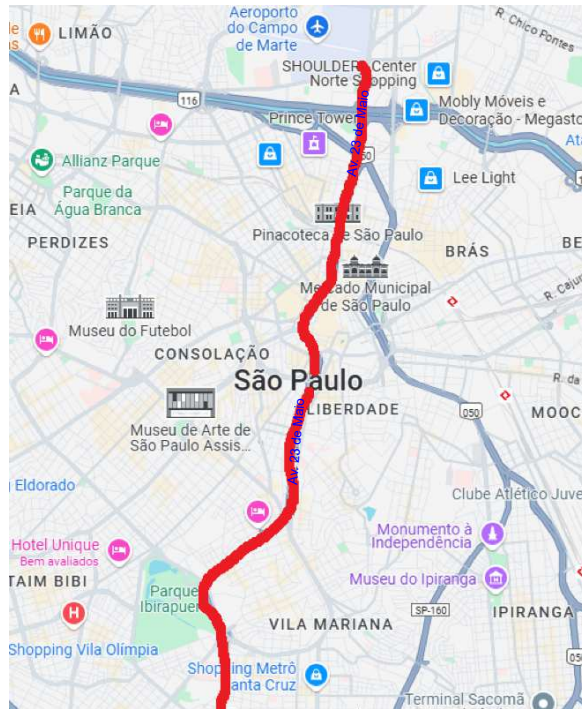


Figura 13 – **(Conjunto de dados 1) São Paulo:** Sensores de radar posicionados ao longo da Avenida 23 de Maio, parte do Corredor Norte-Sul no centro da cidade. Esta via expressa urbana é uma via arterial importante, com alto volume de tráfego e granularidade temporal.

**Conjunto de dados 2:** Um conjunto de dados de tráfego de longo prazo fornecido pelo Departamento de Estradas de Rodagem do Distrito Federal, publicado no Portal de Dados Abertos de Brasília em fevereiro de 2025.<sup>1</sup> O conjunto de dados compreende contagens de veículos amostradas a cada 15 minutos durante quatro dias consecutivos, abrangendo trechos de rodovias estratégicas em Brasília, como a DF-027 (EPJK), a DF-035 (EPCV) e a DF-463, conforme documentado no relatório oficial do Sistema Rodoviário do Distrito Federal [32].

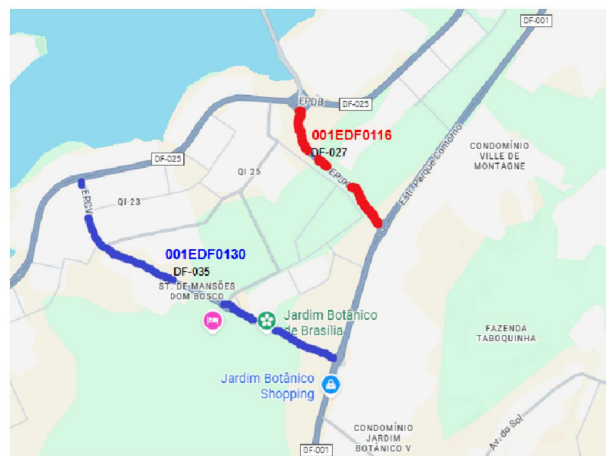


Figura 14 – **(Conjunto de Dados 2) Brasília:** Dados de sensores coletados nos trechos rodoviários DF-027 (EPJK), DF-035 (EPCV) e DF-463 localizados no setor sul do Distrito Federal. Essas vias são oficialmente classificadas como duplicadas e pavimentadas, o que permite alto fluxo de veículos.

Apesar de sua resolução temporal mais grosseira, o Conjunto de Dados 2 oferece dados reais de longa

<sup>1</sup> <http://dados.df.gov.br/dataset>

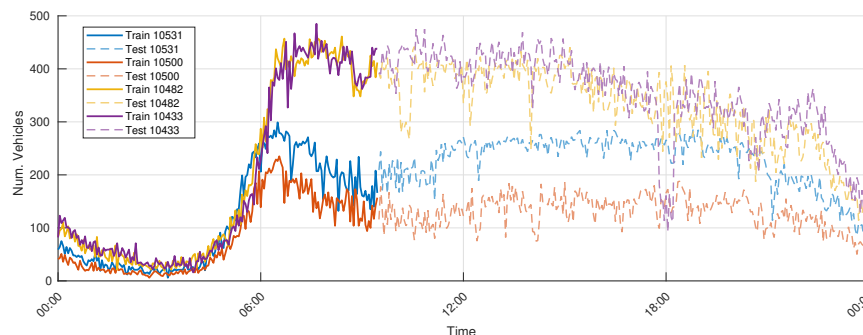


Figura 15 – Divisão de treinamento/teste para o Conjunto de Dados 1 (São Paulo). Os dados compreendem contagens de tráfego baseadas em radar de alta frequência, amostradas em intervalos de menos de um minuto. O conjunto de treinamento abrange aproximadamente 70% da sequência total, com os 30% restantes alocados para testes.

duração, permitindo a análise de tendências de tráfego diárias e de vários dias em segmentos rodoviários estruturalmente significativos da malha viária de Brasília. Esses segmentos, como a DF-027 (Juscelino Kubitschek), conectando a DF-035 e a DF-463, são classificados como corredores arteriais duplicados e pavimentados (DUP, PAV), apoiando o alto fluxo de veículos e o planejamento da mobilidade urbana.

Esses conjuntos de dados, em conjunto, fornecem uma estrutura complementar e balanceada, permitindo-nos avaliar métodos de previsão de tráfego urbano baseados em STARIMA e LSTM. Enquanto o Conjunto de Dados-1 facilita o aprendizado da dinâmica de curto prazo com alta fidelidade temporal, o Conjunto de Dados-2 desafia os modelos a generalizarem em horizontes mais longos e em padrões estruturais de tráfego do sistema arterial de Brasília.

#### 4.3.1 Treinamento e Teste em Modelos Paramétricos e de Aprendizagem

Para garantir o treinamento e a avaliação confiáveis do modelo, cada conjunto de dados foi dividido em subconjuntos de treinamento e teste. Em geral, o conjunto de treinamento é a porção dos dados usada para ajustar o modelo (ou seja, estimar os parâmetros internos que melhor capturam a dinâmica temporal e espacial subjacente da série temporal observada). Para modelos paramétricos como o STARIMA, isso envolve a identificação de dependências de defasagem e a estimativa de coeficientes AR e de MA. Por outro lado, em modelos de aprendizado profundo como o LSTM, o processo de treinamento ajusta os pesos da rede por meio de otimização iterativa baseada em retropropagação.

Por outro lado, o conjunto de teste consiste em dados não observados usados para avaliar a capacidade de previsão do modelo. Ele serve para avaliar o desempenho da generalização, ou seja, a precisão com que o modelo treinado prevê valores futuros com base em padrões aprendidos. A separação dos dados de treinamento e teste é essencial para evitar *overfitting* e garantir que o desempenho relatado reflita a verdadeira capacidade preditiva, em vez da memorização da sequência de entrada.

Para o conjunto de dados de radar de alta taxa de amostragem de São Paulo (Conjunto de Dados 1), 40% da sequência de dados foi usada para treinamento e os 60% restantes para testes. Essa divisão permite que o modelo aprenda com um fluxo contínuo de dados temporais densos, representativos dos padrões reais de fluxo de tráfego. Para o conjunto de dados de Brasília (Conjunto de Dados 2), foram adotadas três divisões diferentes. Especificamente, para a previsão de médio prazo, o conjunto de dados foi dividido de forma que os dois primeiros dias fossem usados para treinamento e os dois últimos para testes. Para a previsão de longo prazo, foram avaliados períodos de treinamento de 2 e 4 dias, seguidos de testes nos dias restantes. Essa divisão balanceada permite a avaliação da robustez do modelo em múltiplos ciclos diários, respeitando a ordem temporal dos eventos.

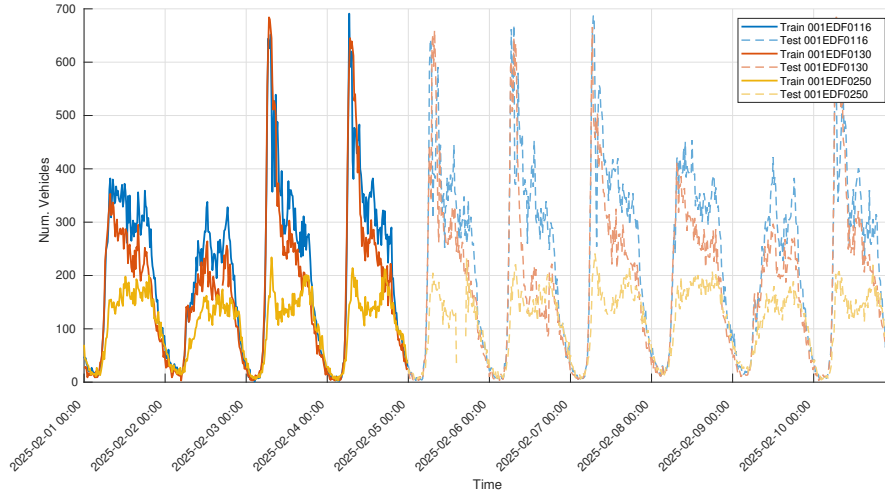


Figura 16 – Divisão de treinamento/teste para o conjunto de dados **Dataset 2** (Brasília). Os dados consistem em contagens de veículos agregadas a cada 15 minutos ao longo de quatro dias (fevereiro de 2025). Os três primeiros dias formam o conjunto de treinamento, enquanto o último dia é reservado para testes.

## 4.4 Métricas de Desempenho

Nesta seção, são apresentadas e discutidas as principais métricas utilizadas para avaliar o desempenho dos modelos de previsão. Essas métricas permitem quantificar o quão próximos os valores estimados estão dos valores reais, permitindo promover análises comparativas entre as abordagens paramétricas dos modelos STARIMA e LSTM. Entre as métricas consideradas estão o Erro Quadrático Médio Normalizado, o Erro Quadrático Médio, o Erro Absoluto Médio, o Erro Absoluto Percentual Médio e o Coeficiente de Eficiência.

### 4.4.1 Erro Quadrático Médio Normalizado (NMSE)

O erro quadrático médio normalizado (NMSE - *Normalized Mean Square Error*) é uma métrica útil para avaliar a precisão de previsões feitas por modelos preditivos. Ele normaliza o erro quadrático médio pela variância dos dados reais, fornecendo uma medida relativa do desempenho do modelo. O NMSE é calculado por:

$$\text{NMSE} = \frac{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}{\frac{1}{N} \sum_{i=1}^N y_i^2} = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N y_i^2}, \quad (4.4)$$

onde  $N$  é número de observações,  $y_i$  representa valores reais e  $\hat{y}_i$  são os valores previstos pelo modelo.

### 4.4.2 Raiz do Erro Quadrático Médio (RMSE)

O RMSE é uma métrica amplamente utilizada para medir a precisão de um modelo de previsão. representa a raiz quadrada da média dos quadrados dos erros, onde os erros são as diferenças entre os valores observados e os valores previstos pelo modelo. Matematicamente, o RMSE é definido como:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \quad (4.5)$$

sendo  $N$  o número de observações,  $y_i$  são os valores observados e  $\hat{y}_i$  são os valores previstos.

A métrica RMSE é útil em diversos contextos, destacando-se pela sua sensibilidade a grandes erros. Diferentemente de outras métricas como MAE, que mede a média das diferenças absolutas entre os valores observados e previstos, o RMSE dá maior peso a grandes desvios devido à operação de elevação ao quadrado

dos erros antes de calcular a média. Portanto, o RMSE é particularmente adequado para situações em que grandes erros são especialmente indesejados ou críticos [30].

#### 4.4.3 Erro Médio Absoluto (MAE)

O MAE é uma métrica amplamente utilizada para avaliar a precisão de modelos de previsão. Ele representa a média das diferenças absolutas entre os valores observados e os valores previstos pelo modelo [29]. Matematicamente, o MAE é definido como:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|. \quad (4.6)$$

A métrica MAE é especialmente útil em contextos onde se deseja uma medida simples e direta da precisão de um modelo. Ao contrário do RMSE, que dá maior peso a grandes erros devido à elevação ao quadrado das diferenças, o MAE trata todos os erros de forma linear, o que pode ser mais intuitivo em muitas aplicações práticas [30] [29].

Além disso, o MAE é menos sensível a *outliers* em comparação com o RMSE, o que o torna uma escolha preferida em situações onde grandes erros não são necessariamente mais importantes que erros menores. No entanto, deve-se considerar que o MAE, ao tratar todos os erros de forma linear, pode não capturar a gravidade de grandes desvios da mesma maneira que o RMSE.

Portanto, a escolha entre MAE e RMSE depende do contexto e dos objetivos específicos da análise. Utilizar ambas as métricas em conjunto pode fornecer uma visão mais abrangente da performance do modelo.

#### 4.4.4 Erro Absoluto Percentual Médio (MAPE)

O Erro Absoluto Percentual Médio (MAPE - *Mean Absolute Percentage Error*) é uma métrica amplamente utilizada para medir a precisão de modelos de previsão. Ele representa a média das diferenças absolutas entre os valores observados e os valores previstos pelo modelo, expressas como uma porcentagem dos valores observados. Matematicamente, o MAPE é definido como:

$$\text{MAPE} = \frac{100}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|. \quad (4.7)$$

A métrica MAPE é particularmente útil em contextos onde a interpretação dos erros em termos percentuais é mais intuitiva e significativa. Por exemplo, ao prever vendas ou demanda de produtos, a porcentagem de erro pode ser mais fácil de comunicar e entender em termos de impacto no negócio [29] [30].

Entretanto, é importante notar que o MAPE pode ser distorcido por valores muito pequenos de  $y_i$ , resultando em valores extremamente altos de porcentagem de erro. Portanto, sua interpretação deve ser feita com cautela em datasets que incluem valores próximos de zero. Para garantir uma avaliação mais robusta da performance do modelo, é comum complementar a análise com outras métricas, como MAE e RMSE, proporcionando uma visão mais completa dos diferentes aspectos dos erros de previsão.

#### 4.4.5 Coeficiente de Eficiência

O Coeficiente de Eficiência (EC), também conhecido como Coeficiente de Nash – Sutcliffe, é uma métrica utilizada para avaliar a acurácia de um modelo de previsão. Este coeficiente quantifica o quão bem os valores previstos por um modelo concordam com os valores observados. O Coeficiente de Eficiência é dado por:

$$\text{EC} = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}, \quad (4.8)$$

sendo:

$O_j$ : representa os valores observados.

$P_j$ : os valores previstos.

$\bar{O}$ : indica a média dos valores observados.

Interpretação do valor de EC:

- O coeficiente EC pode assumir valores entre  $-\infty$  e 1.
- Quanto maior o valor de EC, melhor a concordância entre o modelo e os valores observados.
- Um valor de EC=1 significa que há uma concordância total entre os valores preditos e os observados.
- Um valor de EC=0 significa que o modelo representa os dados tão bem quanto a média dos valores observados.
- Valores negativos de EC indicam que a média dos valores observados seria melhor do que o modelo aplicado para realizar a previsão.

# 5 RESULTADOS E DISCUSSÃO

Neste Capítulo, a partir de exemplos de séries temporais disponíveis na literatura, é corroborado os procedimentos de extração dos parâmetros e a respectiva validação da ordem do modelo paramétrico a partir da aplicação das principais métricas utilizadas na análise de desempenho dos modelos de predição. Também é utilizada séries temporais associadas ao tráfego urbano para validar o modelo paramétrico proposto.

As seções a seguir apresentam resultados de simulações independentes. Cujos objetivos são enfatizar a grade aplicabilidade dos modelos e resultados desenvolvidos e demonstrados.

## 5.1 Geração de Dados Semissintéticos

A robustez e a precisão de modelos de previsão de séries temporais, como os modelos Autorregressivo (AR), de Médias Móveis (MA), ARMA, ARIMA e, notadamente, o modelo STARIMA, são intrinsecamente dependentes da densidade e da qualidade dos dados de entrada. Estes modelos aprendem os padrões de dependência temporal e espacial a partir de sequências de dados contínuas e suficientemente longas. A presença de dados faltantes ou a utilização de uma base de dados esparsa pode levar a uma estimação de parâmetros inadequada, resultando em previsões com baixa acurácia e confiabilidade.

Para mitigar o desafio imposto por conjunto de dados (datasets) de tráfego esparsos, esta dissertação emprega uma metodologia para a geração de dados semissintéticos. O objetivo é adensar a série temporal, ou seja, aumentar sua granularidade, sem, contudo, sacrificar o realismo estatístico dos dados originais. A abordagem se baseia na desagregação de dados de baixa frequência (e.g., contagens a cada 15 minutos) para uma frequência mais alta (e.g., contagens a cada 3 minutos), utilizando um processo estocástico fundamentado na distribuição de Poisson.

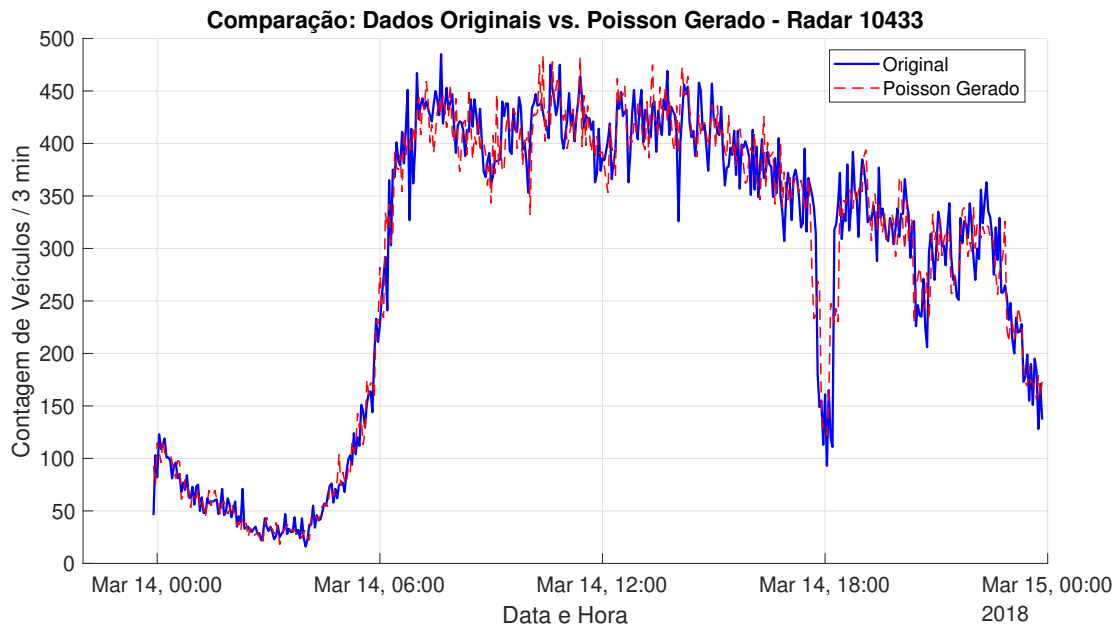
A escolha da distribuição de Poisson é teoricamente justificada, pois a contagem do número de veículos que passam por um determinado ponto em um intervalo de tempo fixo é um exemplo clássico de um processo de Poisson. Este processo modela eventos discretos que ocorrem de forma independente a uma taxa média constante,  $\lambda$ . A principal inovação da metodologia aplicada é que a taxa  $\lambda$  não é considerada constante ao longo de todo o dia, mas sim localmente adaptativa. O procedimento é executado da seguinte forma:

1. **Agregação Local:** Os dados originais de baixa frequência (15 minutos) são primeiramente adaptados em intervalos de alta frequência (a cada 3 minutos). A soma total de veículos neste intervalo é denominada  $C_{15}$ .
2. **Definição da Taxa Poisson Local:** Para cada intervalo de 15 minutos, a contagem agregada  $C_{15}$  é usada como a média esperada de eventos para aquele período.
3. **Geração de Dados Sintéticos:** Para gerar os 5 pontos de dados de 3 minutos dentro do referido intervalo, utiliza-se uma distribuição de Poisson onde o parâmetro de taxa,  $\lambda$ , para cada ponto de 3 minutos é definido como  $\lambda = C_{15}/5$ .

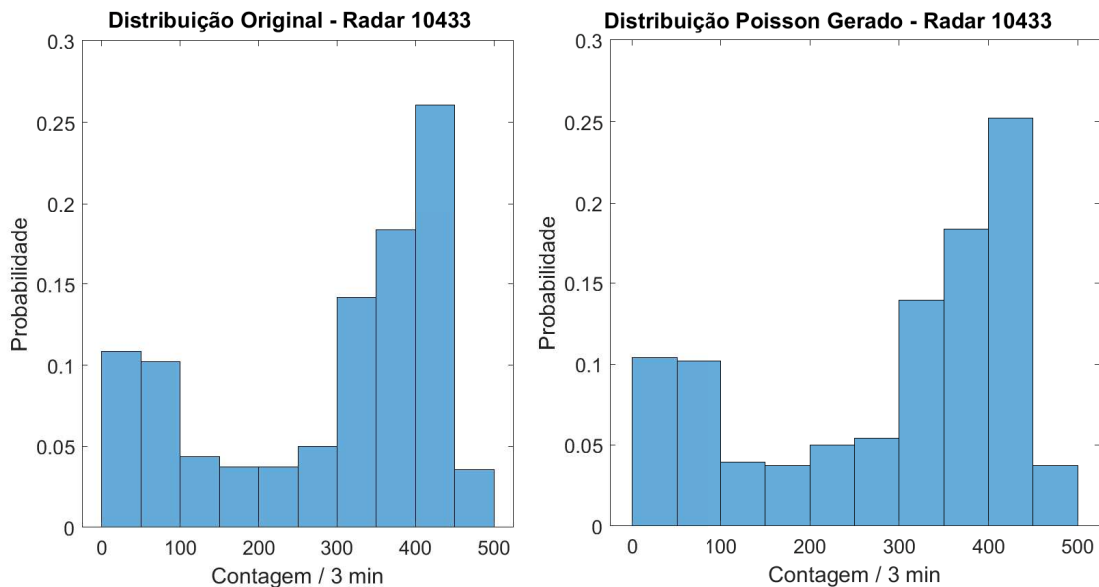
Esta abordagem local garante que as flutuações de grande escala do tráfego, como os picos de congestionamento matutinos e vespertinos, sejam preservadas, uma vez que a geração de dados sintéticos é sempre ancorada por uma medição real ( $C_{15}$ ).

### 5.1.1 Exemplo de Validação

Para validar a eficácia desta metodologia, foi realizado um experimento utilizando um conjunto de dados completo do Radar 10433. A partir dos dados originais, foram gerados dados semissintéticos seguindo o procedimento descrito. A Figura 17 apresenta a comparação visual entre a série temporal original e a série gerada.



a) Dados originais vs. Gerados por Poisson / Sobreposição das séries temporais.



b) Comparação das séries temporais

Figura 17 – Em a) Comparação visual entre os dados originais (em azul) e os dados semissintéticos gerados via Poisson (em vermelho) para o Radar 10433. E a sobreposição das séries temporais. A seguir, b) a comparação das suas respectivas distribuições de probabilidade.

A inspeção visual da Figura 17 revela uma notável semelhança entre os dados originais e os gerados. Tanto a série temporal (17.a) quanto a distribuição de probabilidade (17.b) dos dados sintéticos acompanham de perto as características dos dados reais.

Para uma avaliação objetiva, foram calculadas métricas de erro que quantificam a aderência do modelo. Os resultados foram os seguintes:

- **RMSE (Root Mean Square Error):** 29,02, indicando o erro quadrático médio entre os valores.

- **MAE (Mean Absolute Error):** 20,91, representando o erro absoluto médio.
- **NRMSE (Normalized RMSE):** 10,13%, que normaliza o RMSE pela média dos dados originais, mostrando um erro relativo baixo.
- **NMAE (Normalized MAE):** 7,30%, que normaliza o MAE, também indicando uma excelente aderência.

Adicionalmente, a média da diferença entre as somas agregadas a cada 15 minutos foi de apenas -1,6 veículos, um valor muito próximo de zero, o que demonstra que o método não introduz um viés sistemático de superestimação ou subestimação. Estes resultados confirmam que a metodologia de geração de dados semi-sintéticos é capaz de produzir séries temporais densas e realistas, preservando tanto as tendências macroscópicas quanto as características estatísticas locais. Portanto, constitui uma ferramenta valiosa para o enriquecimento de *datasets* esparsos, viabilizando a aplicação de modelos de previsão mais exigentes como o STARIMA.

### 5.1.2 Estudo de Caso: Adensamento de Dados da CET

Para demonstrar a aplicação prática e a relevância da metodologia de geração de dados, utiliza-se como estudo de caso uma base de dados real fornecida pela Companhia de Engenharia de Tráfego (CET) de São Paulo. O padrão de coleta deste sistema disponibiliza as contagens de tráfego em intervalos de 15 minutos, como ilustrado na Figura 18, que exibe um extrato dos dados para a Av. Prestes Maia.

a) Rota 04G - Av. Zaki Narchi, Av. Santos Dumont, Av. Tiradentes, Av. Prestes Maia, Fonte: [3]

Ponto 1 \* Av. Prestes Maia, entre R. Paula Souza e R. Mauá

Hora	Pico Manhã						Pico Tarde											
	Auto		Ônibus		Caminhões	Motos	Volume Total		FHP	Auto		Ônibus		Caminhões	Motos	Volume Total		FHP
	Urb	Fret	Urb	Fret			Simp.	Equiv.		Urb	Fret	Simp.	Equiv.					
07:00 - 08:00	4.015	57	12	14	468	4.566	4.649	0,94	17:00 - 18:00	3.836	61	13	14	1.145	5.069	5.157		
07:15 - 08:15	4.005	62	14	15	495	4.591	4.682	0,94	17:15 - 18:15	3.650	61	11	11	1.186	4.919	5.002		
07:30 - 08:30	3.910	67	14	13	513	4.517	4.611		17:30 - 18:30	3.532	58	8	9	1.269	4.876	4.951		
07:45 - 08:45	3.794	70	18	11	498	4.391	4.490		17:45 - 18:45	3.359	56	8	10	1.307	4.740	4.814		
08:00 - 09:00	3.629	69	18	11	488	4.215	4.313		18:00 - 19:00	3.156	59	7	10	1.269	4.501	4.577		
08:15 - 09:15	3.439	65	17	11	481	4.013	4.106		18:15 - 19:15	3.455	57	8	10	1.138	4.668	4.743		
08:30 - 09:30	3.390	62	16	11	487	3.966	4.055		18:30 - 19:30	4.066	66	9	9	990	5.140	5.224		
08:45 - 09:45	3.321	61	13	14	507	3.916	4.004		18:45 - 19:45	4.630	68	10	9	842	5.559	5.646		
09:00 - 10:00	3.304	63	12	14	526	3.919	4.008		19:00 - 20:00	5.365	72	13	9	758	6.217	6.311	0,96	
Total 3 horas	10.948	189	42	39	1.482	12.700	12.970	0,92	Total 3 horas	12.357	192	33	33	3.172	15.787	16.045	0,85	

Apesar da riqueza categórica dos dados, que diferenciam o fluxo entre carros, ônibus e motos, a esparsidade temporal, ou seja, a frequência de amostragem de apenas quatro pontos por hora, representa um forte empecilho (*drawback*) para a modelagem com técnicas como o STARIMA. Tais modelos se beneficiam de séries temporais mais densas para capturar com maior fidelidade as dinâmicas de tráfego de curto prazo.

Diante deste cenário, a técnica de geração de dados semi-sintéticos, validada na seção anterior, será aplicada a esta base de dados. O objetivo é adensar a série temporal, convertendo os intervalos de 15 minutos para intervalos de 3 minutos, e posteriormente observar o comportamento da estimação do modelo com os dados enriquecidos.

