



## Intersectando Geoestatística com Modelagem da Demanda por Transportes: um Levantamento Bibliográfico

### *Intersecting Geostatistics with Transport Demand Modeling: a Bibliographic Survey*

Samuel de França Marques<sup>1</sup> e Cira Souza Pitombo<sup>2</sup>

<sup>1</sup> Universidade de São Paulo, Departamento de Engenharia de Transportes, São Carlos-SP, Brasil. [samuelfranca@usp.br](mailto:samuelfranca@usp.br).  
ORCID: <https://orcid.org/0000-0001-5602-3277>

<sup>2</sup> Universidade de São Paulo, Departamento de Engenharia de Transportes, São Carlos-SP, Brasil. [cirapitombo@gmail.com](mailto:cirapitombo@gmail.com).  
ORCID: <https://orcid.org/0000-0001-9864-3175>

Recebido: 08.2020 | Aceito: 11.2020

**Resumo:** O planejamento de transportes depende da modelagem de variáveis que, em função de usualmente exigirem recursos elevados para a sua coleta, dispõem de uma amostragem limitada. Entretanto, uma vez que apresentam dependência espacial, a utilização da Geoestatística na modelagem da demanda por transportes se mostrou bastante conveniente, já que esse ferramental permite a obtenção de estimativas em locais não amostrados. Nesse contexto, a linha de pesquisa voltada a aplicações da Geoestatística na previsão da demanda por transportes se dá no âmbito de três das quatro etapas do modelo sequencial de planejamento (geração de viagens, escolha modal e alocação de fluxo), abrangendo trabalhos que podem ser divididos conforme o suporte, ou escala geográfica adotada, e tipo de modelo utilizado. Dessa forma, no intuito de estabelecer o estado da arte dessa linha de pesquisa, o presente trabalho propôs o levantamento e discussão de estudos no âmbito de zonas de tráfego, áreas regulares, segmentos viários, estações de metrô, pontos de parada, trechos de linha de ônibus e domicílios/indivíduos, os quais se valeram dos interpoladores geoestatísticos Krigagem Simples, Ordinária, Indicativa, Universal e Espaço-temporal, além de Simulação Sequencial Gaussiana. A análise detalhada dos trabalhos permitiu a identificação de lacunas de pesquisa nas etapas de validação dos modelos, comparação com outras abordagens espaciais e não espaciais, utilização de distâncias em rede, aplicações da Krigagem Universal (KU) a variáveis de escolha modal e seleção de preditores para a KU. Atenção especial deverá ser dada à Simulação Sequencial Gaussiana e à Krigagem Espaço-temporal, modelos que devem ditar a evolução da linha de pesquisa nos próximos anos.

**Palavras-chave:** Demanda por transportes. Geoestatística. Mobilidade urbana. Krigagem. Suporte geográfico.

**Abstract:** Transport planning depends on modeling variables, and because their collection usually requires high resources, they have limited sampling. However, since they are spatially dependent, the use of Geostatistics in transport demand modeling has proved to be especially convenient, as this tool allows obtaining estimates in non-sampled locations. In this context, the research line related to applying Geostatistics to travel demand forecasting takes place within the scope of three of the four steps of the traditional planning model (trip generation, modal choice and traffic assignment), covering studies that can be divided according to the support, or geographic scale adopted, and type of model used. Thus, in order to establish the state-of-the-art of this research line, the present study proposed surveying and discussing articles within the scope of traffic zones, regular areas, road segments, metro stations, bus stops, bus line segments and household/individual analysis, which used Simple, Ordinary, Indicator, Universal and Spatio-temporal Kriging geostatistical interpolation, in addition to Gaussian Sequential Simulation. The detailed analysis of the studies allowed identifying research gaps in the models' validation stage, comparison with other spatial and non-spatial approaches, use of network distances, applying Universal Kriging (UK) to modal choice variables and the selection of predictors for UK. Special attention should be given to Sequential Gaussian Simulation and Spatio-temporal Kriging, models that could dictate the evolution of the research line in the coming years.

**Keywords:** Transport demand. Geostatistics. Urban mobility. Kriging. Geographic support.

## 1 INTRODUÇÃO

O problema da escassez de dados em modelagem, de uma forma geral, é e sempre foi um assunto recorrente em pesquisas científicas nas mais diversas áreas de estudo. No que se refere a variáveis espaciais,

diversos artifícios foram sendo desenvolvidos, ao longo dos anos, no intuito de contornar essa limitação e permitir o conhecimento do comportamento da variável de interesse em todo o espaço em que ela atua.

Porém, os primeiros métodos para preenchimento de dados consistem em funções determinísticas, condição que contraria a real natureza das variáveis espaciais, que são espacialmente dependentes, impossibilitando a realização de inferência estatística tradicional nas estimativas provenientes deles. Nesse contexto, em meados dos anos 50 e 60, foi criada a Teoria das Variáveis Regionalizadas (MATHERON, 1963; 1971), na qual baseia-se a Geoestatística, ferramenta que abrange um conjunto de interpoladores espaciais cujos algoritmos se baseiam em uma abordagem probabilística daquelas variáveis espaciais que apresentam um padrão de variação mais ou menos definido no espaço. Entre as vantagens da Geoestatística, destaca-se o fato de ser uma ferramenta que consegue utilizar o máximo de informação disponível sobre a variável de interesse, normalmente de difícil coleta/aquisição, ou seja, um dado escasso, para a estimativa de seus valores em locais não amostrados, a partir da geração de uma superfície contínua de pontos estimados.

O surgimento da Geoestatística remonta ao trabalho de Krige (1951), voltado à resolução de problemas na Engenharia de Minas, e foi pensada, inicialmente, para modelar variáveis espaciais contínuas, isto é, que podem assumir um valor em qualquer ponto do espaço em que elas atuem como, por exemplo, teor de ouro em uma mina terrestre, temperatura, precipitação, concentração de CO<sub>2</sub> na atmosfera, entre outras. Contudo, a partir da observação de que algumas variáveis espacialmente discretas, às quais atribuem-se valores apenas em determinados pontos do espaço, também apresentavam uma estrutura espacial mais ou menos definida, a Geoestatística foi sendo expandida para diversas Áreas de estudo.