A análise dos resultados apresentados na Figura 20, evidencia uma forte correspondência entre os dados gerados pelo modelo de Poisson e as médias originais obtidas experimentalmente para os três experimentos estudados, tanto no período da manhã quanto no da tarde. Observa-se que, em ambos os intervalos, as séries simuladas acompanham de perto as oscilações e tendências das séries originais, validando a adequação do modelo probabilístico para representar a dinâmica observada. Notam-se variações mais acentuadas para os experimentos RadarExp1 e RadarExp2, especialmente no período da tarde, quando há um incremento dos valores ao longo do tempo. Já o experimento RadarExp3 mantém valores inferiores e apresenta

b) Dados do Ponto 1 de Contagem Volumétrica Classificada, Fonte: [3]

Ponto 2 Av. Tiradentes, entre R. João Teodoro e R. Jorge Miranda

Pico Manhã								Pico Tarde									
Hora	Auto	Ônibus		Caminhões	Motos	Volume Total		FHP	Hora	Auto	Ônibus		Caminhões	Motos	Volume Total		FHP
		Urb	Fret			Simp.	Equiv.				Urb	Fret			Simp.	Equiv.	
07:00 - 08:00	3.579	77	8	16	416	4.096	4.197		17:00 - 18:00	4.149	59	16	11	1.227	5.462	5.548	
07:15 - 08:15	3.548	87	9	14	429	4.087	4.197		17:15 - 18:15	3.994	61	14	11	1.220	5.300	5.386	
07:30 - 08:30	3.453	126	10	14	433	4.036	4.186		17:30 - 18:30	3.586	52	12	8	1.383	5.041	5.113	
07:45 - 08:45	3.522	132	16	9	440	4.119	4.276	0,98	17:45 - 18:45	3.244	47	12	9	1.325	4.637	4.705	
08:00 - 09:00	3.410	153	15	8	445	4.031	4.207		18:00 - 19:00	3.405	57	9	11	1.271	4.753	4.830	
08:15 - 09:15	3.290	151	17	7	446	3.911	4.086		18:15 - 19:15	3.487	55	10	10	1.171	4.733	4.808	
08:30 - 09:30	3.329	107	17	6	476	3.935	4.065		18:30 - 19:30	4.123	64	15	11	931	5.144	5.234	
08:45 - 09:45	3.203	120	13	11	486	3.833	3.977		18:45 - 19:45	4.707	65	16	10	859	5.657	5.748	
09:00 - 10:00	3.223	107	14	9	516	3.869	3.999		19:00 - 20:00	5.171	64	19	8	763	6.025	6.116	0,91
Total 3 horas	10.212	337	37	33	1.377	11.996	12.403	0,97	Total 3 horas	12.725	180	44	30	3.261	16.240	16.494	0,90

c) Dados do Ponto 2 de Contagem Volumétrica Classificada, Fonte: [3]

Ponto 3 Av. Santos Dumont, entre Pça. Campo de Bagatelle e R. Santa Eulália

Pico Manhã								Pico Tarde									
Hora	Auto	Ônibus		Caminhões	Motos	Volume Total		FHP	Hora	Auto	Ônibus		Caminhões	Motos	Volume Total		FHP
		Urb	Fret			Simp.	Equiv.				Urb	Fret			Simp.	Equiv.	
07:00 - 08:00	1.294	0	64	17	171	1.546	1.627		17:00 - 18:00	3.321	2	53	44	660	4.080	4.179	
07:15 - 08:15	1.319	0	54	18	189	1.580	1.652		17:15 - 18:15	3.341	2	73	28	694	4.138	4.241	
07:30 - 08:30	1.355	0	45	17	196	1.613	1.675		17:30 - 18:30	3.252	2	73	20	756	4.103	4.198	
07:45 - 08:45	1.411	1	37	17	186	1.652	1.707		17:45 - 18:45	3.072	1	74	22	755	3.924	4.021	
08:00 - 09:00	1.348	2	35	14	178	1.577	1.628		18:00 - 19:00	3.113	0	77	20	751	3.961	4.058	
08:15 - 09:15	1.350	2	36	11	192	1.591	1.640		18:15 - 19:15	3.202	0	68	19	690	3.979	4.066	
08:30 - 09:30	1.365	2	38	15	202	1.622	1.677		18:30 - 19:30	3.478	0	67	15	601	4.161	4.243	
08:45 - 09:45	1.336	1	42	19	225	1.623	1.685		18:45 - 19:45	3.617	0	56	13	526	4.212	4.281	0,93
09:00 - 10:00	1.390	0	35	22	230	1.677	1.734	0,95	19:00 - 20:00	3.504	0	54	13	438	4.009	4.076	
Total 3 horas	4.032	2	134	53	579	4.800	4.989	0,96	Total 3 horas	9.938	2	184	77	1.849	12.050	12.313	0,96

d) Dados do Ponto 3 de Contagem Volumétrica Classificada, Fonte: [3]

Figura 18 – Dados de Velocidade: Tráfego Geral e Volume: Pontos de Contagem Volumétrica.

Rota 04G - Av. Zaki Narchi, Av. Santos Dumont, Av. Tiradentes, Av. Prestes Maia

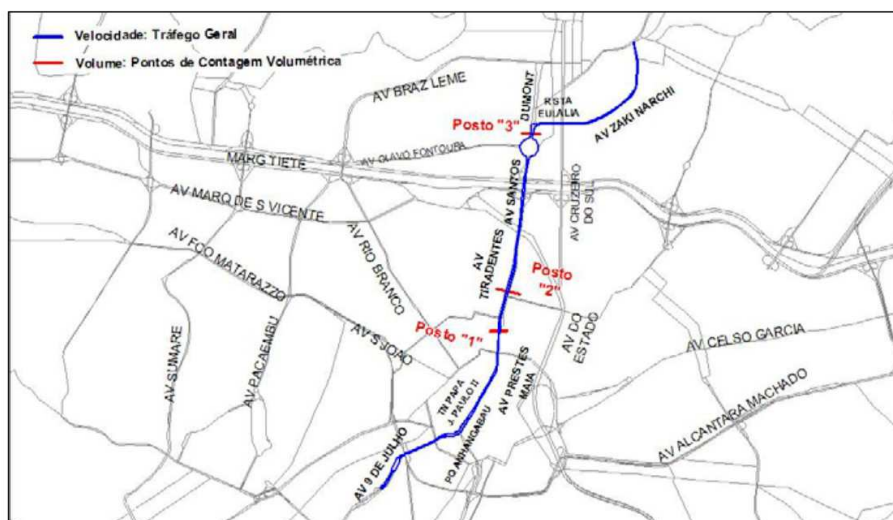


Figura 19 – Rota 04G - Av. Zaki Narchi, Av. Santos Dumont, Av. Tiradentes, Av. Prestes Maia, Fonte: [3]

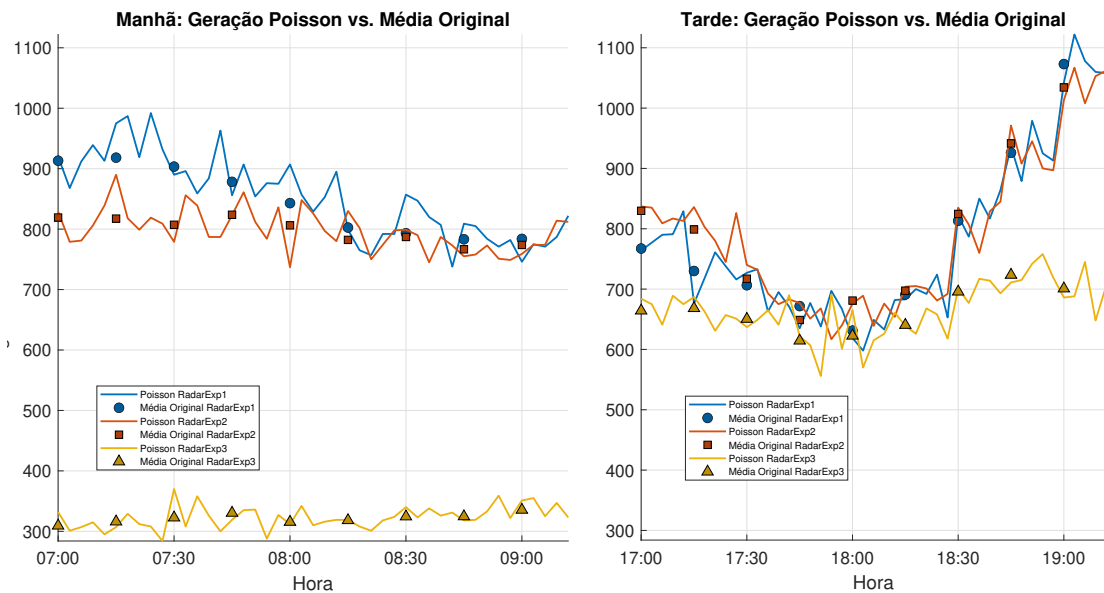


Figura 20 – Comparação entre a geração de dados via processo de Poisson e a média original observada em três experimentos distintos, nos períodos da manhã (07:00–09:00, à esquerda) e da tarde (17:00–19:00, à direita). Cada linha representa a evolução temporal dos valores para cada experimento, permitindo observar a proximidade entre a simulação (Poisson) e os dados reais (Média Original) durante os dois intervalos analisados.

menor dispersão, reforçando padrões distintos de comportamento entre os experimentos. Esses resultados demonstram que a simulação via Poisson é capaz de capturar tanto o perfil médio quanto as flutuações características dos dados reais, o que corrobora a utilização do modelo para futuras análises preditivas e comparativas.

## 5.2 Conjunto de Dados Hipotéticos Gerados a Partir de Modelos ARMA(p,q) e Extração dos Parâmetros com ARIMA(p,D,q)

Esta Seção é composta por conjuntos de dados hipotéticos (séries temporais) gerados a partir dos modelos:

- ◇ ARMA(4,3) não estacionário de ordem  $D = 1 \equiv \text{ARIMA}(4,1,3)$
- ◇ ARMA(4,3) = ARIMA(4,0,3)

A Figura 21 apresenta dois gráficos relacionados à predição da série temporal sintética, com 10k amostras plotadas na Figura 21.a) utilizando um modelo autorregressivo de média móvel (ARMA – *Auto Regressive Moving Average*).

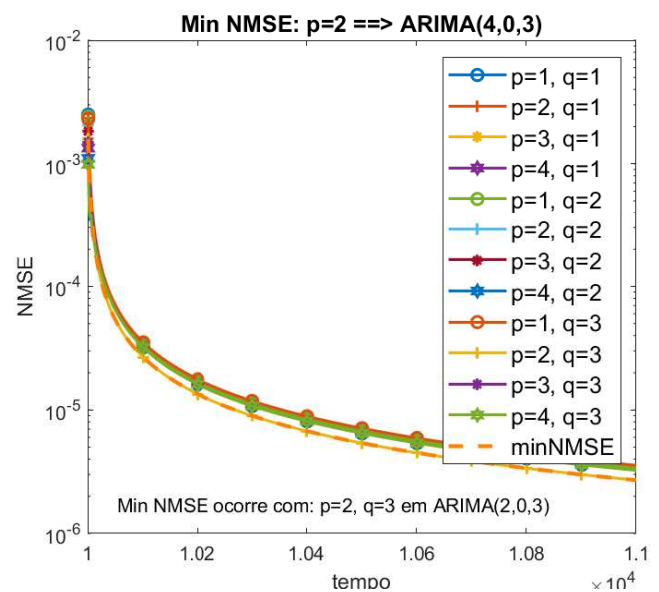
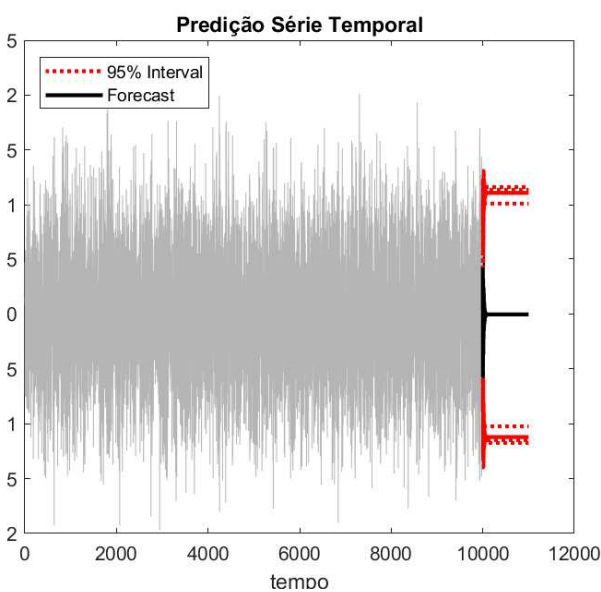
De fato, o gráfico da Figura 21.a) mostra a previsão da série ao longo de 1k amostras de previsão, precedida pelas 10k amostras da série temporal. O valor médio da previsão é representada pela linha preta contínua, e a incerteza associada a essa previsão é mostrada pela faixa vermelha pontilhada, que representa o intervalo de confiança de 95%.

Na Figura 21.b) conclui-se que o menor NMSE ocorre com  $p=2$ ,  $D=0$  e  $q=3$ . A linha tracejada em cor laranja, reforça a visualização para o apontamento da melhor curva desta simulação com os parâmetros adotados. Para confirmar o bom ajuste dos parâmetros  $p$  e  $q$  obtidos, apresenta-se na Figura 21.c) um zoom para o intervalo [9850; 10000] para a série temporal e as respectivas curvas de predição obtidas para três diferentes combinações de parâmetros  $p$  e  $q$ .

A Tabela 3 apresenta os respectivos valores dos parâmetros AR e MA ajustados/preditos em comparação aos valores originais adotados na geração da série temporal original, adotando-se o modelo ARMA, Equação (3.3). Ainda na Tabela 3 comparamos os coeficientes estimados (extraídos do simulador de modelos paramétricos) aos valores originais. Indicamos os respectivos NMSE para as amostras no intervalo de predição.

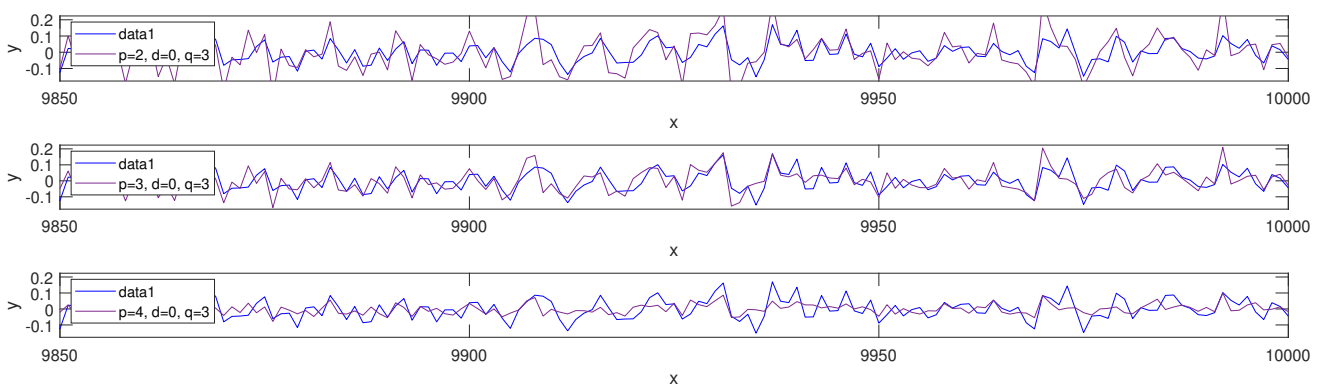
Tabela 3 – Comparação dos coeficientes do modelo original com os ajustados e respectivos NMSE

Modelo $AR\{p\}, D\{d\}, MA\{q\}$	$\sigma_\epsilon^2$	Modelo Ajuste	Coefficientes Ajustados	NMSE
$ARMA\{[.25, -.62, .31, -.71]; [.2, .72, .43]\}, D=0$	0.001	$ARMA(4,0,3)$	$\hat{\beta}^{ar} = [.237, -.624, .300, -.701];$ $\hat{\beta}^{ma} = [.216, .721, .442]$	0.0013
$ARIMA\{[.25, -.62, .31, -.71]; [.2, .72, .43]\}, D=1$	0.001	$ARIMA(4,1,3)$	$\hat{\beta}_1^{ar} = [.237, -.624, .300, -.701];$ $\hat{\beta}_1^{ma} = [.216, .721, .442]$	0.0043



a) 10K amostras seguida pela predição sobre 1000 amostras

b) NMSE para diferentes  $p$  e  $q$ .



c) Predição em 150 amostras considerando três modelos:  $ARMA(2,3)$ ,  $ARMA(3,3)$  e  $ARMA(4,3)$

Figura 21 – Obtenção dos parâmetros  $p$  e  $q$  no modelo  $ARIMA(4,0,3)$  para amostras seguindo Distribuição Gaussiana. MatLab R2024a.

### 5.2.1 Determinação da Ordem AR( $p$ ), MA( $q$ ), de Integração (D) no Modelo STARIMA

A determinação das ordens do modelo STARIMA segue uma metodologia iterativa, semelhante à abordagem de Box-Jenkins para modelos ARIMA [33]. Este processo geralmente envolve três etapas principais: identificação, estimação e diagnóstico. A maior acurácia e eficiência na determinação das ordens vêm da aplicação rigorosa e iterativa destas etapas. O ciclo iterativo é feito nos três estágios iterativos da modelagem paramétrica (identificação, estimação e diagnóstico), até determinarmos um modelo que seja aprovado na fase de diagnóstico, isto é, até que os resíduos sejam considerados ruído branco e todos os coeficientes tenham valores significativos.

Um método para determinar as ordens ( $p, d, q$ ) das componentes temporais e ( $\lambda, m$ ) das componentes espaciais do modelo STARIMA atende a três fases:

#### 1. **Determinação da Ordem de Integração ( $d$ ):**

- ◇ A ordem  $d$  representa o número de vezes que a série temporal precisa ser diferenciada para se tornar estacionária. Uma série estacionária tem propriedades estatísticas (média, variância) que não mudam ao longo do tempo (média) ou que sejam apenas função da diferença temporal (segundo momento estatístico, correlação).
- ◇ A necessidade de diferenciação pode ser verificada inicialmente por inspeção visual da série temporal, observando variações que indicam uma média não constante. Um dos pré-requisitos para aplicar o modelo (ST)ARIMA é que a série temporal seja estacionária. Isso significa que as propriedades estatísticas da série, como média e variância, não devem mudar ao longo do tempo. Para verificar a estacionariedade, podem ser utilizados testes como o teste de Dickey-Fuller Aumentado. Se a série não for estacionária, transformações como diferenciação podem ser aplicadas para atender a esse requisito.
- ◇ O método mais preciso para verificar a estacionariedade é utilizar testes estatísticos. O teste *Augmented Dickey-Fuller* (ADF) é comumente usado. Ele verifica se cada termo da série está linearmente relacionado ao seu termo anterior, testando se o coeficiente do termo defasado é zero.

#### 2. **Identificação Inicial das Ordens Temporais ( $p; q$ ) e Espaciais ( $\lambda; m$ ):** Esta etapa visa identificar uma classe de modelos que melhor se ajusta aos dados. Para modelos STARIMA, utilizam-se as Funções STACF e STPACF. Estas são análogas às ACF e Função de Autocorrelação Parcial (PACF) usadas para modelos ARMA univariados.

- ◇ **STPACF** é utilizada para determinar as ordens da parte autorregressiva (AR), tanto temporal ( $p$ ) quanto espacial ( $\lambda$ ). No gráfico da STPACF, o último lag temporal significativo sugere a ordem  $p$ , enquanto as linhas para diferentes defasagens espaciais ('L1', 'L2', etc.) indicam a ordem espacial  $\lambda$ .
- ◇ **STACF** é utilizada para determinar as ordens da parte de média móvel (MA), tanto temporal ('q') quanto espacial ('m'). No gráfico da STACF, o último lag temporal significativo na linha de defasagem espacial 0 ('L0') indica a ordem 'q', e as linhas para outras defasagens espaciais indicam a ordem espacial 'm'.
- ◇ **Limitação:** para processos mais complexos, como o STARMA (i.e., a versão não integrada do STARIMA), a visualização das ordens a partir dos gráficos STACF e STPACF pode se tornar

confusa, oferecendo apenas um direcionamento em vez de valores exatos para 'p' e 'q' <sup>1</sup>. Isso sugere faixas de valores para as ordens.

### 3. Seleção e Calibração do Modelo para Maior Acurácia e Eficiência:

- ◇ Após a identificação inicial que sugere modelos candidatos, a seleção das melhores ordens é refinada na etapa de calibração.
- ◇ A calibração envolve testar diferentes combinações de ordens (baseadas nas indicações dos gráficos STACF e STPACF).
- ◇ A qualidade do modelo para cada combinação de ordens é avaliada usando métricas de acurácia e critérios de informação. O objetivo é selecionar a combinação que otimiza essas métricas.
- ◇ Métricas de acurácia mencionadas incluem: MAE, RMSE, NMSE, EC [34] e MAPE. A calibração busca minimizar essas medidas de erro.
- ◇ Critérios de informação como o Critério de Informação Bayesiano (BIC) ou o Critério de Informação de Akaike (AIC) são utilizados para comparar modelos candidatos com diferentes ordens e selecionar o melhor. O modelo com o menor valor de BIC ou AIC é geralmente preferido. Para modelos ARIMA com diferentes ordens de diferenciação, o RMSE pode ser usado para comparação.
- ◇ A calibração busca a combinação de ordens que torna o modelo mais eficiente, considerando tanto a complexidade (menor número de parâmetros) quanto a acurácia.
- ◇ Após a seleção das ordens, os parâmetros do modelo são estimados, frequentemente usando métodos de otimização não-linear (já que modelos STARIMA com componente espacial são não-lineares). O algoritmo de Levenberg-Marquardt<sup>2</sup> [35] foi utilizado neste trabalho e em [34].
- ◇ A fase de diagnóstico verifica se os resíduos do modelo ajustado se comportam como ruído branco. Se os resíduos não forem ruído branco (verificado por testes estatísticos e análise das autocorrelações residuais) ou se os parâmetros não forem significativos, o processo iterativo (identificação, estimação, diagnóstico) é repetido, refinando as ordens e/ou a estrutura do modelo.

Portanto, com base no que foi detalhado, a metodologia mais eficiente e de maior acurácia para a determinação das ordens do modelo STARIMA (ou ASTARIMA) é um processo iterativo que combina:

1. Análise de estacionariedade (inspeção visual, teste ADF) para definir 'd'.
2. Identificação inicial baseada em STACF e STPACF para obter ordens candidatas ('p', 'q', 'λ', 'm'), reconhecendo que estas fornecem apenas direcionamento.
3. Calibração sistemática de diferentes combinações de ordens candidatas, avaliando o desempenho usando métricas de acurácia (MAE, RMSE, etc.) e critérios de informação (AIC, BIC) para selecionar o modelo que minimiza os erros e critérios.

<sup>1</sup> Ao operarmos com bancos de dados extensos (1000 *timelags* para cima), tanto STACF quanto STPACF indicam correta e claramente as ordens espaciais e temporais de AR e MA, para processos STAR e STMA. No entanto, para processos STARMA a visualização dessas ordens já se torna confusa, indicando que o método de identificação das ordens não é tão exato, dando apenas certo direcionamento. Isso se intensifica no caso do problema tratado neste trabalho, uma vez que – ao contrário dos dados utilizados para testes preliminares – os dados reais de tráfego urbano não seguem um comportamento totalmente autorregressivo ou de média móvel.

<sup>2</sup> O algoritmo de Levenberg-Marquardt (LMA) é um método de otimização iterativo usado para resolver problemas de mínimos quadrados não lineares [35]. O LMA é uma técnica iterativa que localiza o mínimo de uma função expressa como a soma dos quadrados de funções não lineares. É conhecido por sua convergência rápida quando aplicado corretamente, sendo amplamente utilizado em diversas áreas da ciência e engenharia.

4. Diagnóstico para garantir que os resíduos sejam ruído branco e os parâmetros sejam significativos, refinando o modelo conforme necessário.

A maior acurácia é alcançada pela avaliação empírica do desempenho de modelos candidatos com diferentes ordens, utilizando dados de validação ou critérios de informação, dentro deste ciclo iterativo.

A modelagem STARIMA exige a escolha adequada das ordens de defasagem temporal e espacial dos componentes autorregressivos (AR) e de média móvel (MA). Um recurso fundamental para essa escolha são os gráficos da Função STACF e STPACF.

Algumas orientações importantes para a correta definição das ordens incluem:

- Significância: Valores são considerados significativos quando ultrapassam as bandas de confiança indicadas nos gráficos, geralmente baseadas em um nível de 95%.
- Evitar Sobremodelagem: A escolha de ordens excessivamente altas pode aumentar a complexidade do modelo e prejudicar sua capacidade preditiva.
- Refinamento Posterior: Após a definição inicial, o modelo pode ser ajustado através da análise de resíduos e da comparação de critérios de informação como o AIC (*Akaike Information Criterion*) e o BIC (*Bayesian Information Criterion*).

A análise dos gráficos de autocorrelação e autocorrelação parcial espaço-temporais é uma etapa fundamental para a construção de modelos STARIMA adequados. A utilização de ferramentas computacionais para esse propósito facilita e agiliza o processo de definição das ordens dos componentes, aumentando a robustez dos modelos gerados [36].

### 5.3 Rede Rodoviária Típica Via Simulação Matlab Modelo STARIMA

Tendo em vista a utilização de modelos paramétricos para ajustar séries espaço-temporais, deve-se incorporar ao modelo ARIMA (modelo paramétrico série temporal) a dimensão espacial. Iniciamos revisitando o Modelo STARIMA descrito na Equação (3.11), reproduzida a seguir

$$\mathbf{z}(t) = \sum_{k=1}^p \sum_{\ell=0}^{\lambda_k} \phi_{k\ell} \mathbf{W}^{(\ell)} [\mathbf{z}(t-k)] - \sum_{k=1}^q \sum_{\ell=0}^{m_k} \theta_{k\ell} \mathbf{W}^{(\ell)} [\boldsymbol{\epsilon}(t-k)] + \boldsymbol{\epsilon}(t). \quad (5.1)$$

Seja o exemplo da "Rede rodoviária típica" com  $N = 6$  sites da Figura 22, reproduzida a seguir por conveniência. Desejamos modelar o tráfego nos sites com maiores fluxos (à jusante,  $z_5$  ou  $z_6$ ) conhecendo-se as estatísticas de geração de tráfego nos sites menos densos, geradores de tráfego (à montante).

A Figura 23.b), .c) .d) indica graficamente a evolução espaço-temporal do tráfego da malha viária de 13 sites da Figura 11, reproduzida na Figura 23.a) por conveniência.

Na Figura 24.a), é reproduzido e analisado os nós S1 a S8 da estrutura em árvore da Figura 22, típica da rede rodoviária para o fluxo de tráfego, onde são representados os 8 sites iniciais, denominados pontos de observações. Considerando uma linha horizontal imaginária entre os sites de mesmo lag espacial S6 e S7, podemos definir como Lag 1, enquanto S8 representa Lag 2. Na Figura 24.b), temos cinco séries temporais geradas a partir de uma distribuição Poisson para cinco sites (S1 a S5) e duas séries adicionais,  $z^{(1)}$  e  $z^{(2)}$ .

Cada gráfico/plot representa uma série temporal com 3000 amostras geradas pela distribuição de Poisson, onde a taxa de ocorrência ( $\lambda$ ) varia entre os sites. A variação de intensidade ou densidade ( $\lambda$ ) entre os sites sugere diferentes taxas médias de ocorrência de eventos em cada local. A legenda indica valores específicos de  $\lambda \in \{1, 1/2, 1/3, 1/5\}$ .