Uma das Áreas que merecem destaque nesse âmbito se refere à Epidemiologia, na qual a Geoestatística desempenha um importante papel na estimativa de riscos de doenças, detecção de áreas que apresentam riscos substancialmente maiores e análise da influência de fatores de risco (GOOVAERTS, 2009a). Dessa forma, Goovaerts (2009a) aplicou a Geoestatística ao risco de mortalidade do câncer de colo de útero e na avaliação da incidência do câncer de mama (GOOVAERTS, 2009b). Por outro lado, Stelzenmüller, Ehrich e Zauke (2005) aplicaram, com sucesso, a Krigagem Ordinária e Krigagem Universal à biomassa de duas espécies de peixes da Enseada alemã. Na agricultura, por sua vez, Kerry et al. (2016) destacaram o melhor desempenho da Krigagem na classificação de cultivares de *cranberries* em comparação com outro método baseado em imagens de satélite, e reforçaram a redução de custos na contagem de frutas. Em ciências florestais, Carvalho et al. (2015) também aplicaram duas técnicas geoestatísticas na estimativa da circunferência de árvores. Em todos esses trabalhos, as variáveis de interesse não são espacialmente contínuas, uma vez que se referem a indivíduos nos dois primeiros exemplos, e a peixes, frutas e árvores, nos antepenúltimo, penúltimo e último casos, respectivamente.

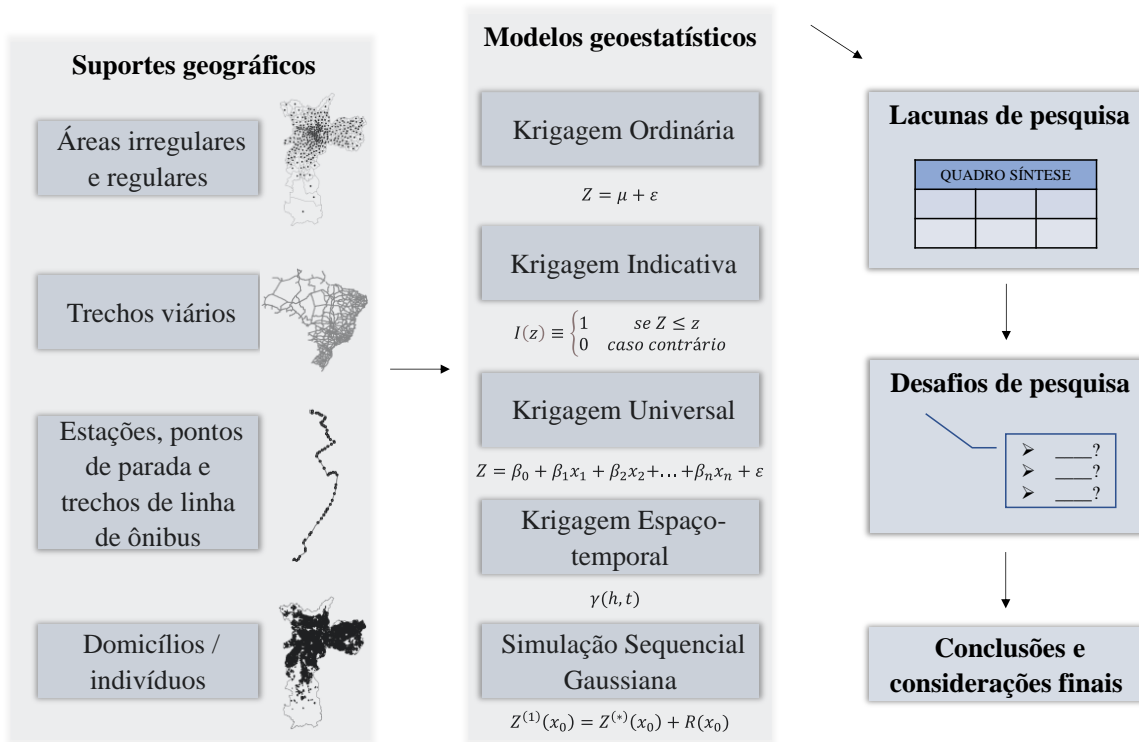
A conveniência de se aplicar a Geoestatística a variáveis espacialmente discretas também foi aproveitada pela Engenharia de Transportes: o emprego desse ferramenta já foi observado na modelagem de acidentes / segurança viária (GOMES et al., 2018; GUNDOGDU; FACULTY, 2014; MAJUMDAR; NOLAND; OCHIENG, 2004; MANEPALLI; BHAM, 2011) e, mais recentemente, na estimativa de variáveis de demanda por transportes (KLATKO et al., 2017; LINDNER; PITOMBO, 2017; SELBY; KOCKELMAN, 2013; ZHANG; WANG, 2014). Nesse contexto, os estudos desenvolvidos se concentram na estimativa de viagens originadas e/ou destinadas a zonas de tráfego, escolha do modo de transporte (associada à coordenada geográfica do domicílio), carregamento de passageiros em trechos da rede de transporte público, quantidade de passageiros embarcando em pontos de parada etc. Dentro desse escopo, as aplicações da Geoestatística a problemas relativos à mobilidade se diferenciam pela unidade de agregação utilizada nos trabalhos: zonas de tráfego, quadrículas de malhas regulares, trechos viários homogêneos, estações de metrô, pontos de parada, segmentos de linhas de ônibus e domicílios.

Enquanto considera-se que todas as abordagens representam, cada uma à sua maneira, uma contribuição para o fortalecimento dessa linha de pesquisa, é possível notar uma evolução, nem sempre cronológica, dos tratamentos e modelos utilizados. Apesar de ainda estar em consolidação, desde os primeiros estudos realizados, nota-se um aumento vertiginoso de trabalhos inseridos nessa temática. O atual estágio de pleno desenvolvimento dessa linha de pesquisa se mostra como um momento oportuno para o levantamento e esquematização de publicações, bem como a pontuação das lacunas existentes na Área, com o propósito de nortear o avanço no estado da arte e fortalecimento da linha de pesquisa. Dessa forma, o objetivo principal do

presente manuscrito reside em tratar do problema de escolha da melhor escala para variáveis de demanda por transportes, bem como o uso de simulação e modelagem geoestatística uni e multivariada.

Esse artigo está dividido da seguinte forma: além dessa breve introdução, a Seção 2 descreve a classificação das abordagens da Geoestatística à demanda por transportes de acordo com o elemento básico de agregação utilizado, também conhecido como suporte. A Seção 3 dá ênfase à evolução do tipo de modelo utilizado nos trabalhos. Por sua vez, a Seção 4 sintetiza os comentários realizados nas seções anteriores, depreendendo, a partir deles, as lacunas ainda existentes nessa linha pesquisa e que poderão servir como base para futuras pesquisas. Com base nos comentários da Seção 4, a Seção 5 lista algumas perguntas desafiantes que devem ser respondidas pelas pesquisas futuras. A Seção 6 encerra o trabalho, resumindo as principais conclusões alcançadas. A Figura 1 ilustra a referida estrutura do artigo.

Figura 1 – Estrutura do artigo.



Fonte: Os autores (2020).