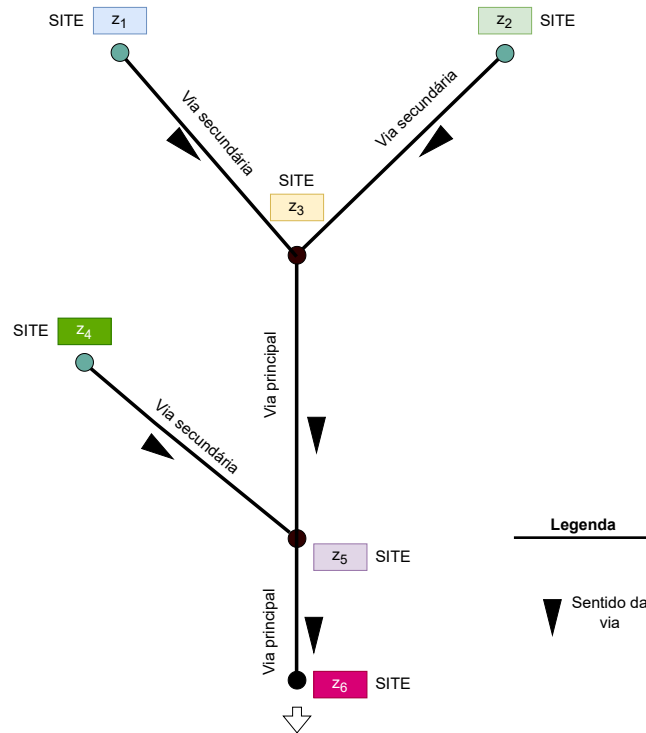
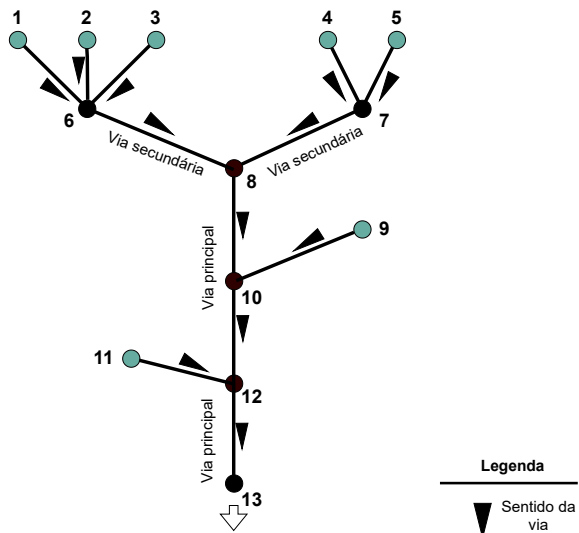


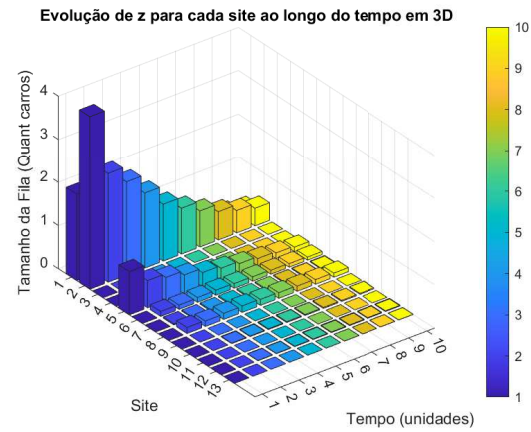
Figura 22 – Estrutura em árvore típica da rede rodoviária para fluxo de tráfego. Os pontos representam locais de medição e setas direção do fluxo.  $N = 6$  sites. Sistema referente à Equação (3.11).

Analisando a Figura 25, podemos ver os seguintes resultados em relação a modelagem preditiva de séries temporais utilizando o modelo ARIMA ( $p, D, q$ ): A Figura 25.a) ilustra a previsão da série temporal após ajuste dos coeficientes do modelo ARIMA, em termos de intervalo de confiança (95%) e a respectiva média dos valores preditos (curva em preto). É visível que as previsões das amostras no intervalo de confiança de 95% (em vermelho) seguem razoavelmente bem a dinâmica da série temporal original (em cinza), com um intervalo de confiança de 95% destacado. A previsão é feita para 1.000 amostras após o ajuste do modelo com 3.000 amostras. Este comportamento indica a capacidade do modelo em capturar tendências principais da série temporal. Já na Figura 25.b), Análise do NMSE para diferentes configurações de  $p$  e  $q$ . Podemos observar que mostra o impacto das combinações dos parâmetros  $p$  e  $q$  no desempenho do modelo, avaliado pelo NMSE. Observa-se que a configuração  $p = 2$  e  $q = 2$  apresenta o menor valor de NMSE, indicando que o modelo ARIMA(2,1,2) é o mais adequado para essa série temporal específica. Finalmente, na Figura 25.c) são exibidos os resultados de previsão do modelo ARIMA para diferentes intervalos de tempo. O gráfico mostra o desempenho do modelo em prever o comportamento da série temporal a partir de 1.000 amostras no intervalo de [2000,3000]. Os dados reais são apresentados como "data1" (em azul), enquanto as previsões ajustadas para combinações de  $p$ ,  $d$  e  $q$  diferentes são representadas em laranja. As previsões mostram boa aproximação ao comportamento dos dados reais, corroborando a eficácia do modelo ajustado.

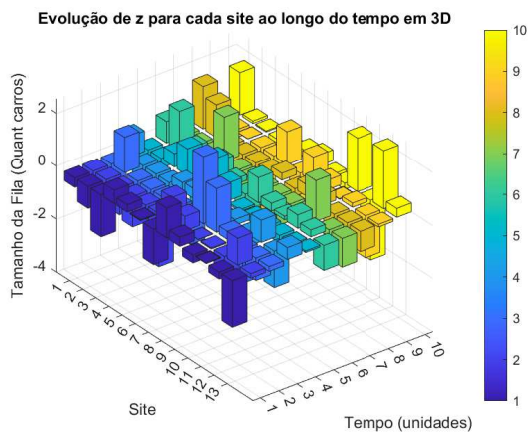
A Tabela 4, resume a comparação entre os coeficientes utilizados no modelo original e os coeficientes ajustados pelo modelo ARIMA(2,1,2). Os coeficientes ajustados ( $\phi$  e  $\theta$ ) são ligeiramente diferentes dos valores teóricos, mas a proximidade sugere que o ajuste foi bem-sucedido. O NMSE para o modelo ajustado foi 0.0525, indicando uma previsão satisfatória no contexto da série temporal analisada.



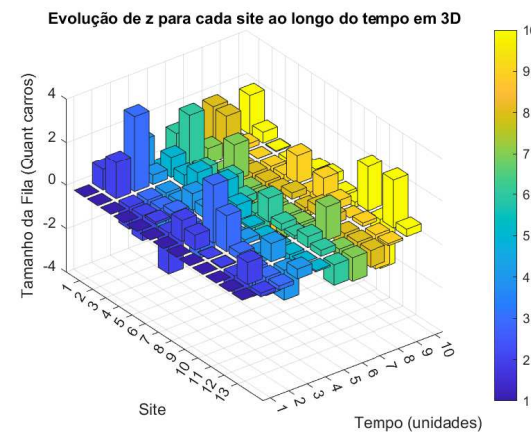
a) Esquemático: fluxo de tráfego rodoviário



b) Fila veículos em cada Site S1–S8



c) Fila veículos em cada Site S1–S13

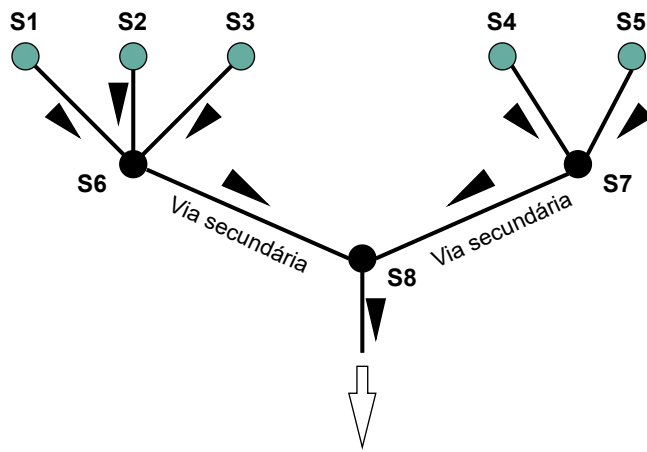


d) Fila veículos em cada Site S1–S13 em regime (últimos 10 [unid. tempo])

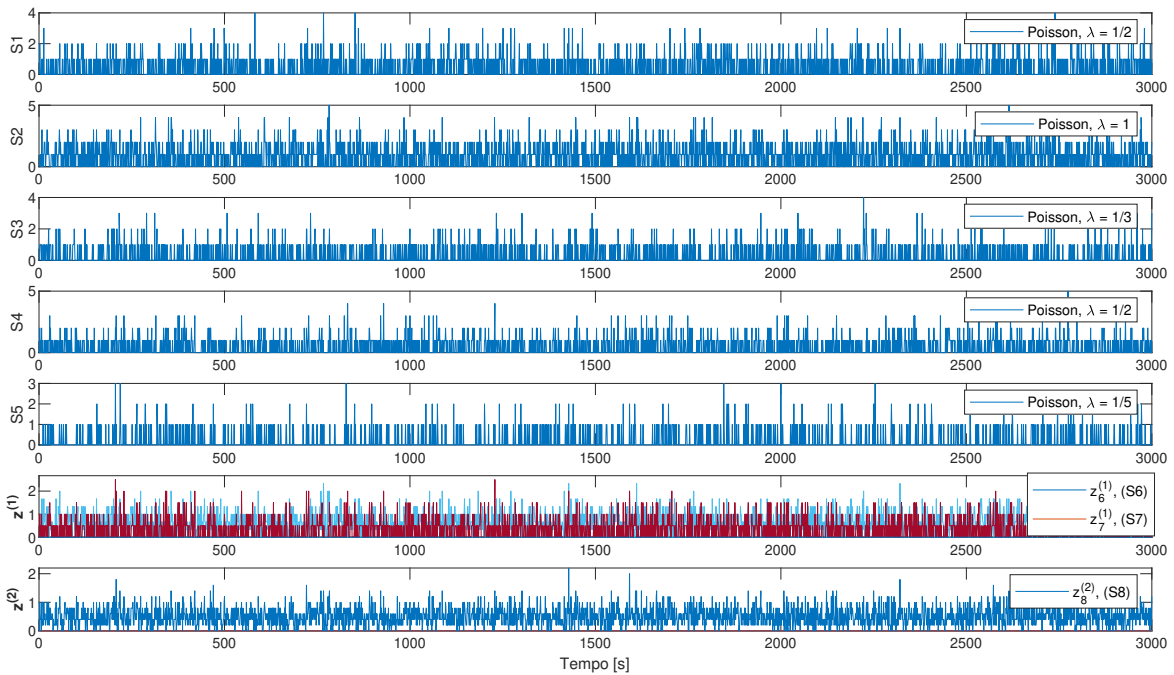
Figura 23 – Estrutura em árvore típica da rede rodoviária para fluxo de tráfego. Os pontos representam locais de medições e setas direção do fluxo ( $N = 13$  sites). Sistema referente à Equação (3.8). Para nós primários (S1-S5, S9 e S11), tráfego é gerado aleatoriamente, com distribuição Poisson de parâmetro  $\lambda$  [veículos/unid tempo]

Tabela 4 – Comparação dos coeficientes adotados no modelo com os coeficientes estimados e respectivos NMSE relativos à série temporal da Figura 25.

Modelo	$\sigma_\epsilon^2$	Modelo Ajuste	Coefficientes Ajustados	NMSE
AR{ $\rho$ }, D{ $d$ }, MA{ $q$ }				
ARIMA{[.25, -.62]; [.2, .62]} $D = 1$	0.001	STARIMA(2,1,2)	$\phi_{ar} = [-.687, -.0005];$ $\theta_{ma} = [-.327, -.672]$	0.0525
ARIMA{[.25, -.62]; [.2, .62]} $D = 0$	0.001	STARMA(2,0,2)	$\phi_{ar} = [.054, .665];$ $\theta_{ma} = [-.079, -.674]$	0.0510

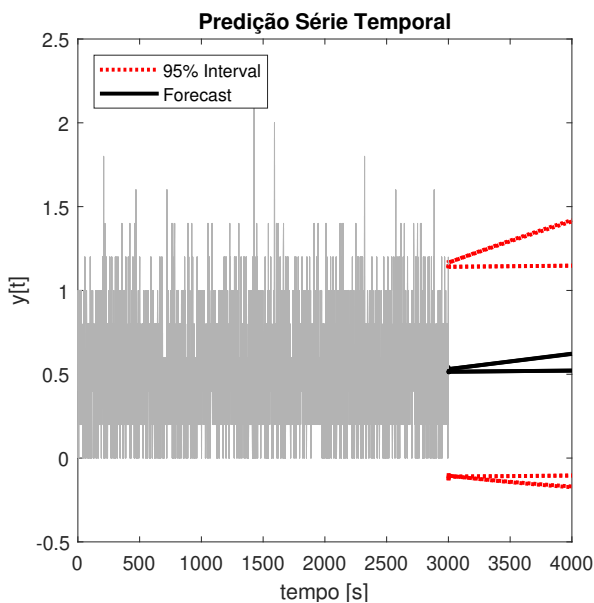


a) Estrutura em árvore típica da rede rodoviária com oito sites.

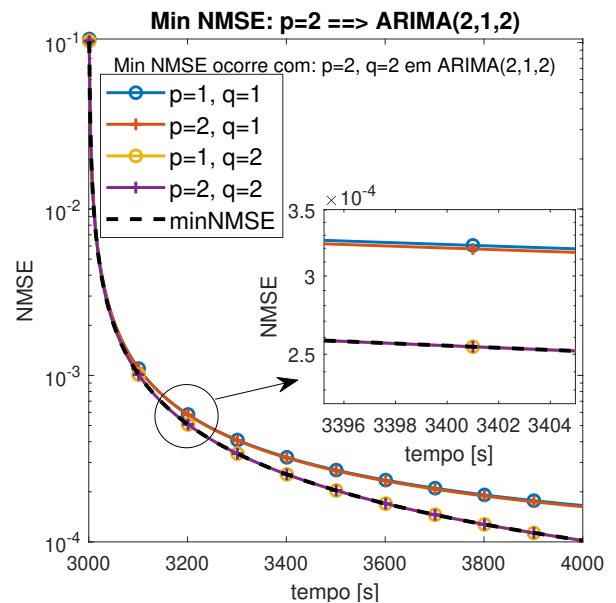


b) 3k amostras Poisson geradas para os sites S1–S5 e tráfego confluyente nos sites S6–S8 ( $z^{(1)}$ ,  $z^{(2)}$ ).

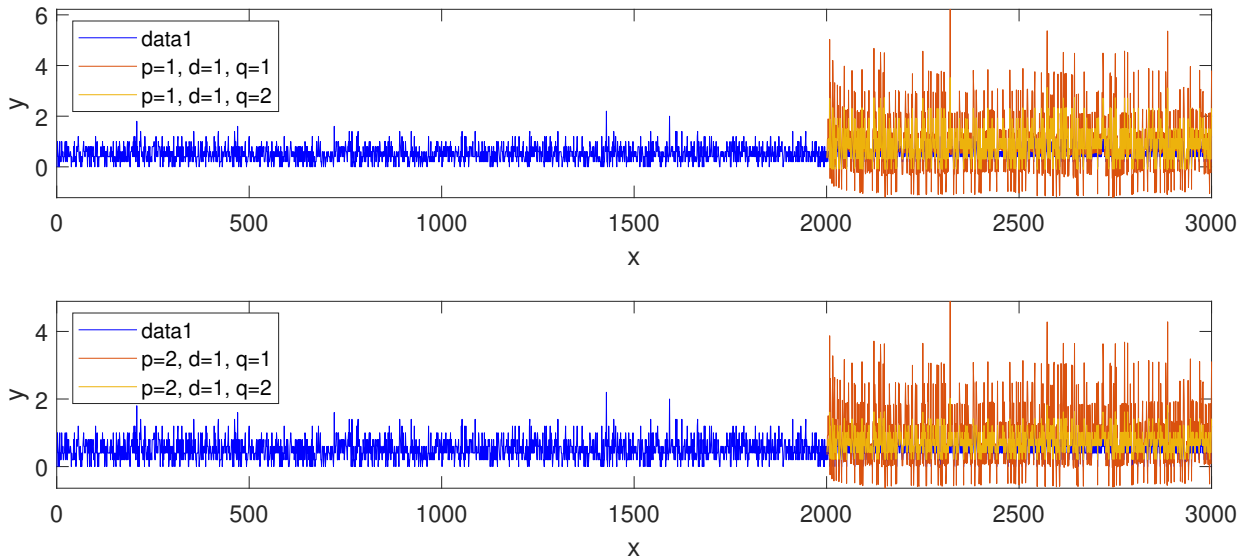
Figura 24 – Séries temporais geradas por uma distribuição Poisson para cinco sites (S1 a S5) e duas séries adicionais,  $z^{(1)}$  e  $z^{(2)}$  no MatLab R2024a.



a) 3k amostras seguida pela predição sobre 1k amostras



b) NMSE para diferentes  $p$  e  $q$ .



c) Predição a partir do modelo ARIMA(p,D,q): 1000 amostras no intervalo [2000; 3000].

Figura 25 – Obtenção dos parâmetros do modelo STARIMA(2,1,2). MatLab R2024a.

A Figura 25.a) apresenta a série temporal com intervalo de confiança de 95%, destacando a robustez da previsão para as 1.000 amostras finais. A Figura 25.b) mostra a análise de erro NMSE para diferentes configurações dos parâmetros  $p$  e  $q$ , sendo identificado o menor NMSE para o modelo ARIMA(2,1,2), posteriormente incorporado ao STARIMA(2,1,2).

As previsões ilustradas nas Figuras 25.c) demonstram que, ao utilizar  $p = 2$ ,  $d = 1$  e  $q = 2$ , o modelo STARIMA apresentou coerência com a estrutura da série temporal, inclusive nas regiões com maior variabilidade do tráfego. Os coeficientes ajustados estão apresentados na Tabela 4, com destaque para os termos autorregressivos  $\phi_{AR} = [-0,687, -0,0005]$  e os termos de média móvel  $\theta_{MA} = [-0,327, -0,672]$ . O erro mínimo obtido foi  $NMSE = 0,0525$ .

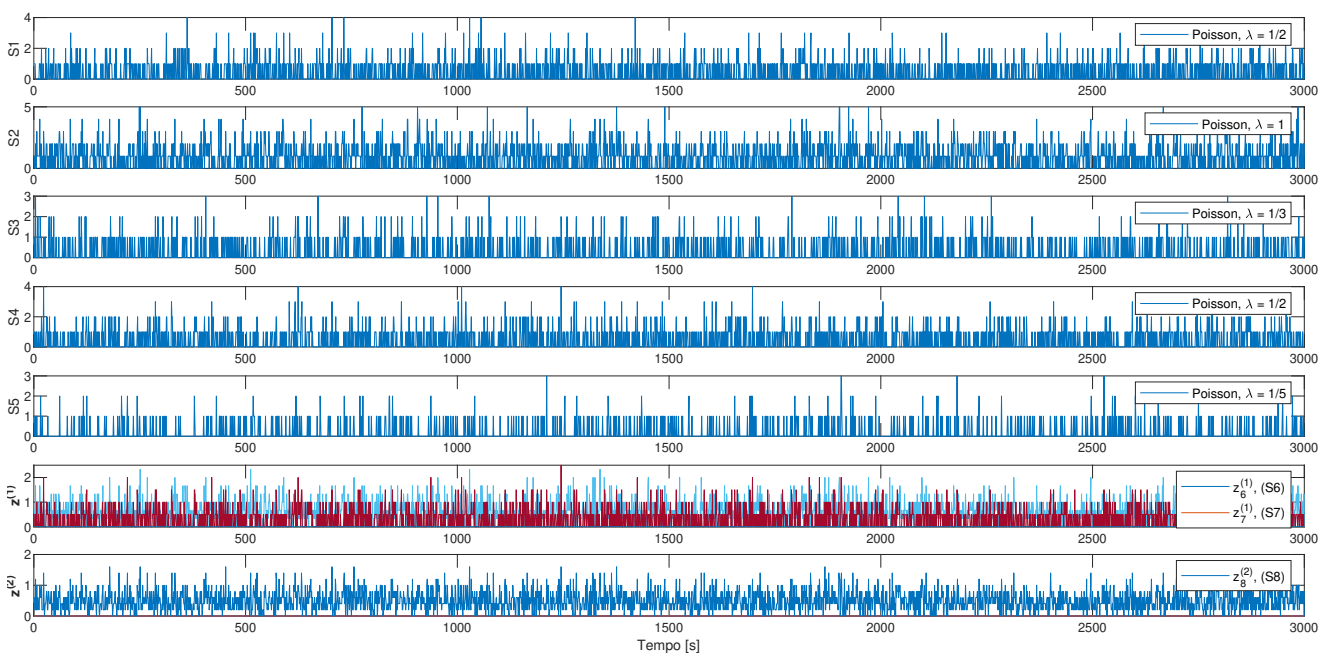
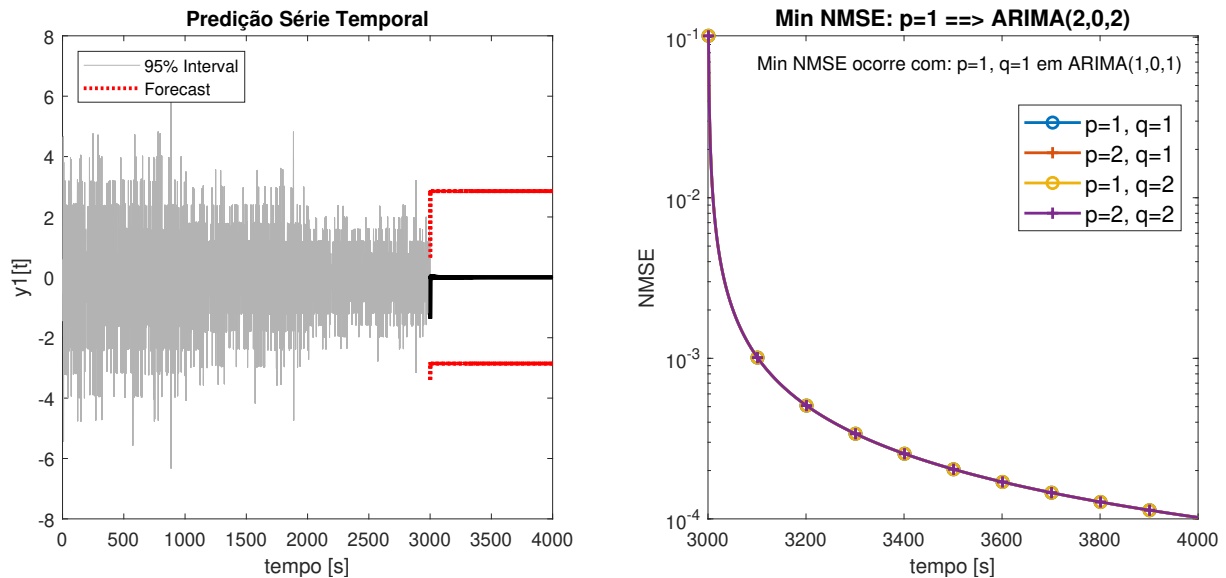


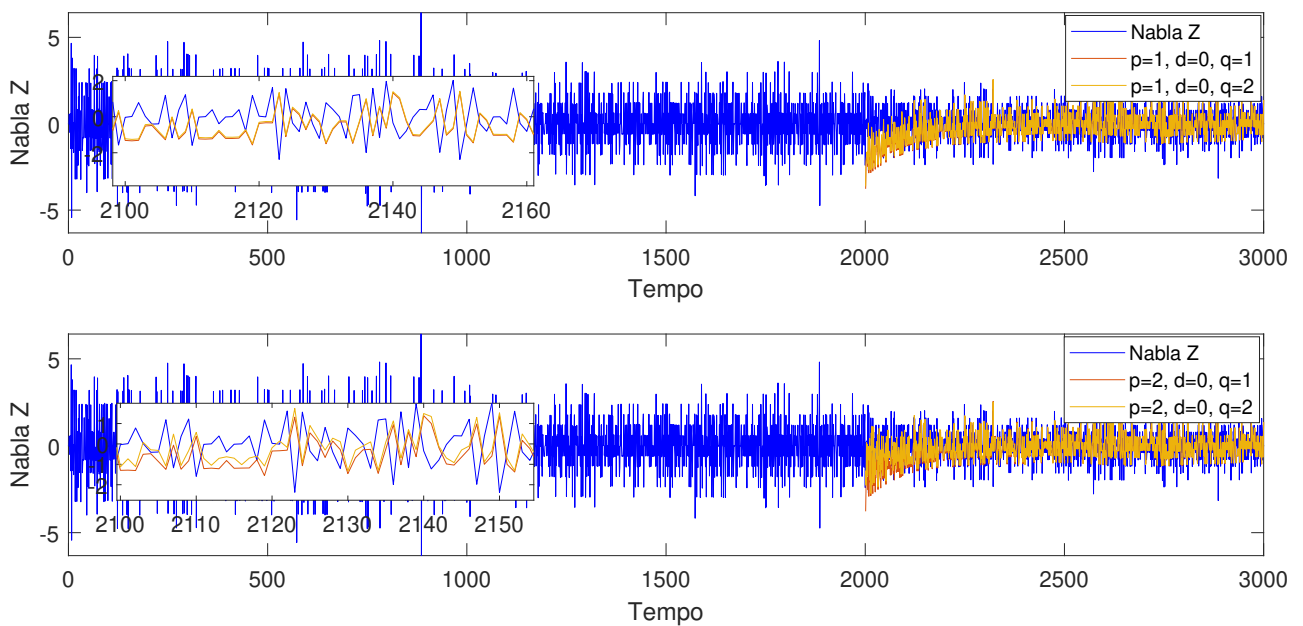
Figura 26 – Séries temporais (3k amostras) geradas por distribuições de Poisson com diferentes  $\lambda$ 's para os cinco sites de entrada (S1 a S5) e três sites de tráfego confluentes intermediários S6 e S7 ( $\mathbf{z}^{(1)}$ ) e S8 ( $\mathbf{z}^{(2)}$ ),  $\mathbf{z}^{(1)}$  e  $\mathbf{z}^{(2)}$ . MatLab R2024a.

A Figura 26 apresenta as séries temporais simuladas com base em distribuições de Poisson, correspondentes aos fluxos de tráfego em cinco pontos de entrada (S1 a S5) e três pontos intermediários (S6 a S8). Cada série contém 3.000 amostras, permitindo uma visualização clara da dinâmica de geração e propagação do tráfego. As taxas médias de ocorrência variam entre os pontos, refletindo diferentes intensidades de fluxo. Observa-se, por exemplo, que S5 possui maior densidade de eventos, enquanto S3 apresenta um comportamento mais esparsos.

Os pontos intermediários S6, S7 e S8 são representados por séries compostas ( $z^{(1)}$ ,  $z^{(2)}$ ,  $z$ ), obtidas da combinação dos fluxos de entrada, o que permite investigar a propagação do tráfego na rede. A presença de múltiplos componentes na série associada a S7 indica que esse ponto concentra sobreposições relevantes, representando um caso interessante para a avaliação de correlações espaço-temporais. Esta configuração serve como base para a aplicação do modelo STARIMA nas próximas análises.



a) 3k amostras seguida pela predição sobre 1k amostras      b) NMSE para diferentes  $p$  e  $q$ .



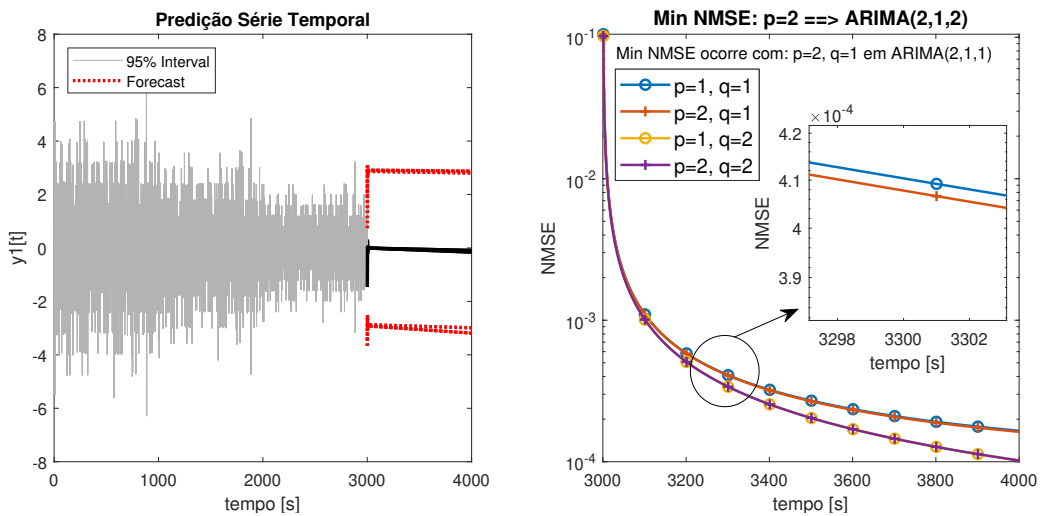
c) Predição a partir do modelo ARIMA( $p,D,q$ ): 1000 amostras no intervalo [2000; 3000].

Figura 27 – Obtenção dos parâmetros do modelo STARMA(2,0,2). MatLab R2024a.

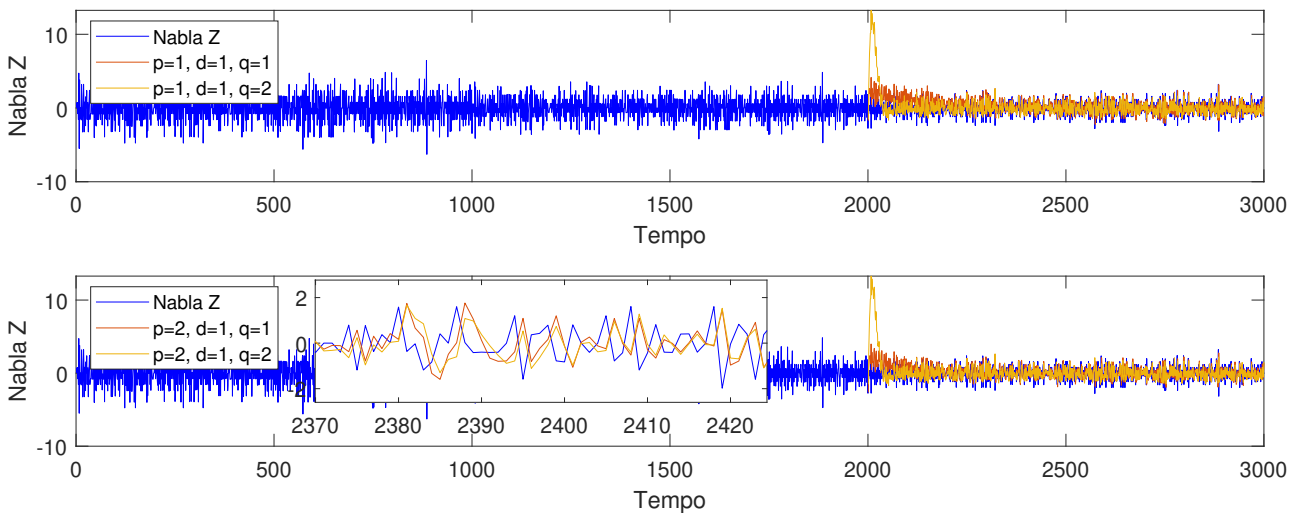
Na Figura 27.c), temos a Série temporal gerada a partir do Modelo STARIMA, com valores de  $p=2$ ,  $d=$

0 e  $q=2$ .

A Figura 27 explora uma segunda configuração com  $d = 0$ , ou seja, sem diferenciação temporal. A previsão, nesse caso, apresentou maior oscilação e menor aderência à série observada. A Tabela 4 confirma essa tendência, com um NMSE = 1,0678, ligeiramente superior ao modelo com diferenciação, mas com maior instabilidade visual nas previsões (Figura 27c).



a) 3k amostras seguida pela predição sobre 1k amostras      b) NMSE para diferentes  $p$  e  $q$ .



c) Predição a partir do modelo STARIMA( $p,D,q$ ): 1000 amostras no intervalo [2000; 3000].

Figura 28 – Obtenção dos parâmetros do modelo STARIMA(2,1,2). MatLab R2024a.

A Figura 28 apresenta a predição da série temporal utilizando o modelo STARIMA(2,1,2), cuja configuração adota diferenciação de primeira ordem ( $d = 1$ ). Observa-se que o modelo apresenta boa aderência à série observada, mesmo nas regiões com maior variação de tráfego, conforme evidenciado na predição das 1.000 amostras finais. A curva de previsão acompanha de forma coerente a estrutura da série original, com incerteza limitada dentro do intervalo de confiança de 95%, indicado pela faixa pontilhada em vermelho. O menor erro obtido para essa configuração foi NMSE = 0,0525, conforme apresentado na Tabela 4.

A fim de validar a consistência do modelo ajustado, a Tabela 5 apresenta uma comparação entre os coeficientes utilizados na geração da série temporal original e aqueles estimados pelo modelo STARIMA nas configurações com e sem diferenciação. Observa-se que, embora a configuração com  $d = 0$  tenha apresentado um valor de NMSE ligeiramente inferior, os coeficientes ajustados apresentam maior variação em relação aos valores originais. Já o modelo com  $d = 1$  mostra maior proximidade nos parâmetros

estimados, reforçando a estabilidade da previsão.

Tabela 5 – Comparação dos coeficientes adotados no modelo original com os coeficientes estimados e respectivos NMSE relativos à série temporal da Figura 28.

Modelo	$\sigma_\varepsilon^2$	Modelo Ajuste	Coefficientes Ajustados	NMSE
ARIMA{[.25, -.62]; [.2, .62]} e D = 1	.001	STARIMA(2,1,2)	$\phi_{ar} = [-.022, .013];$ $\theta_{ma} = [-1.845, -.857]$	1.0152
ARIMA{[.25, -.62]; [.2, .62]} e D = 0	.001	STARMA(2,0,2)	$\phi_{ar} = [-.985, .004];$ $\theta_{ma} = [-.006, -.971]$	1.0678

Na sequência, apresenta-se a saída do MatLab referente ao ajuste do modelo STARIMA(2,1,2), destacando os valores obtidos para os coeficientes autorregressivos ( $\phi_{AR}$ ) e de média móvel ( $\theta_{MA}$ ), bem como os parâmetros estatísticos associados à estimação.

### Saída MatLab – Predição dos parâmetros $p$ e $q$ do modelo STARIMA(2,1,2):

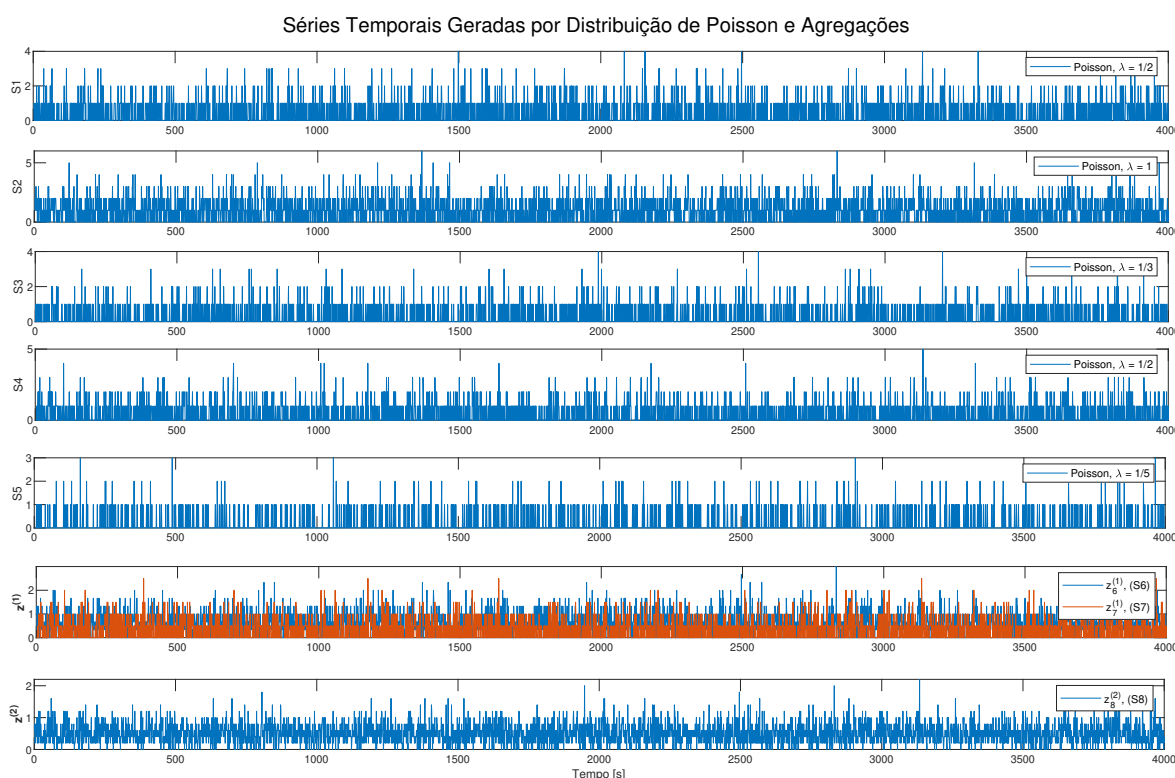


Figura 29 – Série Temporal Geradas por Distribuição de Poisson

A Figura 29 apresenta as séries temporais sintéticas geradas por distribuições de Poisson, representando o fluxo de tráfego em diferentes pontos da malha viária simulada. Os gráficos exibem dados de entrada e agregações em múltiplos sites (S1 a S8), com diferentes taxas de ocorrência  $\lambda$ , conforme especificado em seções anteriores. As amostras foram geradas ao longo de um intervalo de 4k segundos, com o objetivo de reproduzir condições realistas de variação de tráfego ao longo do tempo.

As séries referentes aos sites S1 a S5 representam pontos de entrada independentes, enquanto os sites S6 a S8 correspondem a nós de agregação, resultantes da combinação de múltiplos fluxos. A sobreposição de diferentes componentes nos gráficos inferiores reflete a complexidade da dinâmica de tráfego nesses pontos de confluência, sendo especialmente relevante para a avaliação dos modelos de previsão.

A Figura 30, apresenta uma visualização comparativa entre os dados históricos, os dados observados e as previsões do modelo ARIMA, a Figura 30 divide a série em duas partes: uma fase de treinamento, onde os dados originais são exibidos, e uma fase de teste, onde os dados reais (observados) e as previsões são comparados. O intervalo de tempo vai de 0 a 4k segundos.

A partir do início até cerca de 3k segundos, os dados de treinamento são mostrados em uma linha cinza. Esses dados representam a parte histórica com a qual o modelo foi ajustado.

A Série Temporal apresenta certa variabilidade e ruído, o que é típico de séries com comportamentos dinâmicos.

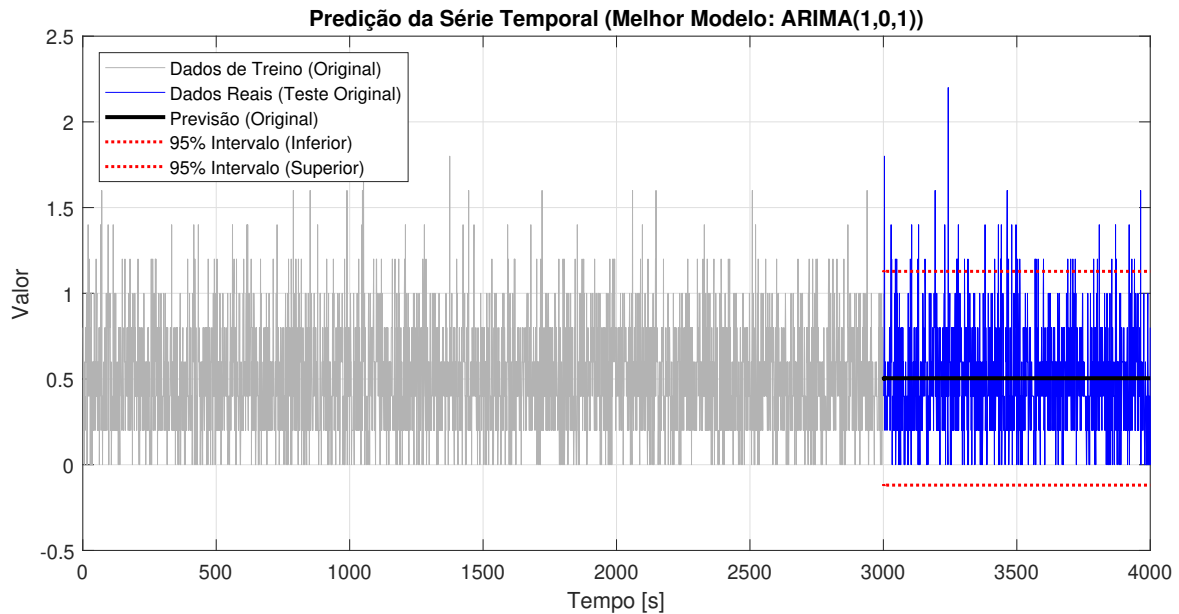


Figura 30 – Predição da Série Temporal

Após os 3k segundos, a figura exibe os valores reais observados em uma linha azul. Esses dados foram reservados para avaliar a capacidade preditiva do modelo. Observa-se que a série real apresenta flutuações mais acentuadas, sugerindo uma variabilidade que pode incluir picos e depressão não tão evidentes na fase de treinamento.

As previsões geradas pelo modelo ARIMA(1, 0, 1) são representados por uma linha preta que se estende sobre o intervalo de teste. Essa linha preditiva é mais suave quando comparada à série real, sugerindo que o modelo, ao tentar capturar o padrão geral, suaviza as variações bruscas observadas nos dados reais.

Para indicar a incerteza das previsões, a figura inclui duas linhas vermelhas tracejadas que delimitam o intervalo de confiança a 95%. A linha superior se aproxima de valores em torno de 1,5, enquanto a inferior se situa próximo a 0. Essa faixa estreita sugere que o modelo apresenta alta confiança em suas estimativas, embora possa não estar capturando toda a variabilidade dos dados observados.

A Figura 31 apresenta a análise comparativa do desempenho de diferentes configurações do modelo ARIMA, considerando as ordens ( $p, q$ ) e avaliando os resultados por meio das métricas NMSE e EC. No gráfico da esquerda, observa-se que o menor valor de NMSE foi obtido com a configuração ARIMA(1,0,1), indicando que esta combinação proporciona a melhor performance preditiva dentre os modelos testados. O valor mínimo obtido foi NMSE = 0,2925.

Já no gráfico da direita, que exibe os valores do EC, nota-se que todos os valores permanecem negativos, o que revela desempenho inferior à simples média da série observada. O modelo com  $p = 1$  ainda apresenta o valor menos negativo (EC  $\approx -0,0001$ ), o que indica, mesmo que de forma marginal, uma vantagem relativa.

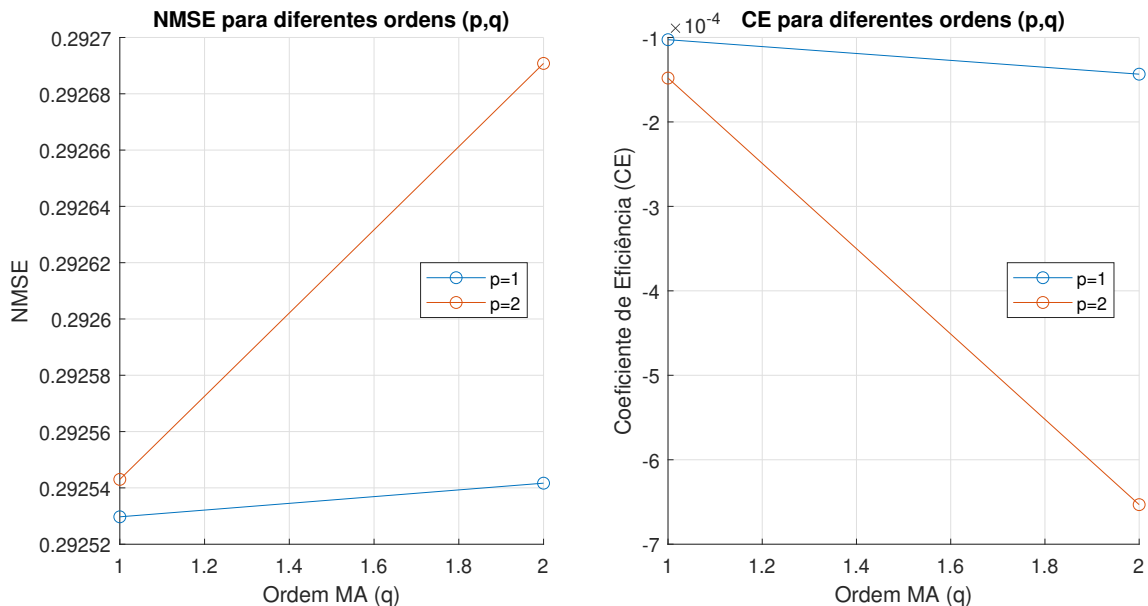


Figura 31 – NMSE e EC para diferentes ordens

Apesar de os valores absolutos das métricas sugerirem limitações no ajuste dos modelos testados, os resultados obtidos com ARIMA(1,0,1) são os mais consistentes entre as opções analisadas. A combinação de menor NMSE e menor magnitude de erro no EC sugere que, para os dados em questão, essa estrutura de ordem oferece a melhor aproximação entre modelo e série observada.

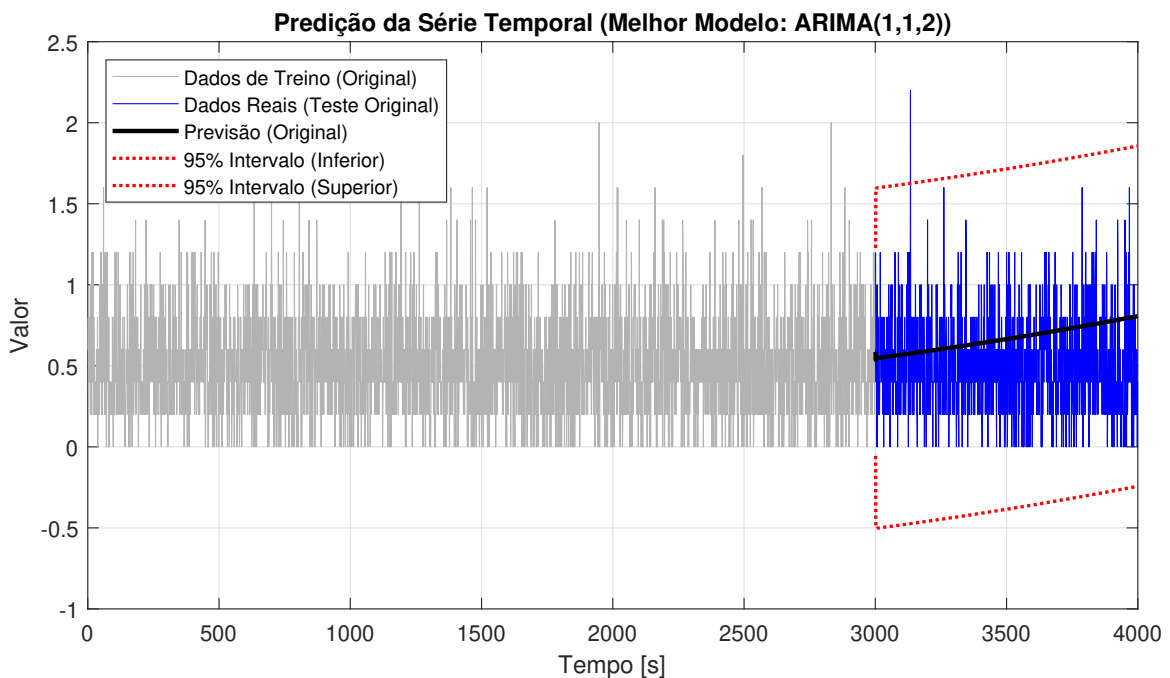


Figura 32 – Predição da Série Temporal

A Figura 33 apresenta a comparação entre diferentes combinações de ordens ( $p, q$ ) no modelo ARIMA, utilizando duas métricas o NMSE e o EC. No gráfico à esquerda, observa-se que o menor valor de NMSE ocorre para a configuração com  $p = 2$  e  $q = 2$ , indicando melhor desempenho preditivo entre as combinações avaliadas. Ainda que a variação no NMSE entre as configurações seja sutil, essa diferença reforça a sensibilidade da métrica na avaliação da qualidade da previsão.

No gráfico à direita, o comportamento do EC segue a mesma tendência, assumindo seu valor mais próximo de zero na mesma configuração ( $p = 2, q = 2$ ). Vale lembrar que, por definição, quanto mais

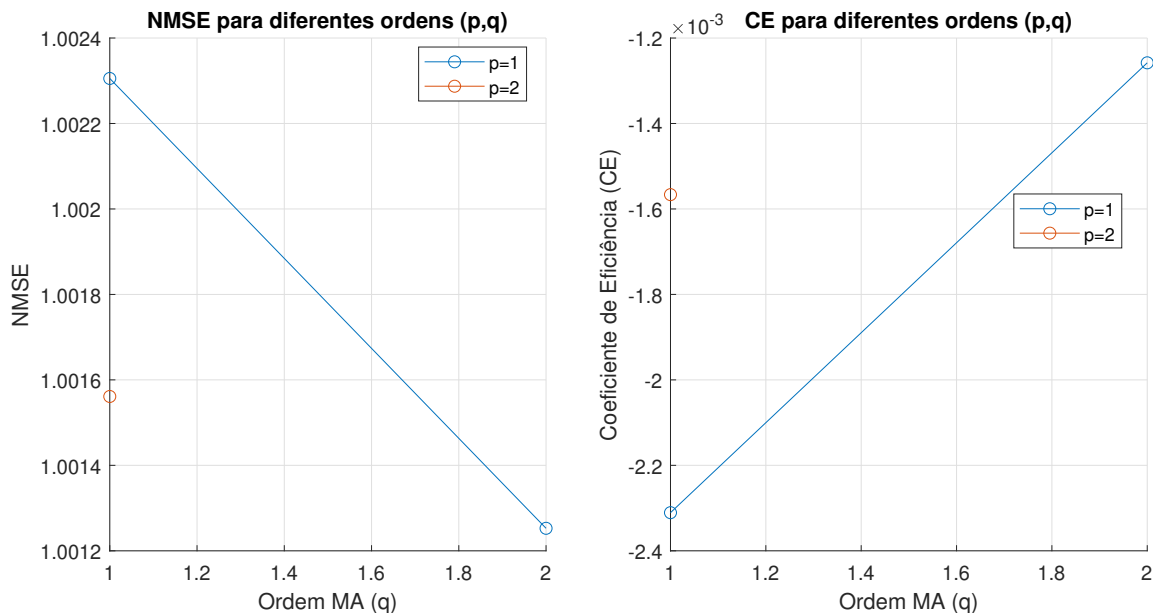


Figura 33 – NMSE e EC para diferentes ordens

próximo de 1 for o valor de EC, melhor o desempenho do modelo em relação à média dos dados observados. No entanto, os valores de EC observados são bastante baixos, próximos de zero e até negativos, o que sugere que o modelo ainda não apresenta uma performance significativamente superior à média da série temporal.

A análise conjunta das duas métricas indica que, embora o modelo ARIMA(2,2) apresente o melhor desempenho relativo dentro das opções testadas, os resultados ainda podem ser considerados modestos em termos absolutos.

#### 5.4 Emulando o Tráfego Urbano na Av. 23 de Maio - São Paulo, BR

No contexto da modelagem paramétrica com o modelo STARIMA, a previsão de tráfego urbano na Avenida 23 de Maio, em São Paulo, representa um estudo de caso relevante para demonstrar sua aplicabilidade. Utilizando dados reais de fluxo veicular obtidos a partir de radares da Companhia de Engenharia de Tráfego (CET) da cidade de São Paulo, como proposto em estudos como o de Zema et al. [4] da Escola Politécnica da USP<sup>3</sup>, o modelo STARIMA pode ser implementado para capturar dependências espaço-temporais do tráfego. Esses dados, que incluem medições de velocidade, densidade e fluxo em intervalos de 15 minutos, permitem a construção de uma matriz de pesos espaciais baseada na conectividade entre pontos de monitoramento ao longo da avenida. A aplicação do STARIMA, conforme discutido em [37]<sup>4</sup>, envolve a estimação de parâmetros autorregressivos e de média móvel (e.g., STARIMA(2,1,2)) utilizando técnicas como autocorrelação (ACF/PACF) e testes de estacionariedade (ADF), garantindo previsões de curto prazo robustas. Essa abordagem possibilita a antecipação de congestionamentos e a otimização de estratégias de controle de tráfego, como ajustes dinâmicos de semáforos, contribuindo para a eficiência da mobilidade urbana em corredores viários de alta demanda como a Avenida 23 de Maio.

<sup>3</sup> TCC de alunos da Eng. Civil, EPUSP relacionado ao Estudo de Caso: Tráfego na 23 Maio, SP, disponível em: Aplicação de ARIMA em previsão de Tráfego de Curto Prazo, EPUSP 2020

<sup>4</sup> Vídeo disponível em [Aplicação de STARIMA em Previsão de Tráfego de Curto Prazo: Um Estudo de Caso na Cidade de São Paulo \(desenvolvimento Tf2\)](#)

### 5.4.1 Métodos

A seguir, apresentam-se os métodos utilizados para medir o volume e a velocidade, assim como os procedimentos adotados para consulta dos dados pesquisados [3].

#### 5.4.1.1 Método - Volume

O objetivo da pesquisa e contagem volumétrica classificada é determinar a quantidade, a composição e o sentido do fluxo de veículos em uma seção do sistema por unidade de tempo.

A quantificação dos veículos é feita por meio de observação direta por pesquisadores equipados com contadores manuais. Os pesquisadores ficam posicionados na seção de contagem, escolhendo locais que proporcionem uma boa visibilidade do fluxo analisado. O número de observadores em cada ponto é definido conforme a quantidade de faixas, o volume e o tipo de tráfego local.

Os veículos são divididos nas seguintes categorias: automóveis, motocicletas, ônibus urbanos, ônibus fretados e caminhões. Em vias consideradas estratégicas, como a Avenida do Estado e a Avenida Jacupêssego, a classificação dos caminhões é feita com base no número de eixos: 2, 3 ou 4 eixos(ou mais)

A contagem é cumulativa durante toda a coleta de dados, ou seja, os contadores não são reiniciados ao longo do levantamento. Os valores acumulados são registrados a cada 15 minutos.

A pesquisa ocorre em dois dias distintos, com medições feitas no período da manhã, das 7h às 10h, e no período da tarde, das 17h às 20h. Os volumes horários com intervalos de 15 minutos(por exemplo: das 7h às 8h, das 7h15 às 8h15, até as 9h às 10h) .

As medições são realizadas em três ou quatro pontos ao longo da rota, sendo escolhido um ponto como principal, onde a contagem é realizada nos dois dias do levantamento .

O ponto principal é assinalado no relatório por um asterisco (\*) e, para ele, são apresentados os valores médios obtidos nos dois dias de observação. Nos demais pontos, a amostragem corresponde a um dia típico do ano. O volume simples de veículos em um ponto representa a soma total dos veículos registrados, sem distinguir o tipo. Já o volume equivalente leva em conta fatores de equivalência conforme a categoria dos veículos: ônibus e caminhões, por exemplo, têm peso dois no cálculo. Neste relatório, motocicletas são consideradas equivalentes a carros de passeio .

A contagem do fluxo de tráfego é realizada separadamente para cada direção, incluindo ambos os sentidos em vias de mão dupla. Os sentidos costumam ser identificados como Bairro–Centro e Centro–Bairro, mas em rotas onde esses nomes não se aplicam, utiliza-se outra denominação, como Ibirapuera–Pinheiros e Pinheiros–Ibirapuera.

Algumas rotas são constituídas por vias de mão única, formando sistemas binários de circulação. Nestes casos, as contagens são feitas em pontos correspondentes de ambas as vias. Um exemplo desse sistema é o binário formado pela R. Cardeal Arcoverde e R. Teodoro Sampaio (Rota 23). Em casos específicos, como a Rota 15 – Rótula Central, existe apenas um sentido de circulação, neste caso, no sentido anti-horário.

Para facilitar a visualização dos dados no relatório, os resultados referentes ao sentido Bairro–Centro aparecem com fundo azul, enquanto os do sentido Centro–Bairro são destacados em laranja. Quando são utilizados outros nomes de sentido, as cores ajudam igualmente a distinguir as direções em cada rota .

O Fator Hora Pico (FHP) expressa a variação do volume de tráfego em períodos de quinze minutos dentro da hora de maior movimento equivalente, podendo variar, teoricamente, entre 0,25 e 1,00. Ele é calculado dividindo-se o volume total da hora por quatro vezes o maior volume de quinze minutos observados naquela hora. O FHP chega a 1 quando o volume é totalmente uniforme nesses intervalos durante a hora de maior fluxo.

Já o FHP de 3 horas corresponde ao volume equivalente registrado em três horas dividido por três vezes o equivalente da hora de pico. Este indicador, que pode variar entre 0,33 e 1,00, tem o objetivo de apontar como o volume de tráfego se distribui ao longo das horas analisadas na pesquisa.