## 2 O SUPORTE GEOESTATÍSTICO NAS ABORDAGENS DE DEMANDA POR TRANSPORTES

A unidade básica de agregação utilizada nas abordagens de Geoestatística é conhecida como suporte. Tendo em vista que o objetivo principal da Geoestatística é, a partir de dados amostrais disponíveis acerca de determinada variável, interpolar os valores dessa variável para todo o campo de interesse, o suporte geoestatístico pode ser definido como o elemento espacial de onde a informação amostral é coletada e onde o valor da variável de interesse será estimado (JOURNAL; HUIJBREGTS, 1978).

A escolha do suporte se dá, geralmente, em função da escala geográfica para a qual há dados amostrais disponíveis, em função da variável de interesse, do tipo de estudo, de métodos já consolidados na literatura, entre outros fatores. No contexto das variáveis de mobilidade, os seguintes suportes já foram observados: áreas (irregulares e regulares), trechos viários, estações, pontos de parada e segmentos de linhas de ônibus, e domicílios e indivíduos, partindo sequencialmente do suporte mais agregado para o mais desagregado. A seguir, os trabalhos encontrados que se utilizam dessas unidades são melhor detalhados.

### 2.1 Áreas

A aplicação da Geoestatística a variáveis de demanda por transportes no âmbito de áreas é

relativamente recente e restrita ao Brasil, considerando-se os trabalhos encontrados. Nesse caso, as contribuições principais se referem ao tratamento adequado das variáveis de demanda por transportes por meio da inclusão da autocorrelação espacial à modelagem, estimativa em locais não amostrados e visualização do comportamento espacial da variável de interesse em todo o espaço em que ela atua.

O suporte de áreas, que pode se referir tanto a áreas irregulares quanto regulares, compõe estudos voltados à demanda por transporte urbano. Nesse contexto, Lindner et al. (2016) utilizaram a Geoestatística para estimar o número de viagens por transporte público produzidas, ou seja, com origem em 320 zonas de tráfego da cidade de São Paulo. Os dados incluídos na modelagem foram coletados a partir da Pesquisa Domiciliar de Origem e Destino de 2007, que, de acordo com as autoras, abrangeu um total de 30.000 domicílios.

As zonas de tráfego são as unidades básicas de agregação do modelo sequencial de planejamento de transportes (ORTÚZAR; WILLUMSEN, 2011) e sua divisão deriva de diversos aspectos, tais como: homogeneidade nas características socioeconômicas e do sistema de transporte, barreiras físicas, relevo, entre outros. Haja vista que a produção de viagens consiste na etapa de geração de viagens do modelo sequencial, a utilização das zonas como suporte no estudo de Lindner e Pitombo (2017) se mostrou conveniente.

Apesar disso, a Geoestatística supõe que o suporte utilizado deve ser regular, tanto em tamanho quanto em forma (GOOVAERTS, 2008). Uma vez que a concepção das zonas de tráfego não se baseia nesse critério, as áreas resultantes exibem tamanhos e formas variados. Além disso, a Geoestatística também assume que os elementos de agregação devem ser consideravelmente pequenos em comparação à malha a ser interpolada (JOURNAL; HUIJBREGTS, 1978), condição facilmente violada em se tratando das zonas de tráfego.

A fim de contornar esses problemas, os trabalhos que abordavam a demanda por transportes no âmbito de zonas propuseram métodos a fim de mover a informação das zonas de tráfego para uma malha de quadrículas regulares (LINDNER; PITOMBO, 2019; ROCHA; PITOMBO; SALGUEIRO, 2016; ROCHA, LINDNER; PITOMBO, 2017). Nos três casos, assumiu-se que a variável de interesse, produção de viagens por transporte público, estava relacionada a uma característica socioeconômica da zona de tráfego, sendo que o primeiro trabalho utilizou a região metropolitana de São Paulo como estudo de caso, e os dois últimos, a região metropolitana de Salvador. Devido à impossibilidade de se realizar a modelagem no âmbito de domicílio, as autoras se valeram de diferentes procedimentos para desagregar a informação das zonas em quadrículas regulares e realizar o tratamento geoestatístico com base na malha mais refinada.

Embora a solução proposta pelos trabalhos anteriores represente, de fato, um aprimoramento da modelagem geoestatística aplicada a variáveis de demanda por transportes, no âmbito de áreas, cabe lembrar que, para gerar a superfície de valores estimados, a Geoestatística faz uso apenas de informações disponíveis na forma de pontos. No que tange às áreas, considera-se que os dados relativos às zonas de tráfego ou quadrículas regulares estão concentradas no centroide dessas unidades. Dessa forma, à medida em que as zonas são desagregadas em unidades de área menores e regulares, a informação que a Geoestatística usa para realizar a modelagem vai se tornando cada vez mais próxima da representação pontual.

Porém, uma vez que os dados ainda provêm de áreas, a simplificação adotada ao se utilizar o centroide como ponto de referência traz consigo o problema da perda da variação intrazonal, já que o centroide carrega o valor médio da variável para toda a zona, ou seja, supõe que todos os pontos contidos dentro da zona se comportam de forma semelhante. Essa condição desfavorável limita a análise dos resultados e pode levar à falácia ecológica. Além disso, embora os variogramas provenientes de áreas associadas a centroides tenham uma boa estrutura e menor efeito pepita, quando comparados aos variogramas provenientes de unidades mais desagregadas, há maior heterogeneidade e variância associada aos valores das variáveis. Nesse sentido, outras unidades de agregação, um pouco menos agregadas, puderam ser encontradas na literatura. A demanda por transportes, no âmbito de trechos viários, está comentada a seguir.

## 2.2 Trechos viários

Compreendendo a maior parte dos trabalhos encontrados na literatura dentro do escopo desse artigo, a aplicação da Geoestatística a trechos viários aborda, em sua maioria, a variável de Volume Médio Diário Anual (VMDA) de veículos em rodovias rurais. Essa variável é de fundamental importância para o dimensionamento

de pavimentos, elaboração, acompanhamento e atualização de projetos, análise de nível de serviço, implantação de pedágios rodoviários, etc., porém, ela está disponível apenas para os trechos que possuem postos de contagem de tráfego e segmentos adjacentes que não sofrem a influência de acessos viários. Nesse contexto, a Geoestatística tem a função primordial de, a partir do VMDA coletado nos postos de contagem, estimar essa variável para os trechos que não dispõem dela, o que se torna ainda mais relevante em se tratando de rodovias rurais, dada a impossibilidade de instalar e manter postos de contagem em toda a extensão da rede rodoviária.

Nesse âmbito, Klatko et al. (2017) aplicaram quatro diferentes métodos de interpolação espacial, incluindo um da Geoestatística, a fim de estimar o VMDA e, conseqüentemente, o total de milhas percorridas por veículos (VMT, *Vehicle Miles Traveled*) para vias locais sem dados de contagem de tráfego. Por sua vez, Shamo, Asa e Membah (2015) propuseram a aplicação da Geoestatística aos dados de VMDA referentes aos anos de 2008, 2009 e 2010 no estado de Washington.