#### 5.4.1.2 Método - Velocidade e Tempo de Retardamento

A pesquisa de velocidade e tempo de retardamento tem como foco medir a velocidade média e os atrasos ao longo de rotas, para avaliar a facilidade de deslocamento. As rotas são divididas em trechos, delimitados por vias transversais chamadas de pontos de controle, e geralmente apresentam características operacionais semelhantes.

O método utilizado é o de veículo-teste com dois pesquisadores e cronômetros: um cronômetro para registrar o tempo de percurso em cada trecho e outro para marcar o tempo parado devido a diversos retardamentos, como congestionamentos, semáforos e outros fatores. Esses dados são anotados em uma planilha específica para posterior análise.

A velocidade média é calculada dividindo a distância do trecho pelo tempo médio gasto. Cada rota é avaliada em ambos os sentidos durante dois dias, nos períodos da manhã e tarde, com pelo menos três viagens para cada sentido em cada período.

Desde 2009, prioriza-se a análise do tráfego geral, pois o monitoramento dos ônibus passou a ser feito principalmente por GPS instalado pela SPTrans nos veículos urbanos.

Os sentidos das rotas são identificados por cores (azul para Bairro-Centro e laranja para Centro-Bairro) e os resultados de velocidade, tempo de percurso e percentuais de retardamento são apresentados consolidados em uma página, facilitando a consulta e a interpretação dos dados [3].

#### 5.4.2 Descrição da Base Experimental e Métodos de Identificação Aplicados

Os dados usados neste estudo são do trabalho da ref. [34] com 482 amostras para cada um dos 4 radares [10531 10500 10482 10433], sentido Aeroporto/Santana, Figura 36. O estudo de previsão de tráfego em uma das importantes avenidas da cidade de São Paulo, a Av. 23 de Maio, responsável pela interligação Norte-Sul e Leste-Oeste, Figura 34 e 36. Os dados são de oito radares presentes na avenida e estão divididos em dois sentidos: Santana/Aeroporto e Aeroporto/Santana. Foram selecionadas as colunas de maior relevância para a análise.

Para a identificação do modelo, foram utilizadas as funções STACF e STPACF, e o método de resolução não-linear Levenberg-Marquardt. Dados conhecidos foram utilizados para entender melhor o funcionamento desses métodos de identificação.

A partir da aba Construção do Modelo (Subseção 4.5.1) desenvolveu-se um modelo do zero - dessa forma, foram gerados os gráficos da STPACF (Autorregressão) e a STACF (Média Móvel), para que seja definido os valores dos coeficientes temporais e espaciais necessários à estimação de parâmetros.

Uma técnica para definir as ordens do modelo é observar nos gráficos: Onde a linha vermelha ( $L_0$ ) apresentar o último valor significativo acima do ruído, é o possível *timelag* da ordem temporal de AR (gráfico STPACF) e de MA (gráfico STACF); A linha de maior índice  $n$  em  $L_n$  que apresentar valores significativos acima do ruído é a ordem espacial de AR (STPACF) e MA (STACF).

O tempo de processamento foi de 1 minuto e para os dados de *input*, o  $d$  necessário para estacionaridade foi: 1.

Os valores dos indicadores de acurácia são analisados da esquerda para direita: O valor de  $E$  adequado é entre 0 e 1, sendo que quando  $E < 0$ , o modelo está muito ruim; quando  $E = 0$ , o modelo é equivalente a simplesmente realizar a média dos dados observados; e quando  $E = 1$ , o modelo é perfeito; Os indicadores

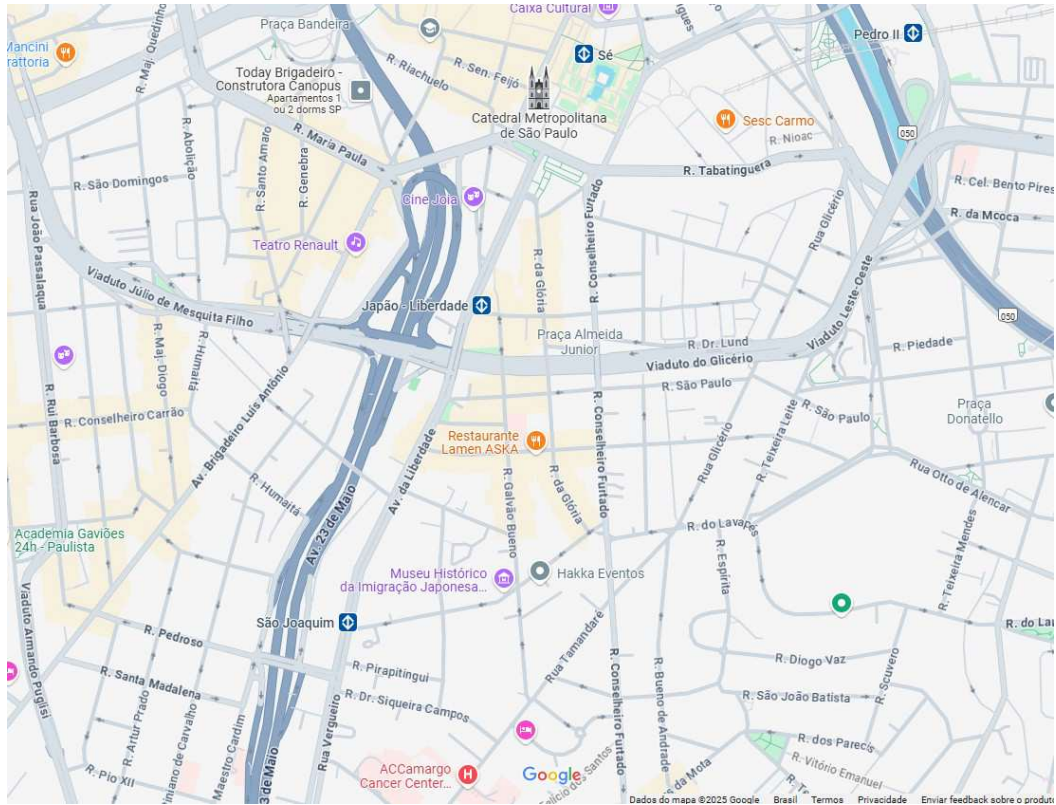


Figura 34 – Av 23 de Maio - SP (Fonte: Google Maps)



Figura 35 – Radares na Av 23 de Maio - SP (Fonte: [4] )

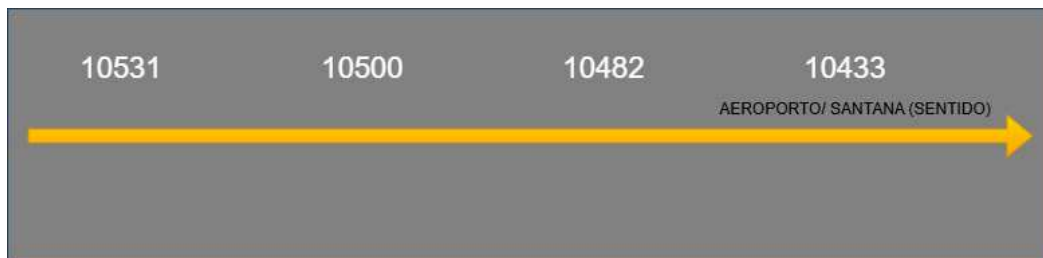


Figura 36 – Distribuição dos radares na Av 23 de Maio por sentido (AS) - (Fonte:Autor)

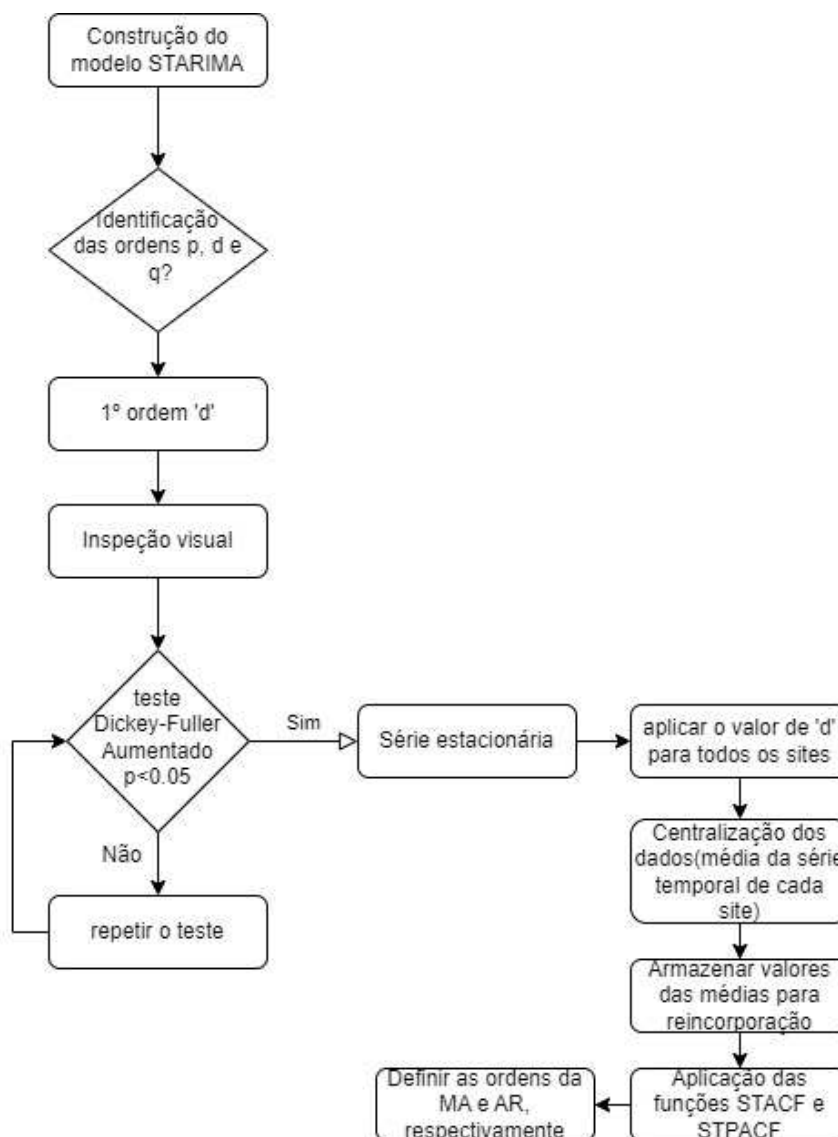


Figura 37 – Fluxograma da construção do modelo

RSS e MAE devem ser minimizados; E o indicador *Box-Pierce* serve como desempate entre os modelos, escolhendo-se de menor valor.

que o valor do  $E$  está menor que zero, logo o modelo está ruim.

#### 5.4.3 Identificação das Ordens $p$ , $d$ e $q$

Define-se a ordem para  $d$  com base no procedimento do fluxograma detalhado na Figura 37. Realizando uma inspeção visual, podemos já verificar a necessidade de diferenciações para tornar a série estacionária. Observa-se que existem muitas variações que não resultam em uma média zero.

Análise dos resultados:

O estudo visa investigar a estacionaridade de séries temporais espaciais extrapoladas de quatro localidades (CET L0 a L3) e construir um modelo de previsão por meio do modelo STARIMA.

- refazer as matrizes  $W^{(\ell)}$  conforme a estrutura do modelo da Rua 23 de Maio;
- Usar a função `curveFitter` do Matlab para encontrar os valores de  $\phi$  e  $\theta$ .
- Teste de estacionariedade (ADF);
- Transformações diferenciais ( $\nabla^1$  e  $\nabla^2$ ),
- Estimativa de parâmetros autorregressivos (AR) e de média móvel (MA),
- Avaliação comparativa entre métodos de modelagem (Curve Fitter vs ARIMA),
- Validação com métricas quantitativas (NMSE).

### Matrizes de peso $W^{(\ell)}$ :

As matrizes de peso foram modeladas de acordo com a arquitetura da via estudada.

$W^{(0)} =$

1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1

$W^{(1)} =$

0	0	0	0
1	0	0	0
0	1	0	0
0	0	1	0

$W^{(2)} =$

0	0	0	0
0	0	0	0
1	0	0	0
0	1	0	0

Nos gráficos superiores das Figura 38.a) e b), cada série temporal (L0, L1, L2 e L3) é apresentada ao longo do tempo, com marcação de *lags* em minutos ( intervalos de 3 minutos).

As séries exibem flutuações regulares, com possíveis tendências ou ciclos sazonais. Verifica-se que há indícios de não estacionaridade, pois o valor médio e a variância parecem variar com o tempo.

No gráfico da Figura 38 e as demais, a média é removida de cada série, realçando as oscilações em torno de zero. O objetivo é simplesmente destacar variações sem o viés do nível médio.

Anota-se que a centralização facilita a identificação de desvios locais e prepara a série para modelagem. Ainda assim, não resolve problemas de tendência ou não estacionaridade, que exigem diferenciação.

No gráfico, Diferença de Primeira Ordem, é aplicada a operação  $\nabla^1 Z(t) = Z(t) - Z(t - 1)$ , Equação (3.13), que remove tendências lineares e ajuda a tornar a série estacionária.

Podemos observar que a variabilidade da série  $\nabla^1$  é mais estável ao longo do tempo. E as flutuações tornam-se mais homogêneas, o que favorece análises de correlação e modelagem ARIMA/STARIMA.

A Figura 38 apresenta uma visualização consolidada das etapas iniciais do processo de modelagem com STARIMA, reunindo em um único painel os principais elementos anteriormente distribuídos entre as Figuras 22 a 25. No primeiro gráfico, observa-se a divisão dos dados em conjuntos de treino e teste, com destaque para os volumes de tráfego coletados por diferentes sensores (identificados por 10531, 10500, 10482 e 10433). A fase de treino abrange o início da madrugada até o meio da tarde do dia 14 de março de 2018, enquanto a fase de teste cobre o intervalo subsequente, entre 15h00 e 00h00 do mesmo dia.

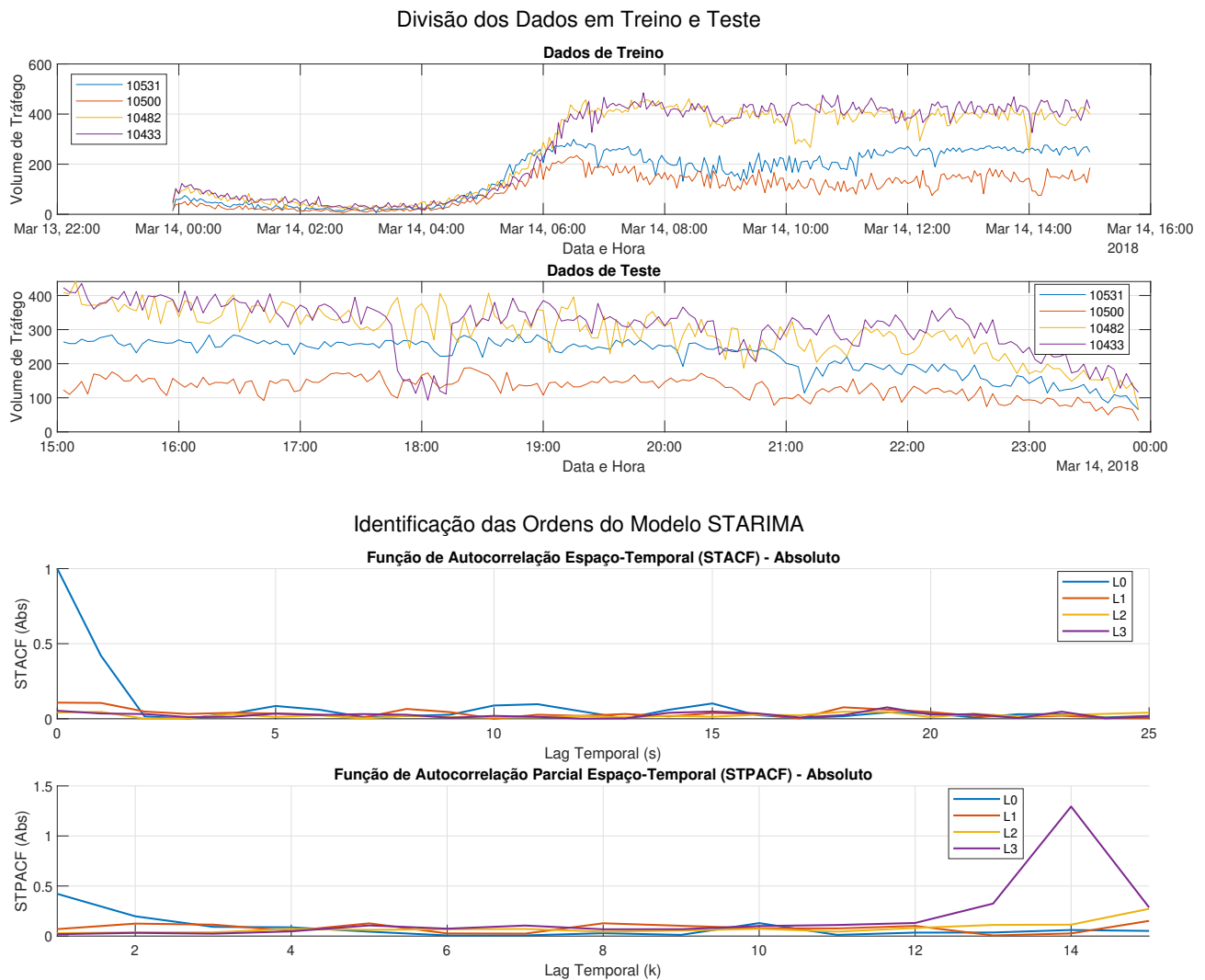


Figura 38 – (a) Série Temporal do dados do CET radar L0; (b) Autocorrelação e Autocorrelação Parcial

O segundo gráfico exibe os dados de teste, que apresentam maior variabilidade e oscilações mais irregulares entre os sensores, o que evidencia a complexidade do comportamento do tráfego urbano em horários de pico e transição, validando a capacidade dos modelos preditivos em lidar com padrões reais.

Na parte inferior da figura, são mostradas as funções STACF e STPACF espaço-temporais, utilizadas para identificar as ordens do modelo STARIMA. A curva da função STACF mostra um rápido decaimento nos primeiros defasamentos, o que sugere a predominância de dependência de curto prazo. Já a função STPACF revela picos significativos para alguns vizinhos espaciais, indicando a presença de correlações relevantes tanto no tempo quanto no espaço.



Neste momento, os dados do Ponto 1 estão sendo analisados com atenção especial aos seguintes aspectos:

- Identificação de padrões de variação no fluxo de veículos ao longo do tempo;
- Análise da fluidez do tráfego, considerando os efeitos de fatores externos, como incidentes ou condições climáticas;
- Identificação de possíveis anomalias nos dados que possam interferir na interpretação dos resultados.

A Figura 40 apresenta a comparação entre os dados expandidos com intervalo de 3 minutos e a taxa média horária original, destacando o ganho de resolução temporal obtido com a expansão dos dados. A Figura 41 ilustra a separação dos dados nos conjuntos de treino (07h–09h) e teste (09h–19h), permitindo visualizar o comportamento do fluxo veicular ao longo do dia. Por fim, a Figura 42 mostra as funções STACF e STPACF aplicadas aos dados, fundamentais para a análise da dependência temporal e espacial entre os pontos de coleta.

A metodologia adotada visa garantir que os dados sejam processados de forma consistente, permitindo a geração de *insights* confiáveis sobre o desempenho do sistema viário. Este tratamento é indispensável para validar as hipóteses do estudo e fornecer uma base sólida para futuras implementações de melhorias na gestão do tráfego urbano.

A análise do Ponto 1 será complementada com modelo de previsão STARIMA, para investigar padrões temporais e prever comportamentos futuros, o que pode ser uma contribuição significativa para estratégias de planejamento viário mais eficientes.

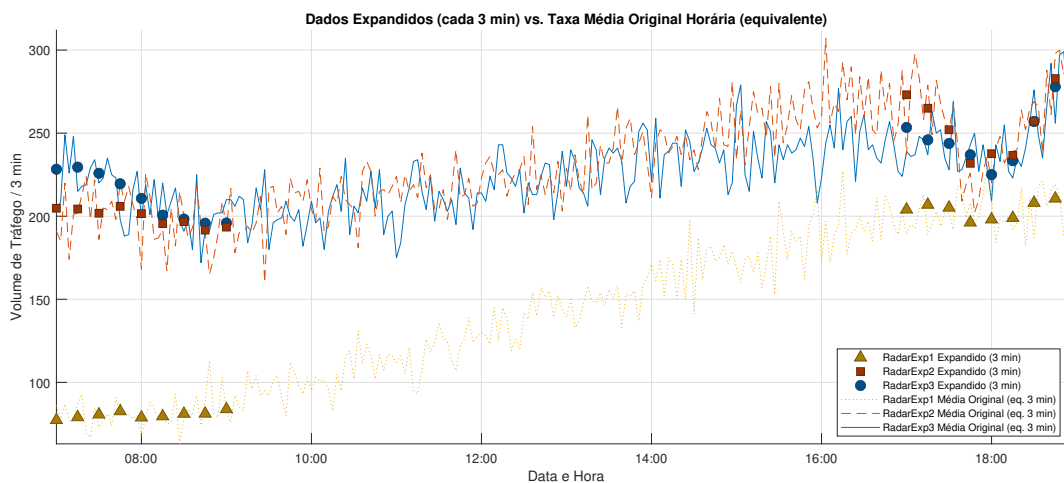


Figura 40 – Dados expandidos (cada 3 min.) vs. Taxa Média Original Horária

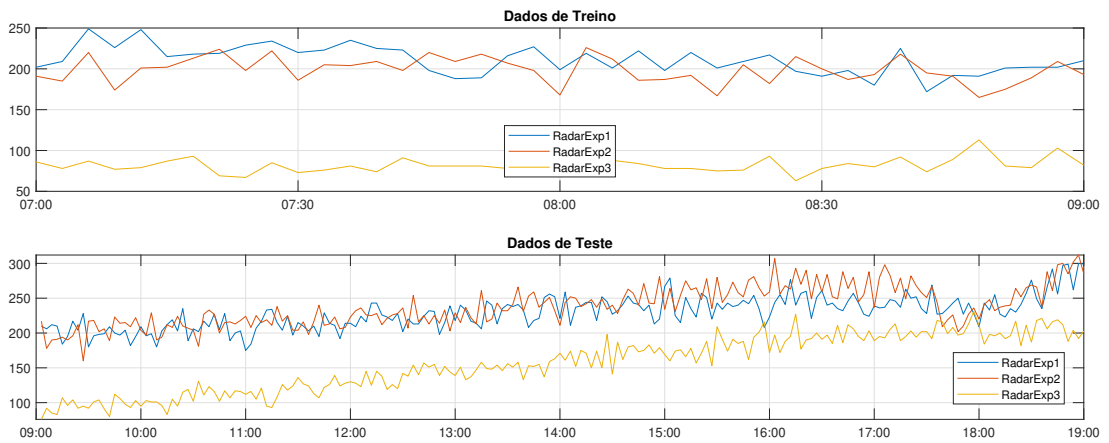


Figura 41 – Dados Treino e Teste

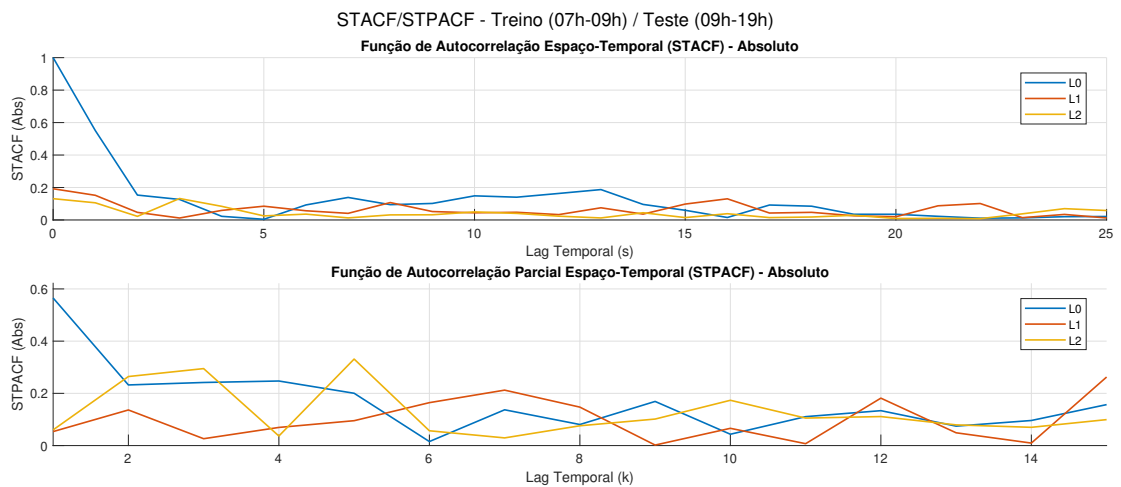


Figura 42 – Função STACF e STPACF

Na Figura 43 observa-se a aplicação do método de previsão *step-by-step*, em que os valores estimados em cada passo são utilizados como entrada para as previsões seguintes. Este tipo de abordagem tende a propagar erros cumulativos, especialmente em séries temporais longas, o que é perceptível no afastamento gradual entre a curva predita e os dados reais ao longo do tempo. Ainda assim, nota-se que o modelo STARIMA consegue captar de forma razoável a tendência geral de crescimento e as flutuações no volume de tráfego, especialmente nos radares 2 e 3, onde a estabilidade do modelo é mais evidente

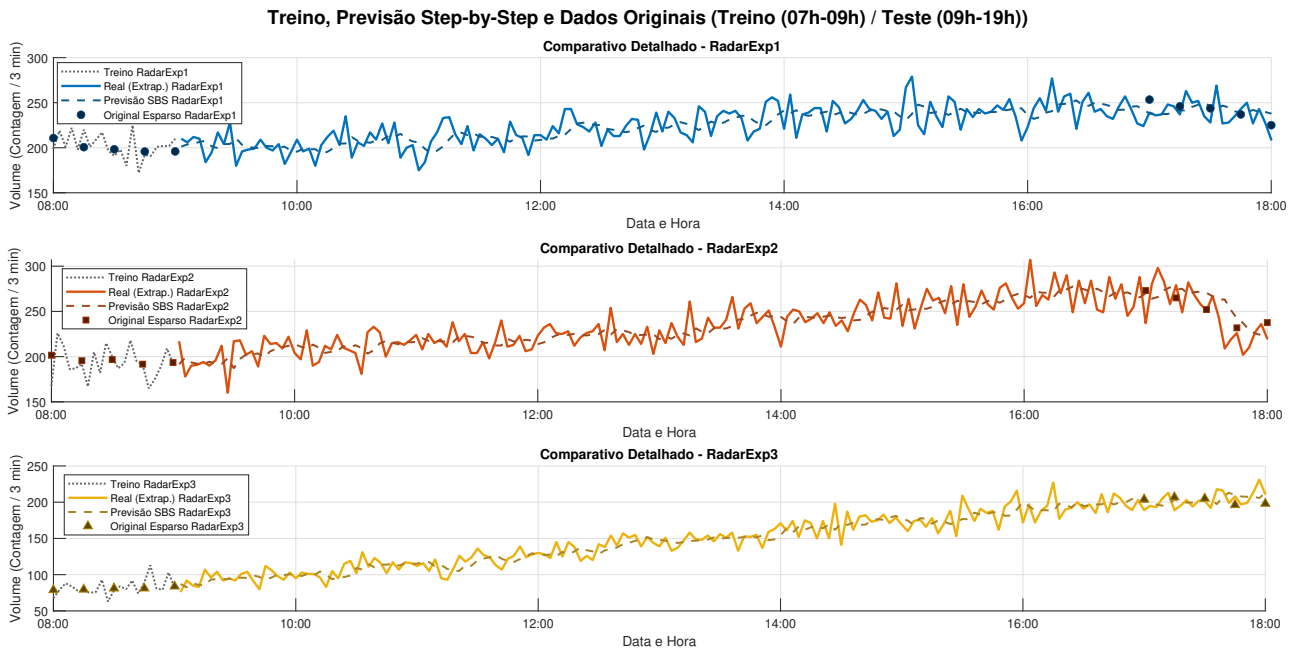


Figura 43 – Treino, Previsão *Step-by-Step* e Dados Originais

Por sua vez, a Figura 44 apresenta os resultados do mesmo modelo aplicando-se uma técnica de alinhamento com deslocamento temporal, adotando-se  $k=2$  como parâmetro. Nessa configuração, as previsões são realinhadas para comparação com os dados reais após um deslocamento de duas unidades de tempo. Tal abordagem permite uma melhor visualização da responsividade temporal do modelo e evidencia que parte dos desvios observados na Figura 43 não decorre de falhas na modelagem da tendência, mas de pequenos atrasos na resposta preditiva.