Selby e Kockelman (2013) utilizaram a Geoestatística na predição/estimativa do volume de tráfego para vias do estado do Texas, nos Estados Unidos. Por sua vez, Eom et al. (2006), que afirmam ser os primeiros autores a publicarem um trabalho dentro desse escopo, se valeram da Geoestatística para a estimativa/predição do VMDA de 5 tipos de vias de Wake County, EUA. Wang e Kockelman (2009) também fizeram uso do ferramental geoestatístico, mas, nesse caso, a ênfase foi dada ao Volume Médio Diário Anual de duas diferentes classes de rodovias do estado do Texas, EUA.

Abordagens indiretas da Geoestatística à demanda por transportes, na qual o interpolador geoestatístico é aplicado a uma variável que é diretamente proporcional ao volume de tráfego, também foram observadas na literatura. Nesse contexto, a fim de avaliar o impacto do volume de tráfego e sua composição na velocidade individual dos veículos e na demanda de carros de passeio (PCU, *Passenger Car Unit*), Biswas et al. (2017) utilizaram a Geoestatística para modelar/estimar a velocidade de veículos por categoria em função do volume de veículos na mesma categoria e nas demais categorias. Posteriormente, a velocidade estimada, juntamente com a área retangular projetada de cada tipo de veículo, foi aplicada no cálculo do PCU. Outra aplicação semelhante pode ser encontrada em Zou et al. (2012).

Além disso, Chi e Zheng (2013) interpolaram a variável *transport footprint* (pegada de carbono devida ao transporte) associada a cada trecho do sistema viário de Houghton County, Michigan, EUA. Os autores colocaram a variável de interesse em função do VMDA e realizaram a modelagem geoestatística a partir de dados de postos de contagem tanto de Houghton quanto de três municípios adjacentes. Outras abordagens baseadas na Geoestatística para o preenchimento de dados de tráfego podem ser encontradas em Offor, Vaci e Mihaylova (2019), Wang e Mao (2019) e Yang et al. (2018).

Retomando a discussão iniciada na subseção 2.1, é importante comentar que, no caso do VMDA de trechos viários, em contraposição às viagens por transporte público produzidas por zona de tráfego, a unidade de agregação é substancialmente menor que a malha a ser interpolada e apresenta relativa regularidade em seu tamanho e forma. Além disso, tendo em vista que o volume de veículos é único ao longo de determinado segmento, os trechos viários não sofrem com a perda da variação intrazonal observada nas zonas e quadrículas regulares. Porém, nesse caso, permanece a simplificação de se atribuir o valor da variável de interesse a um ponto na extensão do trecho, já que se trata de uma condição necessária para a realização do tratamento geoestatístico.

Apesar das vantagens relativas aos trechos viários, em comparação com as unidades de área, algumas limitações, observadas e/ou comentadas nos trabalhos supracitados, podem ser citadas. A precisão das estimativas empreendidas pela Geoestatística depende fortemente da quantidade de informação disponível e arranjo espacial dos pontos amostrais. No caso dos segmentos de vias, os postos de contagem de tráfego, que é de onde coleta-se a informação utilizada na modelagem geoestatística, estão distribuídos de forma visivelmente irregular no interior do espaço a ser interpolado. Dessa forma, no intuito de melhorar a precisão das estimativas, faz-se necessária, frequentemente, a utilização de dados de unidades geográficas vizinhas às de interesse. Além disso, os erros resultantes apresentam uma variação substancial, sendo que os menores erros são atribuídos aos trechos situados em regiões onde há um alta densidade de postos de contagem, e os maiores aos segmentos localizados em áreas com escassez de postos. Tendo em vista que a instalação desses equipamentos não segue, necessariamente, algum critério de disposição espacial regular que beneficiaria a

amostragem sistemática, a abordagem geoestatística de trechos viários sofre com tais restrições, as quais são superadas pelos suportes detalhados a seguir.

### **2.3 Estações, pontos de parada e trechos de linhas de ônibus**

A abordagem geoestatística da demanda por estações de metrô, pontos de parada e trechos de linhas de ônibus é bastante recente e, por esse motivo, poucos trabalhos puderam ser encontrados dentro desse escopo. Nesse caso, as variáveis de interesse (embarques, desembarques e carregamento), de uma forma geral, compõem, também, a etapa de geração de viagens do modelo sequencial, uma vez que se tratam das viagens produzidas e atraídas por transporte público no âmbito de estações ou pontos de parada. A variável de carregamento, apesar de ser semelhante ao VMDA, não foi abordada na subseção anterior em razão de os segmentos de linha serem menores que os trechos viários e em função do carregamento ser derivado do volume de embarques e desembarques, informações provenientes da unidade de agregação ponto de parada.

No que tange às estações, apenas o trabalho de Zhang e Wang (2014) foi encontrado até o presente momento. Nesse estudo, os autores propuseram a aplicação da Geoestatística ao volume de embarques em estações de metrô da cidade de New York, EUA. Tendo em vista que o número de passageiros embarcando em estações de metrô pode ser facilmente coletado a partir do sistema de bilhetagem eletrônica, o objetivo desse trabalho não era interpolar a variável de demanda, mas sim extrapolar o volume de embarques para estações de metrô de uma nova linha, que seria implantada na vizinhança da primeira.

Em se tratando dos pontos de parada e trechos de linha, as variáveis de interesse são o volume de embarques e de desembarques por ponto de parada, e o volume de passageiros dentro do ônibus no segmento de linha compreendido entre dois pontos de parada consecutivos, também conhecido como carregamento. Nesse âmbito, Marques e Pitombo (2019) e Marques (2019) aplicaram a Geoestatística a essas três variáveis de demanda ao longo de uma linha de ônibus da cidade de São Paulo. Tal abordagem se justifica pelos altos recursos financeiros requeridos para a realização da pesquisa sobre/desce, levantamento que dá suporte à coleta das variáveis de interesse e, conseqüentemente, ao planejamento da rede de transporte público por ônibus. Dessa forma, a intenção dos autores era mostrar que a Geoestatística poderia servir para facultar, às cidades, a obrigatoriedade de realização da pesquisa sobre/desce em toda a rede de ônibus, já que, apenas com uma porcentagem de pontos amostrados, o ferramental geoestatístico seria capaz de interpolar a demanda para os pontos de parada e segmentos de linha não amostrados.