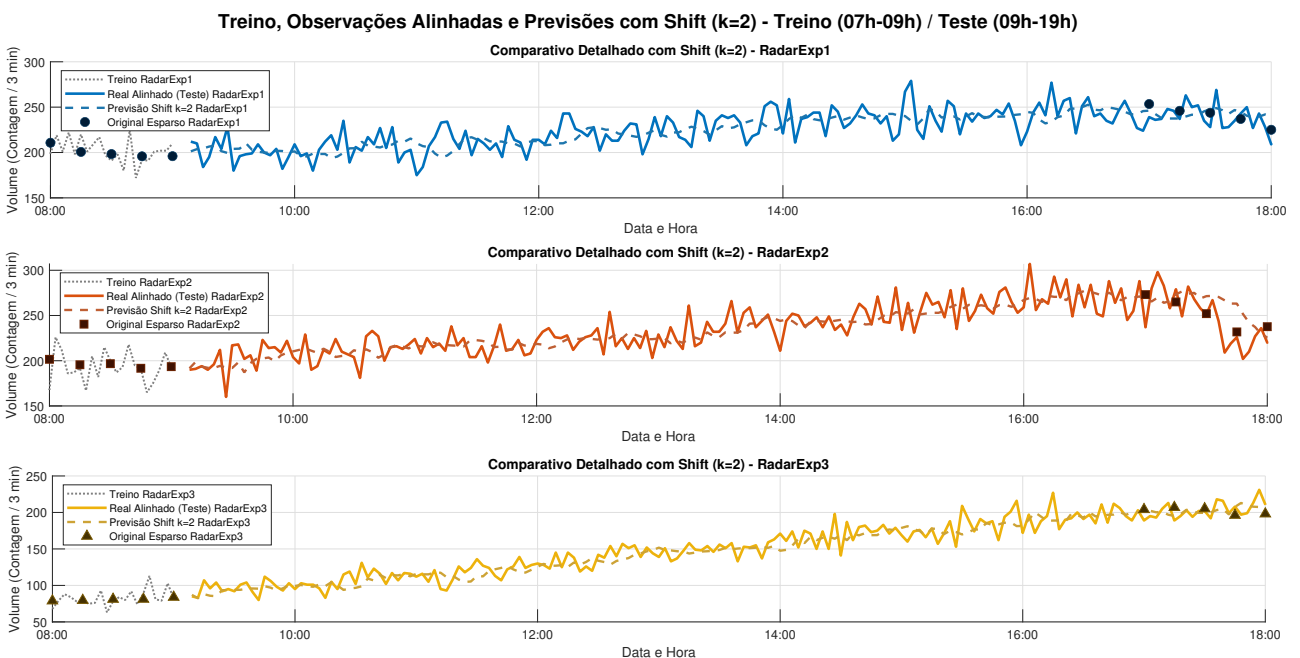


Figura 44 – Treino, Observações Alinhadas e Previsões

A Figura 45 apresenta uma comparação entre os dados expandidos com granularidade de 30 segundos (ou 0,5 min) e a taxa média horária original equivalente, para três radares distintos no período da manhã (07h às 09h). O objetivo da visualização é demonstrar o comportamento mais detalhado e volátil do tráfego veicular quando analisado em janelas temporais menores, em contraste com a suavização natural

observada na taxa média horária. A linha contínua representa os dados expandidos, revelando oscilações rápidas e abruptas no volume de veículos que seriam imperceptíveis na média horária, representada pela linha tracejada.

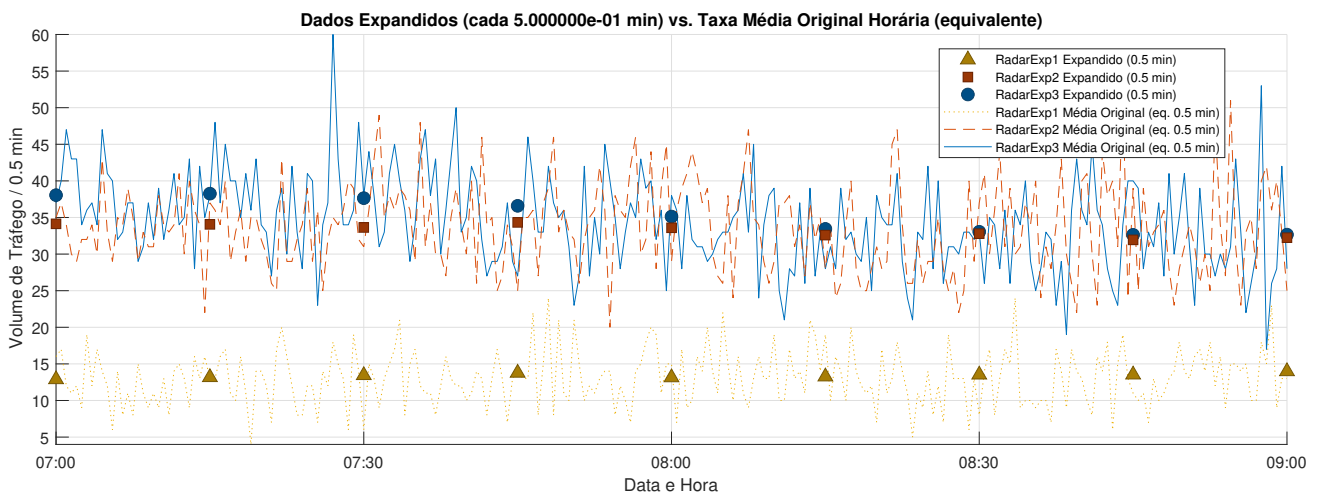


Figura 45 – Dados Expandidos vs. Taxa Média Original Horária - Período da manhã

Esse contraste evidencia a importância de escolher adequadamente a resolução temporal para modelagem e previsão de tráfego: enquanto os dados médios são úteis para tendências globais, os dados expandidos fornecem informações críticas sobre variações locais e comportamento de curto prazo do fluxo veicular fundamentais para aplicações em tempo real, como controle de semáforos adaptativos ou resposta a incidentes.

A Figura 46 exibe as séries temporais correspondentes aos dados de volume de tráfego coletados por três radares distintos, organizados em dois gráficos: o superior representa os dados utilizados para treinamento do modelo STARIMA (intervalo entre 07h e 08h), enquanto o inferior mostra os dados utilizados na fase de teste (08h às 09h). A visualização clara dos padrões individuais em cada radar permite observar variações de comportamento entre os pontos de coleta, o que reforça a natureza espacialmente heterogênea do tráfego urbano característica fundamental para a aplicação de modelos espaço-temporais como o STARIMA, que dependem da estrutura conjunta das relações temporais e espaciais para ajustar seus parâmetros de previsão.

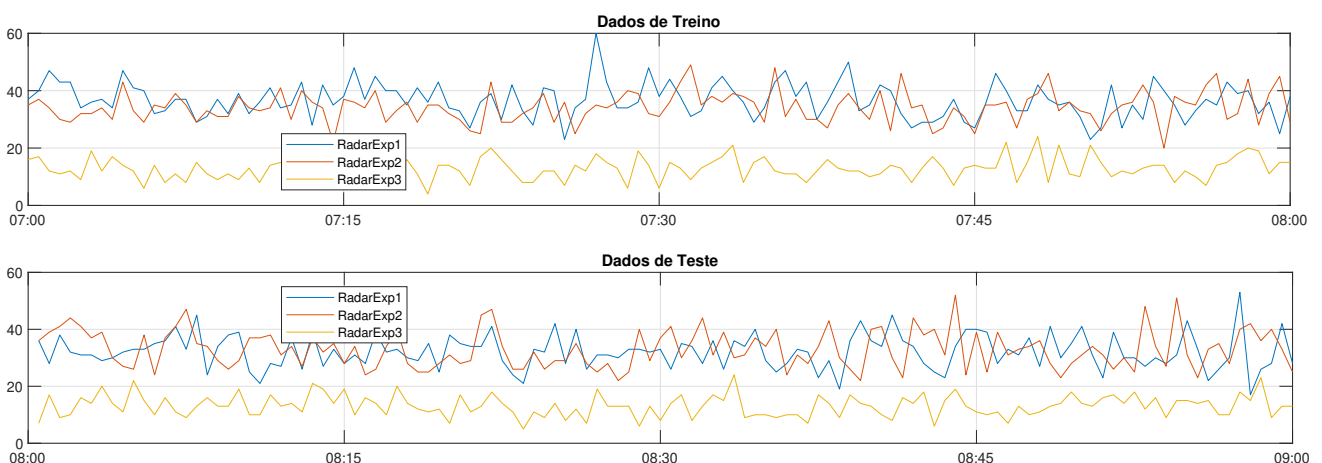


Figura 46 – Dados de Treino e Teste - Período da manhã

A Figura 47 apresenta as funções STACF e STPACF calculadas para os dados de treino (07h–08h) e teste (08h–09h), com o objetivo de identificar a estrutura de dependência temporal e espacial nas séries de volume de tráfego. No primeiro gráfico (superior), a função STACF é exibida em valor absoluto ao longo

dos *lags*, evidenciando que há uma autocorrelação significativa nos primeiros *lags*, especialmente para o radar L1, que apresenta valores mais elevados nos primeiros cinco *lags*. Essa característica indica a presença de forte dependência temporal de curto prazo, uma condição importante para a aplicação de modelos como STARIMA, que exploram justamente essa estrutura para realizar previsões precisas. O gráfico da figura 47 mostra a STPACF, também em valor absoluto, permitindo a identificação das contribuições diretas de cada *lag*, uma processo na escolha da ordem dos parâmetros autorregressivos ( $p$ ) do modelo. Os picos iniciais indicam a ordem até a qual o modelo deve considerar defasagens temporais significativas.

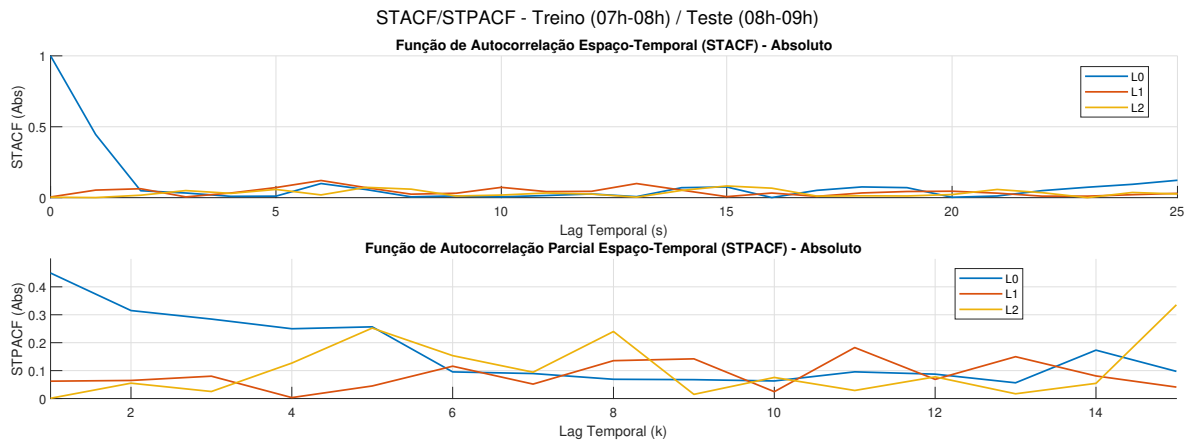


Figura 47 – STACF / STPACF Treino e Teste - Período da manhã

A Figura 48 apresenta os resultados da aplicação do modelo STARIMA ao conjunto de dados de tráfego urbano no período da manhã, com foco no intervalo de teste entre 08h e 09h, após o treino realizado de 07h às 08h. Os três painéis representam os resultados obtidos em três radares distintos (RadarExp1, RadarExp2 e RadarExp3), sendo comparadas as séries reais expandidas (linha pontilhada), os dados reais interpolados (linha contínua), a previsão gerada pelo modelo (linha sólida colorida) e as amostras originais esparsas (marcadores). O método adotado de previsão *step-by-step* evidencia a evolução do modelo ao longo do tempo de teste, utilizando as previsões anteriores como base para as subsequentes.

### Treino, Previsão Step-by-Step e Dados Originais (Treino (07h-08h) / Teste (08h-09h))

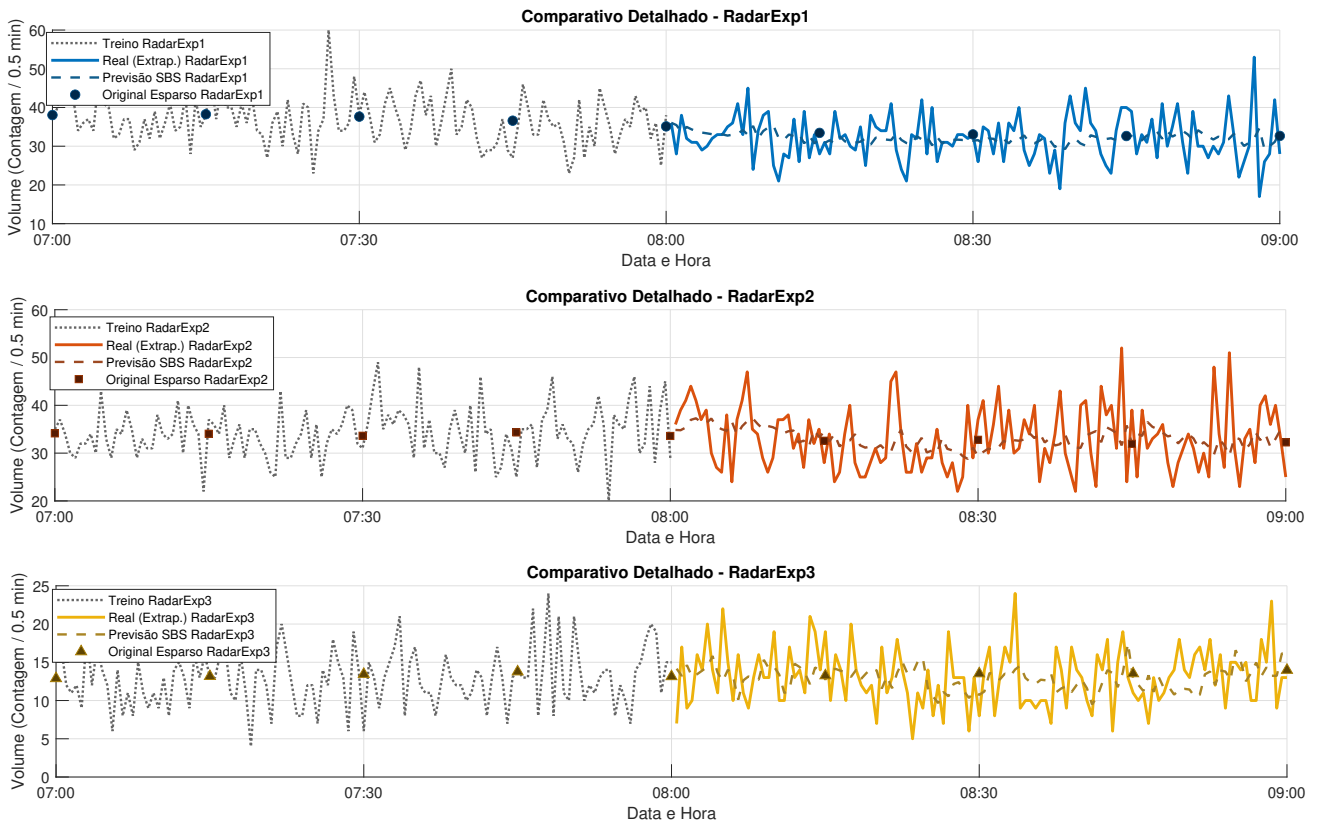


Figura 48 – Treino, Previsão e Dados Originais - Período da manhã

Visualmente, nota-se que o modelo consegue acompanhar razoavelmente bem a tendência dos dados em todos os radares, com destaque para o RadarExp1, cuja previsão (linha azul) permanece próxima da série real ao longo de todo o período. No RadarExp2 e, principalmente, no RadarExp3, observa-se maior dispersão e variações mais bruscas, o que indica maior complexidade local ou ruído nos dados, afetando a precisão das previsões. Ainda assim, o modelo STARIMA demonstra capacidade de representar padrões espaço-temporais com boa estabilidade. A representação gráfica desta figura é importante porque sintetiza o ciclo completo de aplicação do STARIMA: desde o uso de dados reais (ainda que esparsos), passando pela interpolação e expansão, até a geração das previsões e comparação com as medições.

A Figura 49 dá continuidade à avaliação do desempenho do modelo STARIMA, agora aplicando uma estratégia de alinhamento com shift temporal de dois passos ( $k = 2$ ) nas previsões.

## Treino, Observações Alinhadas e Previsões com Shift (k=2) - Treino (07h-08h) / Teste (08h-09h)

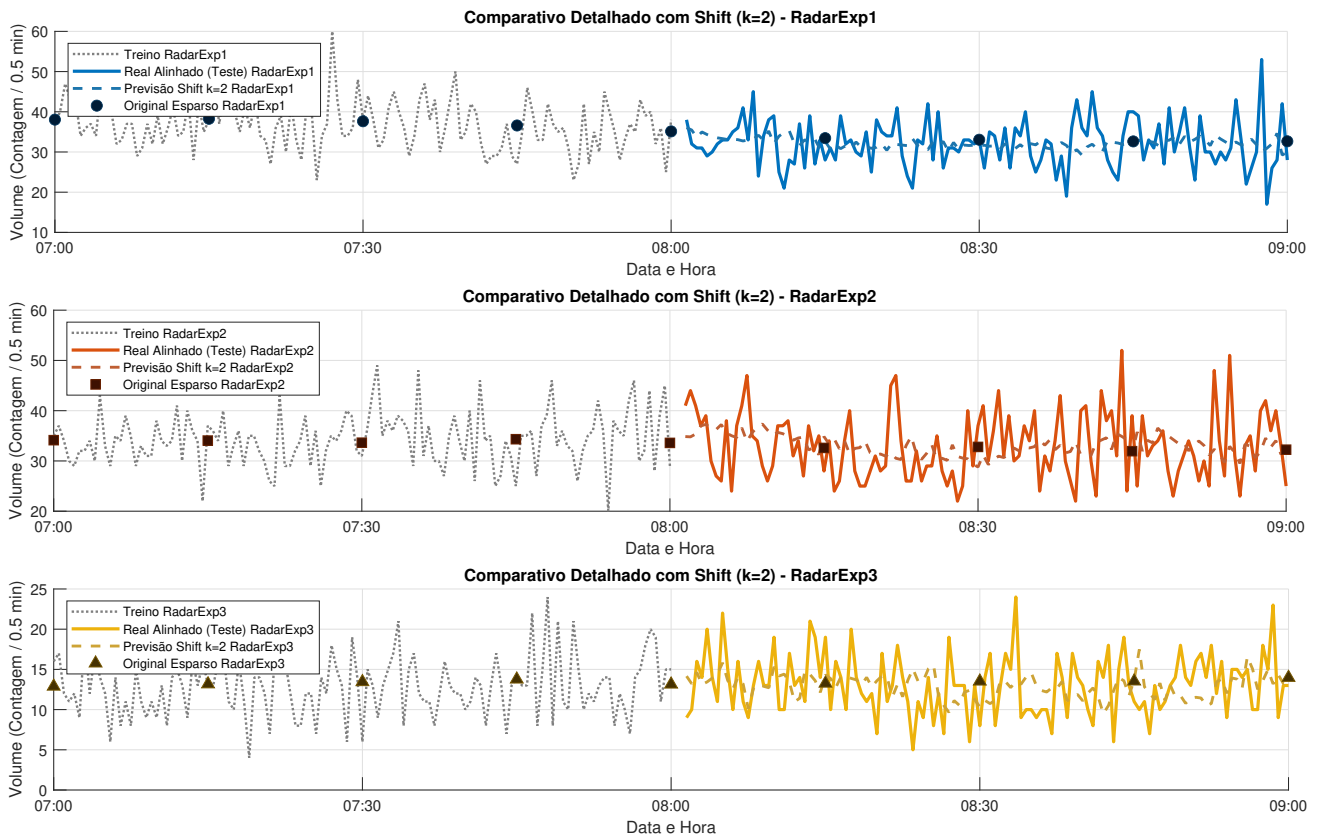


Figura 49 – Treino, Observações e Previsões - Período da manhã

Esta abordagem consiste em ajustar temporalmente as previsões geradas, comparando-as com os valores reais de tráfego deslocados em dois intervalos de tempo à frente. Trata-se de uma técnica usada para compensar atrasos sistemáticos do modelo na resposta aos dados reais.

Dando continuidade à abordagem estatística para previsão de tráfego urbano, inicia-se a aplicação da Rotina STARIMA Aperfeiçoada, cujo objetivo é proporcionar maior robustez ao processo preditivo por meio de ajustes refinados na modelagem espaço-temporal. A rotina contempla a geração de dados semi-sintéticos a partir de contagens horárias esparsas, expandidos com granularidade adequada para capturar variações de curto prazo.

Na Figura 50 utiliza a abordagem de uma rede neural do tipo LSTM na predição do volume de tráfego urbano com base nos três radares distintos. O modelo foi treinado no intervalo das 07h às 09h e validado em um período de teste estendido até às 19h, utilizando abordagem *step-by-step*.

Visualmente, os gráficos revelam uma boa aderência entre as curvas preditas (linha pontilhada) e as observações reais (linha contínua), evidenciando que o modelo foi capaz de capturar as variações sazonais e as tendências locais de forma satisfatória. As métricas numéricas confirmam esse desempenho: MAE de 11,908; RMSE de 15,121; EC de 0,901. Além disso, os indicadores NRMSE (normalizado) apontam valores baixos tanto com base na amplitude (0,0641) quanto na média (0,0728), o que sugere que o modelo possui boa generalização e é pouco afetado por variações extremas nas séries.

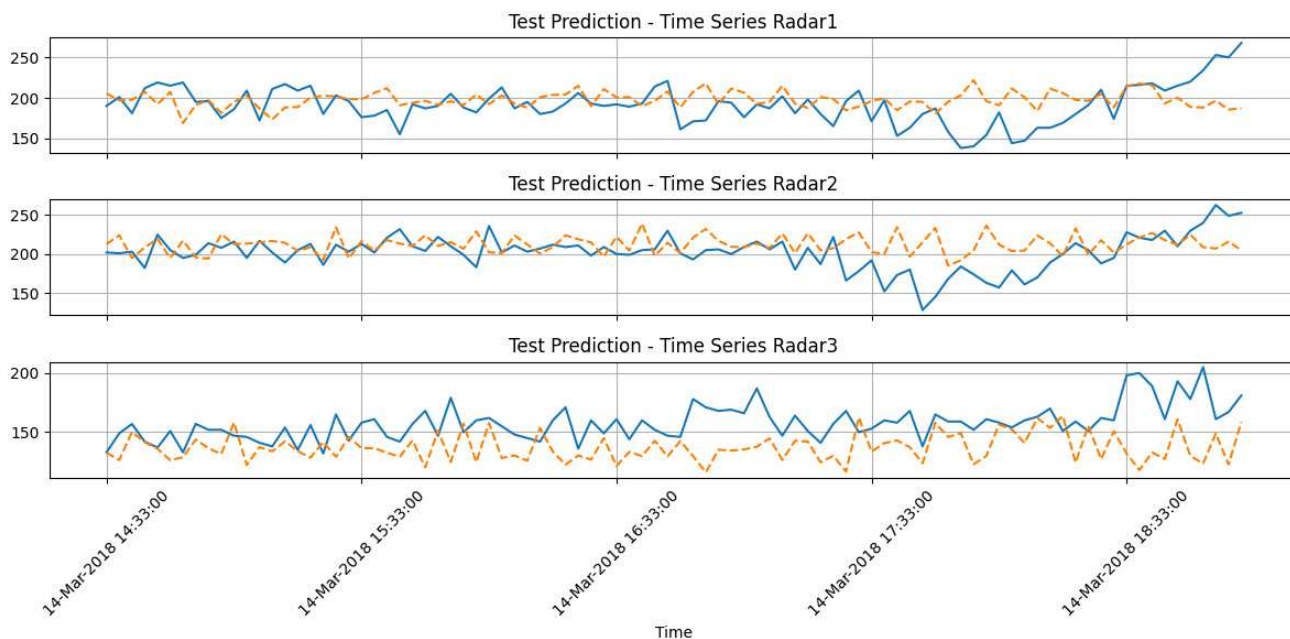


Figura 50 – Resultados em LSTM para 3 radares

As Figuras 51 e 52 apresentam os resultados obtidos por meio do modelo LSTM aplicado à previsão de volume de tráfego veicular em três radares distintos, considerando dois períodos distintos: manhã (Figura 51) e tarde (Figura 52). Em ambos os casos, são comparadas as séries reais (linhas azuis) com as previsões geradas pela LSTM (linhas laranja tracejadas), permitindo avaliar visualmente a aderência do modelo às flutuações do tráfego

Na Figura 51 observa-se que o desempenho da LSTM no período da manhã é razoável, com capacidade de seguir a tendência geral do tráfego, embora com dificuldades em capturar variações mais abruptas, especialmente nos radares 2 e 3. Os erros aparentes ocorrem principalmente em picos de tráfego repentino, que tendem a ser suavizados ou deslocados pela previsão. Apesar disso, a previsão mostra certa estabilidade e acompanha o padrão médio com coerência.

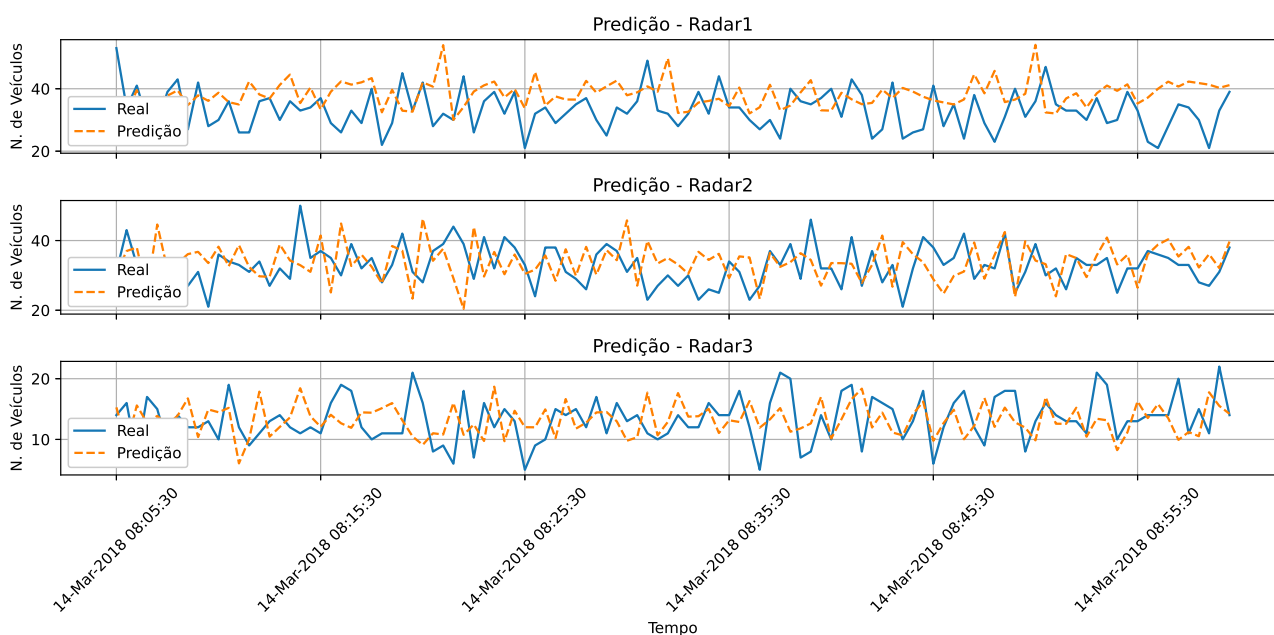


Figura 51 – Resultados em LSTM para 3 radares - Período da manhã

Já na Figura 52, correspondente ao período da tarde, a LSTM apresenta um desempenho visivelmente

inferior em relação ao período da manhã. Nota-se maior divergência entre as curvas reais e previstas, com subestimação frequente dos volumes reais, particularmente no Radar 1. Esse comportamento pode ser explicado por um padrão de tráfego mais irregular durante a tarde, ou ainda pela menor capacidade do modelo em generalizar para janelas temporais diferentes daquelas em que foi originalmente calibrado.

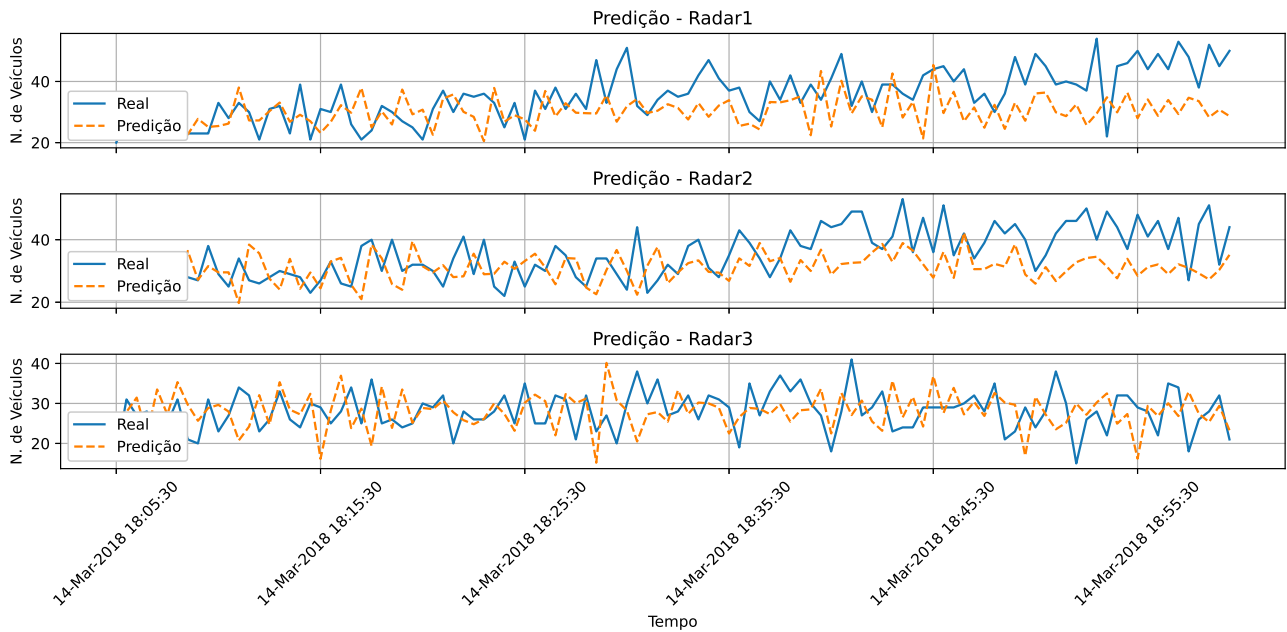


Figura 52 – Resultados em LSTM para 3 radares - Período da tarde

A Figura 53 apresenta a aplicação de redes neurais LSTM na previsão de séries temporais de volume de tráfego urbano captadas por quatro radares distintos. Cada gráfico exibe a série real (linha azul contínua) e a previsão gerada pelo modelo LSTM (linha laranja tracejada) para um radar específico. De forma geral, observa-se que o modelo é capaz de capturar a tendência de queda gradual do volume de tráfego nas últimas horas do dia, com desempenho visualmente satisfatório nas séries com as flutuações reais são suavemente acompanhadas pelas previsões. Esse comportamento indica que, embora o modelo LSTM consiga identificar tendências e padrões sazonais, ele demonstra certa limitação na captura de oscilações de curto prazo ou eventos atípicos. Ainda assim, os resultados demonstram que o modelo possui boa capacidade de generalização e pode ser útil como ferramenta de apoio na previsão de tráfego em ambientes urbanos.

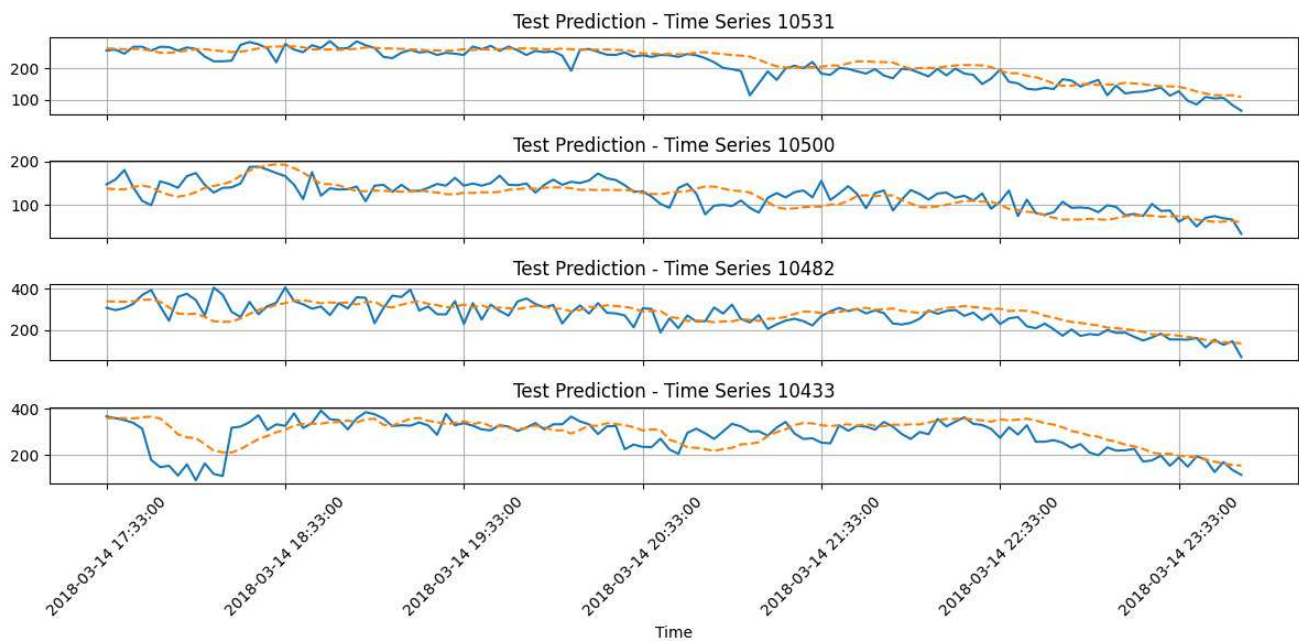


Figura 53 – Resultados em LSTM para 4 radares

## 5.6 Validação do Modelo LSTM

### 5.6.1 Treinamento

Os dados experimentais utilizados neste estudo consistem em observações de tráfego sintéticas coletadas por três sensores de radar, denominados Radar1, Radar2 e Radar3 em dois períodos do dia: manhã e tarde. Cada sensor registra observações em intervalos regulares de 30 segundos. O conjunto de dados completo foi dividido em subconjuntos de treinamento e teste, sendo as últimas 120 amostras (1h) reservadas exclusivamente para a fase de teste, a fim de avaliar o desempenho das previsões.

Antes do treinamento do modelo, os dados brutos passaram por diversas etapas de pré-processamento para garantir estabilidade numérica e um aprendizado eficaz. Primeiramente, foi aplicada a normalização utilizando a técnica *MinMaxScaler*, transformando todos os valores dos sinais de radar para o intervalo  $[0,1]$ . Esse passo reduz a variância entre os sensores e acelera a convergência durante o treinamento.

Em seguida, foi empregada uma abordagem de janela deslizante para estruturar os dados da série temporal em amostras supervisionadas. Especificamente, uma janela de comprimento fixo foi usada para gerar pares de entrada e saída. Cada sequência de entrada consistia em 10 passos de tempo consecutivos, sendo o alvo de predição os valores dos radares no passo de tempo seguinte.

O modelo preditivo utilizado é uma rede *Long Short-Term Memory*. O modelo recebe como entrada uma série temporal multivariada com dimensionalidade três — correspondente aos três sinais dos radares. A arquitetura da rede é composta por duas camadas LSTM empilhadas, cada uma com 64 unidades ocultas. O estado oculto final da segunda camada LSTM é usado para gerar a previsão do próximo passo temporal, produzindo três valores que representam as leituras previstas para o Radar1, Radar2 e Radar3.

O modelo foi treinado utilizando a função de perda MSE, que quantifica a média das diferenças ao quadrado entre os valores previstos e os reais dos radares. A otimização foi realizada com o otimizador Adam, utilizando uma taxa de aprendizado de 0,02. O treinamento empregou gradiente descendente estocástico em mini-batches com tamanho de lote igual a 16, totalizando 200 épocas de treinamento.

## 5.6.2 Resultados Numéricos

As Figuras 54 e 55 apresentam os resultados de predição da rede LSTM para os períodos da manhã e tarde, respectivamente. As Figuras 56 e 57 apresentam os resultados de predição da aplicação do método STARIMA para os períodos da manhã e tarde, respectivamente.

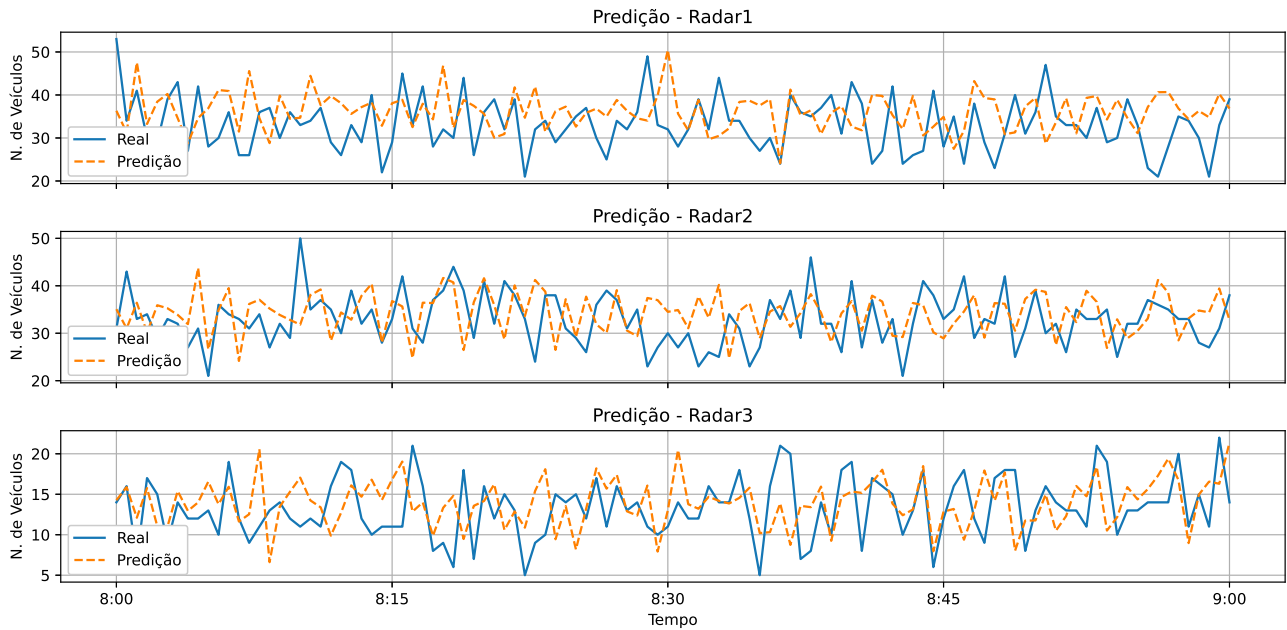


Figura 54 – Comparação entre valores reais e previstos do número de veículos para os radares 1, 2 e 3.

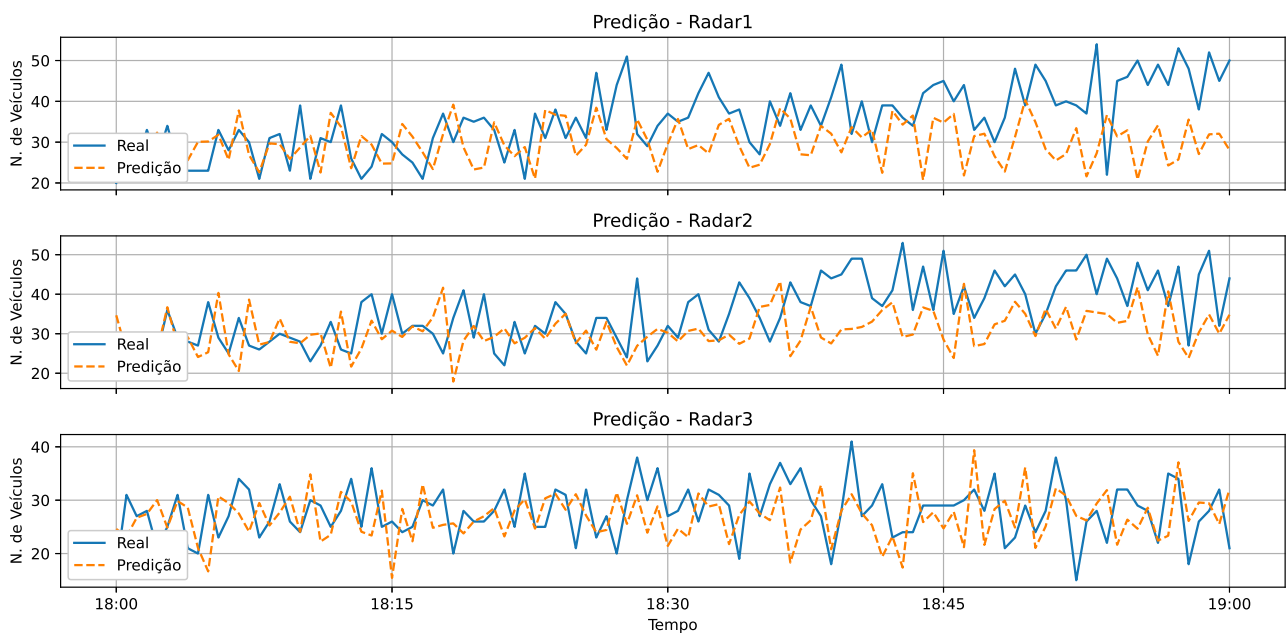


Figura 55 – Comparação entre os valores reais e previstos do número de veículos para os radares 1, 2 e 3.

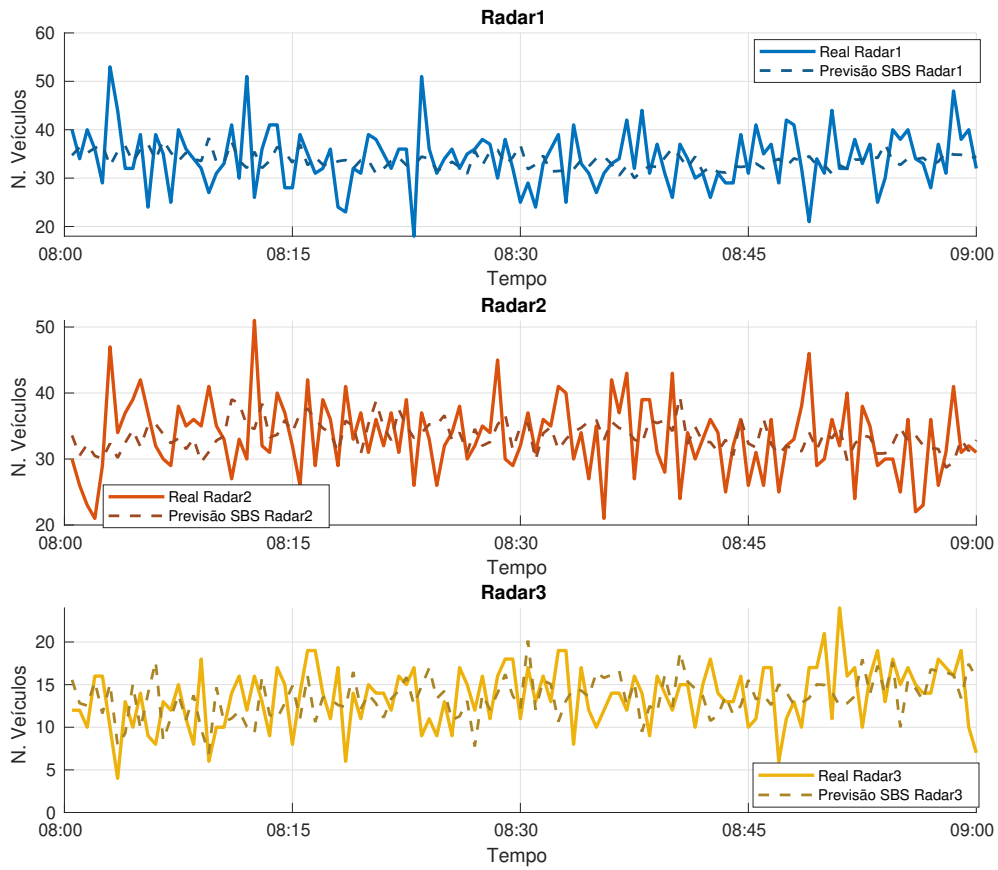


Figura 56 – Comparação entre dados reais e previsões do número de veículos nos radares 1, 2 e 3, entre 08:00 e 09:00.

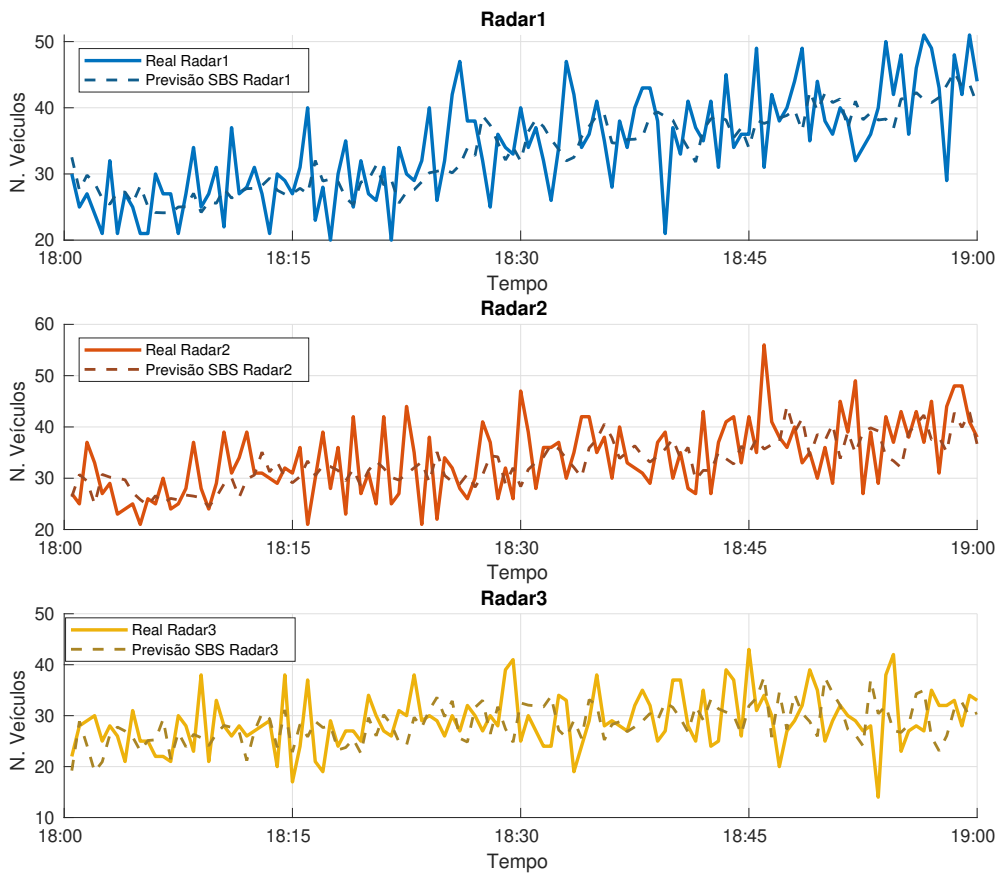


Figura 57 – Comparação entre dados reais e previsões do número de veículos nos radares 1, 2 e 3, entre 18:00 e 19:00.

De modo geral, comparando os resultados LSTM com STARIMA, observa-se que o modelo STARIMA apresenta um comportamento mais aderente à série temporal real, demonstrando maior exatidão nas previsões de curto prazo. Em contraste, a rede LSTM exhibe desvios mais acentuados e uma capacidade reduzida de acompanhar variações imediatas. Apesar disso, o modelo LSTM consegue refletir de forma satisfatória as tendências de longo prazo, capturando corretamente o comportamento geral da série.

Tabela 6 – Comparação de desempenho entre STARIMA e LSTM nos períodos da manhã e da tarde

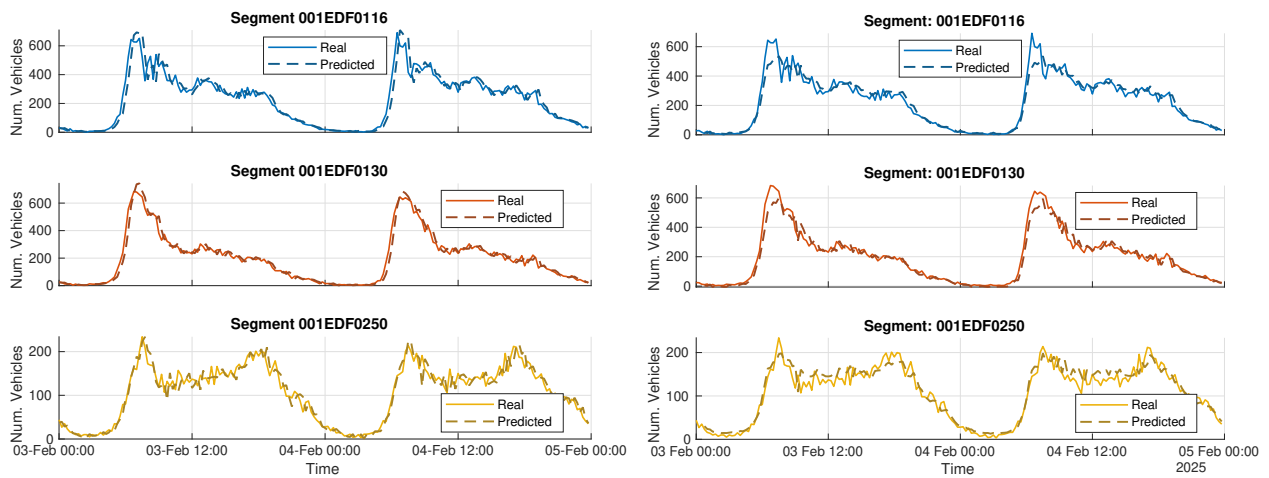
Período	Método	Radar	MAE	RMSE	NRMSE (range)   (mean)
Manhã	STARIMA	Radar1	4,889	6,407	0,1830   0,1878
		Radar2	4,877	6,028	0,2009   0,1812
		Radar3	3,329	4,099	0,2050   0,3014
	LSTM	Radar1	7,0189	8,5936	0,2685   0,2596
		Radar2	5,0397	6,7047	0,2312   0,2042
		Radar3	3,9158	4,7720	0,2807   0,3578
Tarde	STARIMA	Radar1	4,896	6,073	0,1687   0,1784
		Radar2	5,852	6,799	0,1943   0,1927
		Radar3	4,780	5,852	0,2018   0,2041
	LSTM	Radar1	8,3231	10,6047	0,3119   0,2986
		Radar2	7,1753	9,2953	0,2998   0,2624
		Radar3	5,6837	7,2717	0,2797   0,2598

A Tabela 6 apresenta uma comparação detalhada entre os métodos STARIMA e LSTM aplicados aos dados dos períodos da manhã e da tarde para os três radares. Observa-se que o modelo STARIMA apresentou desempenho superior ao da LSTM na maioria dos casos, especialmente em termos de MAE e RMSE. Durante o período da manhã, o STARIMA obteve menores valores de MAE e RMSE em todos os radares, destacando-se no Radar3, onde o RMSE foi de apenas 4,099, enquanto a LSTM apresentou 4,7720. O comportamento se repete no período da tarde, com o STARIMA mantendo um desempenho mais consistente e com menor variação entre os radares. A LSTM, por sua vez, apresentou erros significativamente mais altos, sobretudo no Radar1, com um RMSE de 10,6047 contra 6,073 do STARIMA. Além disso, os valores de NRMSE (normalizado tanto pela amplitude quanto pela média) reforçam a maior estabilidade e aderência do STARIMA às séries temporais observadas, evidenciando sua maior eficácia para a modelagem de séries de tráfego neste contexto.

Diante dos resultados apresentados na Tabela 6, observa-se um comportamento complementar entre os dois modelos avaliados. O modelo estatístico STARIMA apresenta desempenho superior nos períodos de curto prazo, com MAE e RMSE em praticamente todos os radares, especialmente no período da manhã. Essa vantagem decorre de sua estrutura paramétrica, que se beneficia da forte dependência temporal e espacial de curto alcance presente nas séries. Por outro lado, a rede neural LSTM, embora mais suscetível a desvios em janelas curtas, demonstra melhor capacidade de adaptação à variabilidade quando o horizonte de previsão é mais extenso, como evidenciado pela redução relativa dos erros em alguns períodos da tarde. Isso sugere que o modelo LSTM, por sua natureza não linear e flexível, tende a se ajustar mais adequadamente quando dispõe de uma janela temporal mais ampla para captar padrões complexos e não estacionários. Em contrapartida, o STARIMA, apesar de sua precisão em horizontes curtos, tem sua performance limitada quando há aumento na dispersão e irregularidade dos dados ao longo do tempo. Portanto, os resultados reforçam a ideia de que a escolha do modelo mais adequado está diretamente ligada ao horizonte temporal da previsão desejada: o STARIMA se mostra mais eficiente para previsões de curto prazo, enquanto a LSTM tende a apresentar melhor desempenho em contextos com maior abrangência temporal.

## 5.7 Predição de tráfego para o conjunto de dados 2 (Brasília)

O Conjunto de Dados 2 foi projetado para avaliar o desempenho da previsão de longo prazo. Ele é dividido em dois intervalos consecutivos de 2 dias, sendo o primeiro usado para treinamento e o segundo para testes. Essa configuração permite avaliar a capacidade de cada modelo de generalizar padrões de tráfego em períodos temporalmente distantes.



a) Previsões STARIMA no conjunto de dados 2

b) Previsões LSTM no conjunto de dados 2

Figura 58 – Previsões dos modelos a) STARIMA e b) LSTM no conjunto de dados 2 nos três segmentos de estrada com radares.

A Figura 58 descreve os resultados de previsão de médio prazo dos modelos LSTM e STARIMA, respectivamente, avaliados no Conjunto de Dados 2. Cada modelo foi treinado no primeiro intervalo de 2 dias e testado no intervalo de 2 dias subsequente, destacando sua capacidade de generalização em intervalos temporais estendidos.

Tabela 7 – Comparação de desempenho entre os modelos STARIMA e LSTM nas três séries de radares de Brasília

Segmento	Modelo	MAE	NRMSE <small>(<math>\frac{\text{range}}{\text{mean}}</math>)</small>
001EDF0116	STARIMA	31.693	0.0758 / 0.2311
	<b>LSTM</b>	<b>30.378</b>	<b>0.0761 / 0.2207</b>
001EDF0130	<b>STARIMA</b>	<b>22.103</b>	<b>0.050 / 0.173</b>
	LSTM	25.424	0.057 / 0.185
001EDF0250	STARIMA	12.670	0.074 / 0.157
	<b>LSTM</b>	<b>12.372</b>	<b>0.073 / 0.150</b>

A Tabela 8 fornece uma comparação quantitativa dos modelos STARIMA e LSTM nos três segmentos de radar selecionados. Os resultados dos testes/previsões revelam que ambos os modelos alcançaram desempenho comparável em termos de MAE e NRMSE. O STARIMA superou ligeiramente o LSTM nos segmentos 001EDF0130 e 001EDF0116, enquanto o LSTM apresentou desempenho marginalmente melhor no segmento 001EDF0250.

Embora o STARIMA tenha se beneficiado, de modo geral, de sua capacidade de modelar dependências espaciais e dinâmicas temporais de curto prazo, sua vantagem diminuiu em previsões de longo prazo, onde tais dependências enfraquecem ao longo do tempo. Por outro lado, o modelo LSTM, apesar de ser mais suscetível a ruídos e exigir mais dados, mostrou-se eficaz na captura de tendências temporais mais amplas, o que explica seus resultados competitivos em todos os segmentos. O conjunto de dados 2, com o primeiro

intervalo de 2 dias dedicado ao treinamento e o intervalo subsequente de 2 dias para teste/previsão, representa um cenário de previsão de tráfego urbano de médio prazo.