O tratamento geoestatístico da demanda por transporte público no âmbito de estações, pontos de parada e segmentos de linha de ônibus apresenta as seguintes vantagens: 1) as variáveis de interesse se referem a pontos propriamente ditos, ao contrário das abordagens anteriores (somente a variável de carregamento, dada por trecho de linha, se assemelha ao caso dos segmentos viários). 2) as variáveis de interesse estão disponíveis em um nível de agregação favorável à aplicação da Geoestatística. Na abordagem por áreas, ocorre um elevado grau de agregação, o que prejudica a discriminação das variações que ocorrem no interior do suporte utilizado e pode levar à falácia ecológica. No caso dos domicílios/indivíduos, conforme será visto adiante, os quais também se tratam de pontos propriamente ditos, surge o problema da alta aleatoriedade na vizinhança próxima.

### **2.4 Domicílios / indivíduos**

O tratamento geoestatístico mais detalhado pode ser atribuído aos suportes de domicílio ou indivíduo. Enquanto as abordagens anteriores se concentravam nas etapas de geração de viagens e alocação de fluxo em rede, a utilização da Geoestatística no âmbito individual ou de domicílio diz respeito à escolha modal, com apenas uma exceção entre os trabalhos encontrados.

No caso da escolha do modo de transporte, a Geoestatística exerce uma importância altamente significativa. Os dados coletados para a realização da modelagem são provenientes de pesquisas domiciliares ou questionários, cuja cobertura muito dificilmente abrange toda a população envolvida no estudo. As visitas necessárias para a aplicação de questionários são consideravelmente custosas, o que permite considerar tais dados como escassos. Conhecer o modo de transporte, por outro lado, de toda a população e não apenas dos pontos amostrados, é de fundamental relevância para o desenvolvimento de projetos urbanos, implantação de

novos modos de transporte, alteração das características dos modos já existentes, conhecimento dos fatores intervenientes na escolha modal etc.

Nesse âmbito, Pitombo et al. (2015), com base em uma pesquisa domiciliar, aplicaram a Geoestatística à probabilidade de escolha do carro como modo de viagem, de escolha de um modo não motorizado e de escolha do transporte público na cidade de São Carlos, SP, Brasil. Tendo em vista a não visualização de uma estrutura espacial aparente nessas três variáveis considerando todos os domicílios entrevistados, os autores tiveram que propor a divisão da cidade em seis regiões homogêneas de acordo com a renda e utilizar apenas duas delas na modelagem geoestatística.

Utilizando 110 pontos amostrados da cidade de São Carlos, Gomes et al. (2016a, 2016b) interpolaram, a partir da Geoestatística, a probabilidade de escolha do modo individual motorizado (automóvel ou motocicleta) por domicílio. Nesse caso, a desagregação do banco de dados em seis regiões homogêneas, conforme renda, também foi inicialmente adotada. Porém, as autoras identificaram a existência de um padrão espacial mais ou menos definido na variável de interesse apenas em uma das regiões, a qual foi a única utilizada no estudo de caso.

Ainda relativo ao Brasil, o trabalho de Lindner e Pitombo (2017) pode ser encontrado. Utilizando a pesquisa domiciliar de origem e destino de 2007 da região metropolitana de São Paulo, as autoras geraram mapas da probabilidade de escolha do modo de transporte público em detrimento do individual motorizado. Apesar de ter apresentado erros maiores que a regressão logística tradicionalmente utilizada no modelo sequencial (ORTÚZAR; WILLUMSEN, 2011), as autoras ressaltaram que, em se tratando da Geoestatística, observa-se a importante contribuição da superfície contínua de valores interpolados, isto é, o conhecimento da escolha do modo de transporte mesmo em domicílios não amostrados, considerando a dependência espacial presente na escolha do modo de transporte.

Enquanto nos trabalhos anteriores, a escolha modal foi abordada, de uma forma geral, sem se preocupar com qual era o destino das viagens, Chica-Olmo, Rodríguez-López e Chillón (2018) propuseram a aplicação da Geoestatística à seleção do modo de transporte com destino ou origem em quatro escolas de ensino médio da cidade de Granada, na Espanha. Nesse artigo, o objetivo dos autores era encontrar o tipo de distância que melhor explicava a utilização do modo a pé para viagens às escolas e, reconhecendo que, não somente a distância influenciava essa decisão, mas também outras variáveis difíceis de se mensurar, eles modelaram a probabilidade de escolha pelo modo a pé, em detrimento do motorizado, a partir da Geoestatística adaptada para variáveis de interesse binárias. Dessa forma, além de se conhecer a probabilidade de escolha do modo a pé em domicílios não entrevistados, a superfície contínua de valores estimados permitiu a delimitação da área de influência das escolas, ou seja, as regiões no plano cuja probabilidade de utilização da caminhada para o trajeto casa-escola e/ou vice-versa são mais próximas de 1 (maiores que 50%).

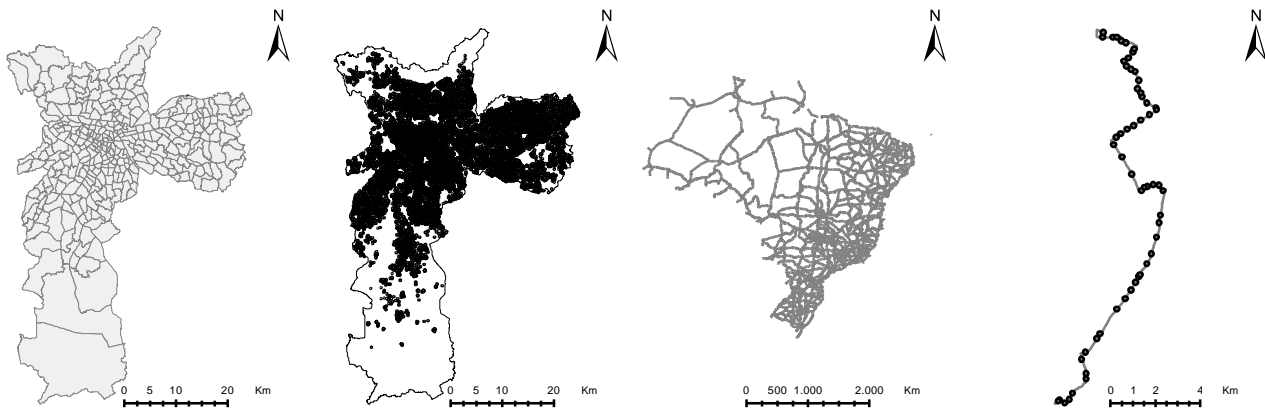
Apesar das contribuições incontestáveis das abordagens geoestatísticas no âmbito de domicílios, algumas limitações podem ser observadas nos trabalhos: ao utilizar informações com um elevado nível de desagregação, a ocorrência de um padrão espacial aparente na variável de interesse se torna difícil. Mesmo que haja alguma dependência espacial na escolha do modo de transporte, o âmbito do domicílio acaba demonstrando uma alta aleatoriedade na vizinhança imediata, uma vez que o comportamento humano, reconhecidamente complexo, está envolvido. Tendo em vista que a Geoestatística assume dependência espacial das variáveis de interesse, ou seja, que pontos próximos entre si no espaço apresentam um comportamento semelhante, os resultados da modelagem por domicílio acabam demonstrando erros elevados. Além disso, o espaço a ser interpolado pode abranger regiões cujos resultados da superfície contínua não fazem sentido, tais como áreas verdes. Operacionalmente, este tipo de análise gera variogramas experimentais e teóricos bem estruturados, porém com alto efeito pepita.