O NRMSE calculado em relação à amplitude e à média, destaca ainda mais a estabilidade de ambos os modelos. O STARIMA tendeu a oferecer previsões ligeiramente mais consistentes em segmentos com correlação espacial mais forte, enquanto o LSTM manteve sua robustez em segmentos com dinâmica temporal mais pronunciada.

A Tabela e a Figura apresentam o desempenho preditivo dos modelos STARIMA e LSTM em três segmentos rodoviários monitorados por radar em Brasília. Os experimentos comparam modelos LSTM treinados com 2 e 4 dias de dados, bem como modelos STARIMA usando 2 e 4 dias de treinamento, e os avaliam em um horizonte de previsão de 10 dias.

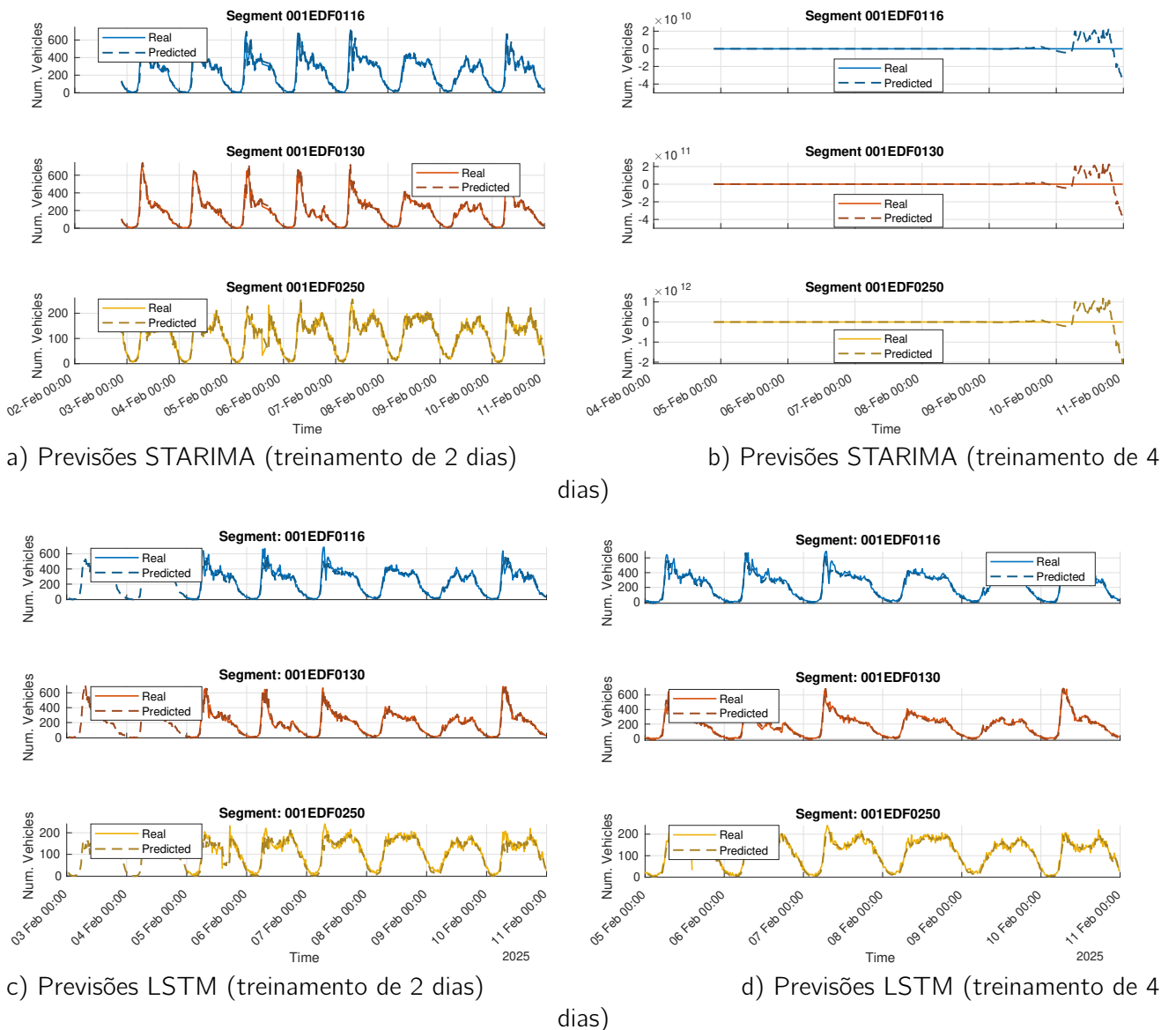


Figura 59 – Previsões dos modelos STARIMA e LSTM nos segmentos do radar de Brasília, comparando dados de treinamento de 2 e 4 dias. Os painéis a) e b) mostram os resultados do STARIMA, enquanto os painéis c) e d) mostram os resultados do LSTM.

Para o segmento 001EDF0116 (Brasília), o modelo LSTM treinado com 4 dias obteve a melhor precisão geral, com o menor MAE (28,85) e NRMSE (0,0626 / 0,1835). O modelo STARIMA com treinamento de 2 dias obteve um MAE ligeiramente maior (30,13), enquanto a variante STARIMA de 4 dias apresentou desempenho insatisfatório, com um erro significativo ( $MAE = 2,2 \times 10^9$ ),  $NRMSE = 8,8 \times 10^6 / 2,6 \times 10^7$ ), sugerindo instabilidade numérica ou sobreajuste.

No segmento 001EDF0130 (rodovia Brasília), o LSTM (4 dias) novamente superou as configurações LSTM e STARIMA de 2 dias, obtendo um MAE de 25,43. O STARIMA com treinamento de 2 dias apresentou desempenho razoável (MAE = 24,09), mas o STARIMA de 4 dias apresentou degradação severa (MAE =  $2,3 \times 10^{10}$ ), confirmando sua inconsistência com o aumento do tamanho do treinamento.

Para o segmento 001EDF0250 (Estrada Brasília), o STARIMA (2 dias) apresentou bom desempenho (MAE = 13,90), e o LSTM (2 dias) apresentou resultados semelhantes (MAE = 13,64). No entanto, o STARIMA com treinamento de 4 dias novamente falhou drasticamente e tornou-se inviável em tal horizonte temporal de previsão (MAE =  $1,1 \times 10^{11}$ , NRMSE =  $1,3 \times 10^9 / 2,8 \times 10^9$ ).

No geral, os números confirmam que aumentar a duração do treinamento de 2 para 4 dias melhora a capacidade de previsão do LSTM, especialmente para segmentos complexos. No entanto, o desempenho do STARIMA se deteriora acentuadamente quando treinado por 4 dias devido à instabilidade dos parâmetros durante o processo de otimização. Esses resultados destacam a robustez e a escalabilidade de modelos de aprendizado profundo, como o LSTM, para lidar com horizontes de treinamento mais longos, ao mesmo tempo em que reforçam a inadequação de modelos paramétricos como o STARIMA para previsões de longo prazo.

Tabela 8 – Comparação de desempenho entre os modelos STARIMA e LSTM nas três séries de radares de Brasília; Resultados de treinamento de 2 dias vs. treinamento de 4 dias

Segmento	Modelo	MAE	NRMSE (range / mean)
001EDF0116	STARIMA (2-day)	30.126	0.0734 / 0.2221
	STARIMA (4-day)	$2,22 \times 10^9$	$8,78 \times 10^6 / 2,63 \times 10^7$
	LSTM (2-day)	33.000	0.0718 / 0.2138
	<b>LSTM (4-day)</b>	<b>28.847</b>	<b>0.0626 / 0.1835</b>
001EDF0130	<b>STARIMA (2-day)</b>	<b>24.088</b>	0.0582 / 0.2121
	STARIMA (4-day)	$2,27 \times 10^{10}$	$9,13 \times 10^7 / 3,40 \times 10^8$
	LSTM (2-day)	26.528	0.0586 / 0.2103
	<b>LSTM (4-day)</b>	25.428	<b>0.0527 / 0.1922</b>
001EDF0250	<b>STARIMA (2-day)</b>	13.897	<b>0.0798 / 0.1693</b>
	STARIMA (4-day)	$1,14 \times 10^{11}$	$1,33 \times 10^9 / 2,80 \times 10^9$
	<b>LSTM (2-day)</b>	<b>13.638</b>	0.0821 / 0.1754
	LSTM (4-day)	13.824	0.0874 / 0.1849

A drástica degradação do desempenho do modelo STARIMA com treinamento de 4 dias, evidenciada pelos valores extremos de erro na Tabela 8, pode ser atribuída à instabilidade no procedimento de estimativa de parâmetros do modelo. O STARIMA se baseia em ajustes não lineares de mínimos quadrados para estimar coeficientes autorregressivos e de defasagem espaço-temporal. Quando a janela de treinamento é estendida para quatro dias, a matriz de dados se torna maior e potencialmente mais mal condicionada, especialmente se os padrões de tráfego contiverem alta variabilidade (não estacionários) ou colinearidade entre localizações espaciais. Isso pode levar à divergência no processo de otimização, onde as equações normais se tornam instáveis ou a solução superajusta o ruído em vez de capturar com precisão a dinâmica temporal real. Nesses casos, pequenas imprecisões numéricas podem ser amplificadas, resultando em coeficientes com magnitudes extremamente grandes e previsões que estão várias ordens de magnitude distantes dos valores reais. Isso sugere que o STARIMA, embora eficaz em configurações de curto prazo, pode exigir técnicas de regularização, estimativa robusta de parâmetros ou normalização de dados para manter a estabilidade numérica quando treinado em horizontes de tempo mais longos.

# 6 CONCLUSÕES E PERSPECTIVAS

A crescente complexidade da mobilidade urbana, impulsionada pelo aumento do número de veículos e pela dinâmica do tráfego nas cidades, exige soluções analíticas precisas e adaptáveis aos diferentes contextos temporais e espaciais. Este trabalho apresentou uma abordagem comparativa entre dois modelos preditivos distintos: o modelo STARIMA, que explora dependências espaço-temporais lineares, e o modelo LSTM, baseado em redes neurais recorrentes capazes de capturar relações não lineares e de longo prazo em séries temporais.

Através de uma base experimental construída a partir de dados reais de radares em vias urbanas, que foram interpolados e expandidos para permitir análises em diferentes escalas temporais com maior granularidade, a metodologia envolveu desde a pré-modelagem com testes de estacionariedade, construção de matrizes espaciais e análise de autocorrelação, até a calibração de parâmetros, previsão e validação dos resultados por métricas como MAE, RMSE e EC.

Os resultados evidenciaram um comportamento diferenciado entre os modelos conforme o período analisado. O modelo paramétrico para o Conjunto de Dados 1, o modelo STARIMA demonstrou maior eficiência em previsões de curto prazo, especialmente no período da manhã (07h–09h), quando os padrões de tráfego são mais regulares e previsíveis. Sua estrutura paramétrica permitiu capturar com precisão as flutuações locais e temporais, resultando em baixos níveis de erro e resíduos estatisticamente não significativos. A forte autocorrelação nas séries temporais analisadas, característica bem aproveitada pela modelagem espaço-temporal clássica, favoreceu a aplicabilidade do STARIMA. Em contraste, o modelo LSTM mostrou maior capacidade de adaptação em períodos de maior variabilidade, como no intervalo da tarde (17h–19h), onde o tráfego urbano é influenciado por múltiplos fatores contextuais, como condições climáticas, eventos locais e comportamentos não sistemáticos dos condutores. Nesses cenários, a arquitetura da rede neural revelou-se mais flexível para lidar com irregularidades e capturar padrões não lineares, mesmo que sua performance em janelas curtas não tenha superado a do STARIMA. Esta análise comparativa sublinha a natureza complementar dos dois modelos. Enquanto o STARIMA é ideal para previsões de curto alcance com alta precisão em ambientes estáveis, o LSTM oferece vantagens em horizontes de tempo mais amplos ou em cenários de maior complexidade dinâmica. Essa complementaridade sugere a possibilidade de abordagens híbridas, que combinem as forças de ambos os modelos para aprimorar a acurácia e a robustez das previsões.

Em contraste, para o Conjunto de Dados 2 (Brasília), projetado para previsões de médio prazo em intervalos consecutivos de vários dias, a diferença de desempenho entre STARIMA e LSTM diminuiu consideravelmente. Ambos os modelos alcançaram precisão preditiva comparável, com o LSTM superando ligeiramente o STARIMA em alguns casos e vice-versa. Isso sugere que a capacidade do LSTM de aprender padrões temporais complexos e dependências de longo prazo oferece uma vantagem quando os dados de tráfego apresentam maior variabilidade ao longo de períodos prolongados ou quando as correlações espaciais se tornam menos pronunciadas. Além disso, experimentos que variaram a duração dos dados de treinamento (2 dias vs. 4 dias) mostraram que o aumento dos dados de treinamento geralmente beneficia os modelos LSTM devido à sua abordagem de aprendizado orientada por dados.

Os resultados reforçam o papel da modelagem preditiva como instrumento estratégico para a gestão de tráfego urbano em cidades inteligentes. As análises conduzidas demonstram que a escolha do modelo preditivo deve considerar fatores como o volume e a qualidade dos dados disponíveis, o horizonte de previsão e as características dinâmicas do ambiente monitorado. A capacidade de prever o comportamento do tráfego permite que gestores tomem decisões proativas, otimizando o fluxo veicular, reduzindo congestionamentos

e contribuindo para a sustentabilidade e a qualidade de vida nas cidades.

## 6.1 Trabalhos Futuros

Com base nos resultados e discussões apresentados, diversas direções para trabalhos futuros podem ser exploradas para aprimorar a modelagem e a previsão do tráfego urbano:

### 6.1.1 Modelos Híbridos e Ensembles

Uma área promissora é o desenvolvimento e a avaliação de modelos híbridos que combinem as vantagens do STARIMA e do LSTM. Por exemplo, um modelo pode utilizar o STARIMA para capturar as dependências lineares de curto prazo e o LSTM para modelar os padrões não lineares e de longo prazo, ou para processar os resíduos do STARIMA. A integração de técnicas de ensemble learning, onde múltiplos modelos são combinados para produzir uma previsão mais robusta, também pode ser investigada. Isso poderia levar a uma melhoria significativa na acurácia e na capacidade de generalização das previsões em diferentes condições de tráfego.

### 6.1.2 Integração de Fontes de Dados Heterogêneas

Expandir as fontes de dados utilizadas nos modelos preditivos é crucial. Além dos dados de volume e velocidade de radares, futuros trabalhos podem incorporar informações de GPS de veículos, dados de transporte público, informações de redes sociais (para eventos que afetam o tráfego), dados meteorológicos em tempo real, e informações sobre obras e incidentes. A fusão desses dados heterogêneos, possivelmente utilizando técnicas de data mining e machine learning para extrair características relevantes, pode enriquecer os modelos e torná-los mais resilientes a variações inesperadas no tráfego.

### 6.1.3 Aplicações em Tempo Real e Computação de Borda

Explorar a implementação dos modelos STARIMA e LSTM em plataformas de edge computing ou diretamente em dispositivos IoT (Internet das Coisas) para permitir previsões e tomadas de decisão em tempo real. Isso reduziria a latência e a dependência de infraestruturas de nuvem centralizadas, tornando os sistemas de gestão de tráfego mais ágeis e eficientes. A otimização dos algoritmos para execução em ambientes com recursos computacionais limitados seria um foco importante.

### 6.1.4 Expansão para Outros Desafios de Mobilidade Urbana

Os modelos e metodologias desenvolvidos para a previsão de tráfego veicular podem ser adaptados e aplicados a outros desafios da mobilidade urbana. Isso inclui a previsão do fluxo de pedestres em áreas de grande circulação, a otimização de rotas de transporte público em tempo real, a gestão de frotas de veículos autônomos, e a previsão da demanda por serviços de micromobilidade (e.g., bicicletas e patinetes compartilhados). A versatilidade desses modelos pode contribuir para uma gestão mais integrada e inteligente de todo o ecossistema de mobilidade urbana.

### 6.1.5 Interpretabilidade e Explicabilidade (XAI)

Para modelos mais complexos, como as redes LSTM, a pesquisa em Explainable AI (XAI) é fundamental. Desenvolver métodos para tornar as previsões desses modelos mais interpretáveis e transparentes para os gestores de tráfego permitiria uma maior confiança nas decisões baseadas em IA. Compreender o raciocínio

por trás das previsões pode ser crucial para a validação e aprimoramento contínuo dos sistemas de gestão de tráfego.

### 6.1.6 Otimização de Parâmetros e Arquiteturas de Modelos

Investigar métodos mais avançados para a otimização dos hiperparâmetros e arquiteturas dos modelos STARIMA e LSTM. Isso pode incluir o uso de algoritmos genéticos, otimização bayesiana ou outras técnicas de busca automatizada para encontrar as configurações ideais que maximizem a performance preditiva. A exploração de novas arquiteturas de redes neurais, como Transformers ou Graph Neural Networks (GNNs), que são particularmente adequadas para dados espaço-temporais, também pode ser um caminho promissor.

### 6.1.7 Análise de Cenários e Simulações Preditivas

Desenvolver ferramentas que permitam a simulação de diferentes cenários de tráfego (e.g., impacto de grandes eventos, fechamento de vias, mudanças na infraestrutura) utilizando os modelos preditivos. Isso possibilitaria que os planejadores urbanos avaliassem o impacto de suas decisões antes da implementação, otimizando o planejamento de longo prazo e a resiliência da rede viária. A criação de um ambiente de simulação interativo, onde diferentes políticas de gestão de tráfego possam ser testadas, seria de grande valor.

### 6.1.8 Impacto de Eventos Extremos e Anomalias

Estudar a robustez dos modelos preditivos frente a eventos extremos e anomalias no tráfego, como acidentes graves, desastres naturais ou grandes manifestações. Desenvolver mecanismos para que os modelos possam detectar e se adaptar rapidamente a essas situações inesperadas, fornecendo previsões mais precisas e auxiliando na resposta a emergências. A incorporação de técnicas de detecção de anomalias e aprendizado contínuo pode ser explorada para este fim.

Essas direções de pesquisa, baseadas nas descobertas e limitações do presente trabalho, visam contribuir para o avanço da modelagem preditiva do tráfego urbano, tornando-a uma ferramenta ainda mais poderosa e indispensável para a construção de cidades inteligentes e sustentáveis.

# Referências

- [1] Shyam Shankaran R and Logesh Rajendran. Real-time adaptive traffic control system for smart cities. In *2021 International Conference on Computer Communication and Informatics (ICCCI)*, pages 1–6, 2021. Citado 5 vezes nas páginas [8](#), [18](#), [19](#), [20](#) e [21](#).
- [2] Alex Graves. *Supervised Sequence Labelling with Recurrent Neural Networks*. Studies in Computational Intelligence 385. Springer-Verlag Berlin Heidelberg, 1 edition, 2012. Citado 5 vezes nas páginas [8](#), [33](#), [34](#), [35](#) e [36](#).
- [3] Companhia de Engenharia de Tráfego. *Desempenho do Sistema Viário Principal - 2012: Volume e Velocidade*. Companhia de Engenharia de Tráfego – CET, Diretoria de Planejamento e Educação no Trânsito – DP, Superintendência de Planejamento – SPL, Gerência de Planejamento, Logística e Estudos de Tráfego – GPL, Departamento de Simulação e Pesquisa de Tráfego – DSP, São Paulo, Brazil, March 2013. Relatório técnico sobre a pesquisa de monitoração da fluidez viária em São Paulo, abrangendo o ano de 2012. Citado 7 vezes nas páginas [8](#), [9](#), [54](#), [55](#), [71](#), [72](#) e [77](#).
- [4] Domênico Zema, Marcella Gomes da Costa Pereira, and Yuri Gil Ferreira. Aplicação de arima em previsão de tráfego de curto prazo. Trabalho de conclusão de curso (graduação), Escola Politécnica, Universidade de São Paulo, São Paulo, 2020. Citado 5 vezes nas páginas [9](#), [23](#), [46](#), [70](#) e [73](#).
- [5] Dariusz Swiatek, Melissa S Pokorny, Jilmar A Tatto, VAZC Melo, and EM Dias. Traffic management solutions in large cities-the integrated centre of urban mobility (cimu) in são paulo. In *Proceedings of the 18th International Conference on Systems*, 2014. Citado 2 vezes nas páginas [14](#) e [17](#).
- [6] J Sloboden, J Lewis, V Alexiadis, YC Chiu, and E Nava. Guidebook on the utilization of dynamic traffic assignment in modeling. *Federal Highway Administration, US Department of Transportation*, 120, 2012. Citado 2 vezes nas páginas [14](#) e [17](#).
- [7] Eleni I Vlahogianni. Computational intelligence and optimization for transportation big data: challenges and opportunities. *Engineering and Applied Sciences Optimization: Dedicated to the Memory of Professor MG Karlaftis*, pages 107–128, 2015. Citado 2 vezes nas páginas [14](#) e [17](#).
- [8] Matthew G Karlaftis and Eleni I Vlahogianni. Statistical methods versus neural networks in transportation research: Differences, similarities and some insights. *Transportation Research Part C: Emerging Technologies*, 19(3):387–399, 2011. Citado 2 vezes nas páginas [14](#) e [17](#).
- [9] Y. Kamarianakis and P. Prastacos. Space–time modeling of traffic flow. *Computers & Geosciences*, 31(2):119–133, 2005. Geospatial Research in Europe: AGILE 2003. Citado 2 vezes nas páginas [14](#) e [29](#).
- [10] Yiannis Kamarianakis and Poulicos Prastacos. Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches. *Transportation Research Record*, 1857(1):74–84, 2003. Citado 5 vezes nas páginas [17](#), [25](#), [26](#), [27](#) e [28](#).
- [11] Eleni I. Vlahogianni, John C. Golias, and Matthew G. Karlaftis and. Short-term traffic forecasting: Overview of objectives and methods. *Transport Reviews*, 24(5):533–557, 2004. Citado na página [17](#).
- [12] Mohammed Sarrab, Supriya Pulparambil, and Medhat Awadalla. Development of an iot based real-time traffic monitoring system for city governance. *Global Transitions*, 2:230–245, 2020. Citado na página [17](#).
- [13] Shanthi D L, Keshava Prasanna, Vishwas Desai, Sakshi Agarwal, V Manish M Shetty, and Rakesh A S. Traffic prediction system using iot in smart city perspective. In *2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNBC)*, pages 1–6, 2021. Citado na página [18](#).
- [14] Vikram Bali, Sonali Mathur, Vishnu Sharma, and Dev Gaur. Smart traffic management system using iot enabled technology. In *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*, pages 565–568, 2020. Citado na página [18](#).
- [15] Attila M. Nagy and Vilmos Simon. Survey on traffic prediction in smart cities. *Pervasive and Mobile Computing*, 50:148–163, 2018. Citado na página [21](#).
- [16] Ambreen Joyo, Raziq Yaqub, and Nicholas Madamopoulos. Intelligent traffic-lights management by exploiting smart antenna technology (itsat). *IEEE Intelligent Transportation Systems Magazine*, 13(4):154–163, 2021. Citado na página [21](#).

- [17] R. Islam-Khan, B. Landfeldt, and A. Damdhare. Predicting travel times in dense and highly varying road traffic networks using starima models. Technical report, School of Information Technologies, University of Sydney, February 2012. Citado 5 vezes nas páginas 22, 24, 25, 27 e 28.
- [18] OLÍMPIO MENDES DE BARROS. Estimação das características de tráfego em tempo próximo ao real para uso em sistemas de previsão e controle de tráfego. Technical report, Escol Politécnica da USP, São Paulo, February 2022. Citado na página 24.
- [19] Xinyu Min, Jianming Hu, and Zuo Zhang. Urban traffic network modeling and short-term traffic flow forecasting based on gstarima model. In *13th International IEEE Conference on Intelligent Transportation Systems*, pages 1535–1540, 2010. Citado na página 25.
- [20] Peibo Duan, Guoqiang Mao, Changsheng Zhang, and Shangbo Wang. Starima-based traffic prediction with time-varying lags. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1610–1615, 2016. Citado na página 29.
- [21] Peibo Duan, Guoqiang Mao, Shangbo Wang, Changsheng Zhang, and Bin Zhang. Starima-based traffic prediction with time-varying lags, 2017. Citado na página 29.
- [22] Xinyu Min, Jianming Hu, Qi Chen, Tongshuai Zhang, and Yi Zhang. Short-term traffic flow forecasting of urban network based on dynamic starima model. In *2009 12th International IEEE Conference on Intelligent Transportation Systems*, pages 1–6, 2009. Citado na página 30.
- [23] Wenchang Duan and Zhiqiang Gong. Research and application of urban real-time traffic flow prediction based on starima. In *2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, pages 1521–1525, 2022. Citado na página 30.
- [24] Peibo Duan, Guoqiang Mao, Wenwei Yue, and Shangbo Wang. A unified starima based model for short-term traffic flow prediction. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 1652–1657, 2018. Citado 2 vezes nas páginas 30 e 31.
- [25] Athanasios Salamanis, Polykarpos Meladianos, Dionysios Kehagias, and Dimitrios Tzovaras. Evaluating the effect of time series segmentation on starima-based traffic prediction model. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 2225–2230, 2015. Citado 2 vezes nas páginas 31 e 32.
- [26] Charu C. Aggarwal. *Neural Networks and Deep Learning: A Textbook*. Springer, 2018. Citado na página 33.
- [27] Sony Sumaryo, A. Halim, and K. Ramli. Improved discrete event simulation model of traffic light control on a single intersection. In *2013 International Conference on QiR*, pages 116–120, 2013. Citado na página 37.
- [28] H. Murat Gursu, M. Cagatay Moroglu, Mikhail Vilgelm, Federico Clazzer, and Wolfgang Kellerer. System Level Integration of Irregular Repetition Slotted ALOHA for Industrial IoT in 5G New Radio. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, 2019-September:2–8, 2019. Citado na página 37.
- [29] S. Sharma. *Applied Multivariate Techniques*. John Wiley & Sons, 1996. Citado 3 vezes nas páginas 40, 41 e 50.
- [30] G. James, D. Witten, T. Hastie, and R. Tibshirani. *An Introduction to Statistical Learning*. Springer, 2013. Citado 3 vezes nas páginas 40, 41 e 50.
- [31] W.W.S. Wei. *Multivariate Time Series Analysis and Applications*. Wiley Series in Probability and Statistics. Wiley, 2019. Citado na página 46.
- [32] Departamento de Estradas de Rodagem do Distrito Federal (DER/DF). Sistema rodoviário do distrito federal – srdf 2020. Technical report, Governo do Distrito Federal, Secretaria de Estado de Transporte e Mobilidade, Brasília, DF, Brazil, 2020. Coordenação de Planejamento, Gerência de Geoprocessamento. Documento técnico atualizado conforme Portaria MT nº 197/2006. Citado na página 47.
- [33] Guilherme Biz, Daniel Lima Miquelluti, Eli Carlos de Oliveira, and Vitor Augusto Ozaki. Aplicação do modelo starima para dados de temperaturas mínimas mensais. *Revista Brasileira de Climatologia*, 27:463, jul/dec 2020. Eletrônica. Citado na página 58.
- [34] Kaiene Domingues Paz, Marco Aurélio Tavares Marques Da Silva, and Michelle Mieko Taniguchi Trintinalia. Aplicação de starima em tráfego de curto prazo: Um estudo de caso na cidade de são paulo. Trabalho de formatura, Escola Politécnica da Universidade de São Paulo, São Paulo, Brasil, 2020. Citado 2 vezes nas páginas 59 e 72.

- [35] Donald W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441, 1963. Citado na página [59](#).
- [36] Paul E. Pfeifer and Steven J. Deutsch. A three-stage iterative procedure for space-time modeling. *Technometrics*, 22(1):35–47, 1980. Citado na página [60](#).
- [37] Kaiene Domingues Paz, Marco Aurélio Tavares, Marques Da Silva, and Michelle Mieko Taniguchi Trintinalia. Aplicação de starima em previsão de tráfego de curto prazo: Um estudo de caso na cidade de são paulo (desenvolvimento tf2), 2022. Acessado em: 2025-05-05. Citado na página [70](#).