Uma solução para essas restrições foi proposta por Rocha (2019). Com base na pesquisa de origem e destino de 2007 da região metropolitana de São Paulo, a autora apresentou um método de agregação da informação domiciliar (porcentagem de viagens por automóvel) em quadrículas regulares. Nesse trabalho, que incluiu apenas o centro da cidade como estudo de caso, novamente devido à ausência de estrutura espacial aparente considerando toda a cidade, Rocha (2019) apresentou um modelo de otimização que encontrasse a grade regular (ou suporte) mais apropriada para o tratamento geoestatístico empreendido. Em outras palavras, a grade ótima seria nem tão agregada quanto as zonas de tráfego, nem tão desagregada quanto os domicílios,

sendo que o mapa resultante da interpolação deveria reproduzir um padrão semelhante ao do nível mais detalhado, porém com erros menores. A agregação da demanda por atividades, levantada a partir de uma pesquisa domiciliar, em células regulares, também pode ser encontrada em Yoon, Ravulaparthi e Goulias (2014), com estudo de caso voltado a uma cidade da Pensilvânia, EUA.

Em síntese, a Seção 2 dedicou-se à discussão sobre as abordagens geoestatísticas da demanda por transportes em diferentes escalas geográficas. Tais suportes geoestatísticos estão ilustrados, em sequência, na Figura 2.

Figura 2 – Da esquerda para a direita: zonas de tráfego de São Paulo, com respectivos centroides; domicílios amostrados na pesquisa de origem e destino de 2017 em São Paulo; trechos viários do Sistema Nacional de Viação; e linha 856R-10-1 do município de São Paulo, com respectivos pontos de parada e segmentos de linha.



Fonte: Os autores (2020).

Tendo em vista as limitações impostas pelos pressupostos da Geoestatística e pelo próprio nível de agregação das variáveis, foi observada uma melhor adequação do tratamento geoestatístico no âmbito de pontos de parada e trechos de linha de ônibus, bem como estações e segmentos viários. Tais suportes encontram-se em um nível de agregação considerado como intermediário entre o de zonas de tráfego, mais agregado, e o de domicílios, mais desagregado, sendo que, tanto no caso das zonas quanto dos domicílios, a utilização de grades regulares se mostra como uma solução possível. A seção a seguir busca classificar os trabalhos conforme o interpolador geoestatístico utilizado, comentando, conforme realizado na presente seção, suas vantagens e limitações.

### 3 MODELOS GEOESTATÍSTICOS UTILIZADOS NAS ABORDAGENS

Os interpoladores tradicionais da Geoestatística receberam o nome de Krigagem, em homenagem a Krige (1951), engenheiro de minas precursor desse ferramental. Conforme mostra a Eq. (1), tais estimadores compreendem, basicamente, uma combinação linear de valores vizinhos da variável de interesse, em pontos amostrados, associados a pesos ótimos, os quais dependem do arranjo espacial do banco de dados e dos variogramas teóricos obtidos em etapa prévia.

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (1)$$

em que  $Z^*(x_0)$  é valor estimado da variável regionalizada  $Z$  na posição geográfica  $x_0$ ; e  $\lambda_i$  é o peso ótimo atribuído pela krigagem à observação do vizinho  $i$ . Dessa forma, a Krigagem tem por objetivo principal calcular esses pesos, levando em consideração que pontos amostrais mais próximos daquele onde a estimativa será realizada devem receber um peso maior que os pontos situados mais distantes, seguindo a lógica de que a dependência espacial é mais intensa na vizinhança imediata. A evolução da modelagem geoestatística da demanda por transportes, considerando os diferentes tipos de Krigagem utilizados, é comentada a seguir.



### 3.1 Krigagem Ordinária

A Krigagem Ordinária (KO) é um dos interpoladores mais simples da Geoestatística e está presente em quase todas as abordagens descritas na Seção 2. Sua simplicidade provém do fato de que se trata de uma técnica univariada, isto é, que depende somente dos valores da variável de interesse em pontos do espaço e respectivas coordenadas geográficas. A KO ainda assume que a variável de interesse apresenta uma média local constante ( $\mu$ ) e que os valores dessa variável oscilam em torno dessa média em função da distância entre os pontos do banco de dados (CRESSIE, 1993; MATHERON, 1971), de acordo com a Eq. (2).

$$Z = \mu + \varepsilon \quad (2)$$

em que  $\varepsilon$  são os resíduos do modelo. No intuito de estimar o Volume Médio Diário Anual (VMDA) e, conseqüentemente, o total de milhas percorridas por veículos para vias locais sem dados de contagem de tráfego, Klatko et al. (2017) aplicaram a KO ao VMDA dividindo-se as vias em categorias de acordo com suas características (urbana ou rural e de tráfego intenso ou reduzido). Nesse caso, 90% do banco foi utilizado na etapa de modelagem, enquanto os 10% restantes foram reservados para validação. Os resultados da KO foram comparados aos de três outros interpoladores e, apesar de nenhuma das técnicas ter demonstrado um desempenho predominantemente superior ao das demais na validação, os autores afirmam que a interpolação espacial, fundamental do qual a KO faz parte, pode levar a estimativas mais confiáveis e precisas do que outros métodos existentes.

Marques e Pitombo (2019) aplicaram a KO às variáveis de demanda por transporte público no âmbito de pontos de parada (embarques e desembarques) e de trechos de linha (carregamento). No intuito de mostrar todo o potencial da KO em estimar tais variáveis, os autores preferiram calcular as métricas de aderência com base na validação cruzada ao invés da validação propriamente dita. Em Geoestatística, a validação cruzada é realizada por meio do teste do ponto fictício (CRESSIE, 1993), no qual todo o banco de dados é utilizado na etapa de modelagem. A KO apresentou um desempenho satisfatório, nesse caso, principalmente no que diz respeito à variável de carregamento, que, por demonstrar uma variação espacial suave ao longo da linha de ônibus, foi a que mais se destacou em comparação às variáveis embarques e desembarques. Os autores afirmaram que a utilização de variáveis explicativas poderia melhorar o desempenho de embarques e desembarques.

No que tange a variáveis ao longo de rede de transporte, a Geoestatística permite a expansão da abordagem tradicional, baseada na distância euclidiana (em linha reta) entre os pontos do banco de dados, para a modelagem ao longo da rede. Uma vez que a atividade de transporte ocorre ao longo de uma rede viária, autores como Wang e Kockelman (2009) afirmaram que utilizar distâncias em rede em substituição às tradicionais euclidianas poderia melhorar os resultados dos interpoladores. A partir dessa sugestão, Marques (2019) aplicaram a KO, com validação cruzada, às mesmas variáveis de Marques e Pitombo (2019), mas, dessa vez, utilizando as distâncias ao longo da linha de ônibus. Comparando, porém, os resultados dessa abordagem com os da KO tradicional, o autor notou pouca ou nenhuma melhoria significativa.

Chi e Zheng (2013) também utilizaram a KO com validação cruzada e distâncias em rede para interpolar espacialmente, a variável *transport footprint* associada a cada trecho do sistema viário de Houghton County, Michigan, EUA, colocando a variável de interesse em função do VMDA. Além de verificar que a porcentagem de erro variava de acordo com a magnitude dos valores reais, isto é, que os erros cresciam à medida que o valor da variável de interesse aumentava, os autores apontaram como limitação o fato de a Krigagem Ordinária assumir que a variável resposta possui média constante, o que pode ter afetado o desempenho da técnica. Nesse caso, a abordagem tradicional com distâncias euclidianas não foi realizada.

A KO foi utilizada, também, por Pitombo et al. (2015) para a interpolação da escolha de três diferentes modos de transporte no âmbito do domicílio para duas regiões de São Carlos, SP, Brasil. A etapa de validação, para a qual reservou-se 30% do banco de dados, apontou uma porcentagem de acerto inferior a 50% apenas em um dos seis casos analisados. Porém, tendo em vista que a Krigagem Ordinária é uma técnica apropriada para variáveis contínuas, o fato de a escolha do modo ter, usualmente, uma natureza discreta/categórica, acaba exigindo que os dados originais sejam transformados antes de serem utilizados na KO. Nesse caso, Pitombo et

al. (2015) se valeram de Árvores de Decisão (AD) para converter a variável dicotômica em contínua. Uma vez que os resultados das ADs trazem consigo alguns erros, as estimativas de Krigagem Ordinária apresentam erros acumulados, somando a incerteza das ADs à da própria KO.

Além disso, por ser univariada, não é possível utilizar a KO para previsões futuras da variável de interesse (justamente pela ausência de variáveis explicativas). Dessa forma, a superfície obtida pode ser utilizada para fins de planejamento operacional para períodos curtos ou diminuição de locais amostrados. As limitações da KO são superadas pelos interpoladores detalhados nas subseções a seguir.

### 3.2 Krigagem Indicativa

Diferentemente da Krigagem Ordinária, a Krigagem Indicativa (KI) foi desenvolvida para tratar justamente de variáveis categóricas e, mais especificamente, binárias. Conforme mostra a Eq. (3), a variável dicotômica  $I$  varia em função de um *cut-off*  $z$ , que divide a variável regionalizada  $Z$ , em dois grupos que podem assumir apenas os valores 0 ou 1 (CRESSIE, 1993; JOURNEL, 1983).

$$I(z) \equiv \begin{cases} 1 & \text{se } Z \leq z \\ 0 & \text{caso contrário} \end{cases} \quad (3)$$

Dessa forma, a modelagem geoestatística da escolha modal não depende de etapas anteriores e a superfície contínua de valores estimados possui apenas duas indicações: uma para os pontos onde a probabilidade de escolha de determinado modo é maior que 50% (ou outro patamar especificado); e outra para os pontos onde a probabilidade é inferior a esse *cut-off* estabelecido, que seria a região na qual prioriza(m)-se outro(s) modo(s) de transporte.

Nesse âmbito, Lindner e Pitombo (2017) compararam o desempenho da KI ao da KO na estimativa da preferência pelo transporte público, em detrimento do individual motorizado, na cidade de São Paulo para o ano de 2007, reservando 30% do banco de dados para validação. De forma semelhante à adotada por Pitombo et al. (2015), os dados utilizados na KO foram derivados do modelo de escolha discreta tradicional (*logit*, regressão logística), que utiliza a informação de variáveis socioeconômicas para explicar a escolha modal. Apesar de exibir propagação de erros, os resultados da KO foram semelhantes aos da KI, porém, nesse caso, a KI só consegue gerar mapas para a situação analisada, enquanto que a KO, associada ao modelo *logit*, é capaz de realizar estimativas da probabilidade de escolha em cenários alternativos/futuros. Por outro lado, informações adicionais são requeridas, o que não ocorre na KI.

No caso de Gomes et al. (2016a, 2016b), aos invés de utilizar o modelo *logit*, a variável de escolha discreta foi transformada em contínua a partir de Árvores de Decisão. Dessa forma, a KO foi aplicada às probabilidades de escolha do modo individual motorizado, sendo, posteriormente, comparada à KI, que utiliza os dados originais em sua forma bruta. Em contraposição aos resultados de Lindner e Pitombo (2017), a simplicidade da KI apresentou um melhor desempenho, na validação cruzada, que o método de dois estágios envolvendo ADs e Krigagem Ordinária, que sofreu prejuízos devido à propagação de erros.

A associação da KO ao modelo *logit*, conforme verificado em Lindner e Pitombo (2017), busca, basicamente, superar a limitação da KO e KI em modelar apenas a situação sob escrutínio, o que impede a sua aplicação ao planejamento de médio e longo prazo. Além disso, esses interpoladores univariados, por não dependerem de variáveis explicativas, também impossibilitam o conhecimento das relações entre as variáveis de interesse e os fatores que as afetam. Tais restrições são contornadas por meio do interpolador Krigagem Universal, a ser comentado adiante.

### 3.3 Krigagem Universal

Retomando a introdução da Seção 3.1, a Krigagem Ordinária assume que a variável de interesse possui uma média local constante e que as estimativas em pontos não amostrados dependem desse valor central e das distâncias entre tais pontos e seus vizinhos (CRESSIE, 1993; MATHERON, 1971). Essa suposição é capaz de gerar resultados mais precisos apenas quando a amplitude de variação local da variável de interesse não é elevada. À medida em que começam a ocorrer mudanças bruscas na variável de interesse, principalmente na

vizinhança próxima, o estimador de KO passa a ser ineficiente, como é o caso de Marques e Pitombo (2019) e Chi e Zheng (2013), já comentados anteriormente.

O estimador de Krigagem Universal (KU), por sua vez, flexibiliza essa suposição ao assumir que há uma variação de longo alcance na variável de interesse, que surge em função da influência de variáveis externas. Dessa forma, conforme mostra a Eq. (4), a média local constante é substituída, na maioria das vezes, por uma combinação linear de variáveis explicativas  $x_k$  e parâmetros  $\beta$  a serem estimados (CRESSIE, 1993).

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (4)$$

em que  $\varepsilon$  são os resíduos do modelo, nos quais assume-se existir uma estrutura espacial. Devido a essa flexibilidade, mas com o custo adicional das variáveis explicativas, a KU normalmente alcança resultados melhores que os da KO e também permite a previsão dos valores de interesse em outros cenários, além do analisado.

Contudo, para evitar a necessidade de coleta de dados secundários para a modelagem geoestatística, é possível utilizar, como variáveis explicativas, apenas as coordenadas geográficas dos pontos, conforme realizado por Wang e Kockelman (2009), referente a aplicações da KU ao volume de tráfego. Wang e Kockelman (2009) verificaram que o método utilizado retornava resultados razoavelmente confiáveis para volumes de tráfego intermediários, enquanto que, para baixos VMDAs e VMDAs elevados, a técnica superestimou e subestimou, respectivamente, os valores reais. Segundo eles, esses resultados poderiam ser melhorados com a aquisição de dados acerca de variáveis explicativas e utilização de distâncias em rede.

Uma das formas mais simples de Krigagem Universal, em que realmente se faz uso de uma variável explicativa externa, se refere à Krigagem com Deriva Externa (KDE). A KDE requer um modelo de deriva, deduzido a partir de uma função linear da variável secundária (CRESSIE, 1993; HENGL; HEUVELINK; ROSSITER, 2007). Assim, uma variável secundária, que deve estar altamente correlacionada com a variável de interesse, é selecionada para auxiliar na modelagem. Dessa forma, a KDE já foi utilizada por Rocha, Pitombo e Salgueiro (2016) e Rocha, Lindner e Pitombo (2017) na estimativa da produção de viagens por transporte público, com população como variável secundária. Na validação, para a qual 40% do banco de dados foi reservado, observaram-se erros próximos de zero e alta aderência entre valores reais e estimados.

Lindner et al. (2016) também aplicaram a KDE à produção de viagens por transporte público, mas, dessa vez, os autores utilizaram, como variável secundária, um dos fatores resultantes da agregação de múltiplas variáveis socioeconômicas a partir de uma Análise em Componentes Principais. Comparando os resultados dessa abordagem com os da regressão linear simples (RLS), as autoras verificaram um desempenho semelhante de ambas as técnicas, porém reforçaram que somente a KDE é capaz de gerar, com base na autocorrelação espacial dos dados, mapas com valores estimados, inclusive em pontos não amostrados. A regressão linear, por sua vez, não contabiliza a dependência espacial da variável de demanda.

Valendo-se de múltiplas variáveis explicativas, Eom et al. (2006) aplicaram a KU para a estimativa/previsão do VMDA de cinco tipos de vias de Wake County, EUA. A partir de uma amostra com apenas 200 de 1154 postos de contagem, a variância explicada apenas pelas variáveis explicativas foi de aproximadamente 65%, deixando uma contribuição de 55% à interpolação espacial. Em comparação com a regressão linear múltipla (RLM), os resultados da KU demonstraram um desempenho superior. Biswas et al. (2017) também aplicaram a KU para a modelagem indireta do volume de tráfego, em unidades de carros de passeio, a partir da velocidade dos veículos. Nesse caso, a Krigagem Universal foi validada a partir de dados coletados de uma via arterial com características semelhantes às da via cujos dados foram utilizados no treinamento do modelo. Os autores também compararam os resultados da KU com aqueles obtidos a partir de regressões baseadas no volume e na densidade de tráfego, sendo que a KU demonstrou melhor desempenho na previsão da velocidade para as cinco categorias de veículos consideradas.

Conforme sugerido por Wang e Kockelman (2009), a KU com distâncias em rede foi comparada à KU tradicional (com distâncias em linha reta) e à regressão geograficamente ponderada (GWR, *Geographically Weighted Regression*) na previsão/estimativa do VMDA para vias do estado do Texas nos Estados Unidos (SELBY; KOCKELMAN, 2013). Variáveis explanatórias como velocidade, número de faixas, acessibilidade populacional, densidade de empregos e classe funcional foram incluídas na modelagem. Baseados em métricas

de erros aplicadas na etapa de validação, os resultados mostraram um desempenho ligeiramente melhor da Krigagem Universal sobre a GWR, sendo que ambas as técnicas foram superiores a um terceiro modelo não espacial que considerava apenas a heterocedasticidade presente na variável de interesse. A alternativa de distâncias em rede, por sua vez, além de ter exigido expressivos recursos computacionais e de tempo de processamento, não demonstrou impacto positivo nos resultados, embora os autores afirmem que modelos baseados em dados de contagem densamente distribuídos no espaço poderiam ser beneficiados por tal recurso. Eles, porém, ressaltaram que a utilização de distâncias não euclidianas pode comprometer a obtenção de uma matriz de covariância positiva, necessária à aplicação da Krigagem.

Distâncias em rede também foram utilizadas por Zhang e Wang (2014) em modelagem espacial do número de embarques em estações de metrô na cidade de New York. Nesse caso, a abordagem tradicional de KU com distâncias euclidianas, juntamente com a de regressão linear clássica, foi comparada à KU com distâncias em rede. Apesar de essa última abordagem ter estimado parâmetros com estatísticas *t* consideravelmente maiores que as demais, o desempenho dos três modelos não variou significativamente. Tal resultado pode ser atribuído à alta porcentagem de variação já explicada pela tendência de longo alcance (a combinação linear de parâmetros), que acabou deixando pouca variabilidade para a estrutura espacial da variável, refletida na função semivariograma. Por sua vez, a pequena diferença no desempenho da KU com e sem distâncias em rede colocou, novamente, em questão, a real eficácia da utilização das distâncias em rede.

Por fim, Shamo et al. (2015) propuseram o ajuste de três tipos de Krigagem (Simples, Ordinária e Universal) associados a três modelos de variograma (exponencial, esférico e gaussiano) aos dados de VMDA referentes aos anos de 2008, 2009 e 2010 no estado de Washington. A Krigagem Simples está pautada em uma suposição ainda mais rigorosa que a da KO: além de ser univariado, esse interpolador considera que a variável de interesse possui média global constante e conhecida (YAMAMOTO; LANDIM, 2015). No geral, baseado em um processo de validação no qual aplicaram-se várias métricas de erros, não foi identificada uma combinação ótima para o banco de dados. Dessa forma, os autores reforçaram que a melhor técnica de Krigagem e melhor variograma só podem ser determinados a partir da estrutura presente nas informações disponíveis sobre a variável de interesse.

Apesar de a equação que descreve a tendência de longo alcance na Krigagem Universal ser, normalmente, semelhante a uma regressão linear, Chica-Olmo, Rodríguez-López e Chillón (2018) modelaram a probabilidade de escolha pelo modo a pé a partir da Krigagem Universal (chamada, pelos autores, de *regression-kriging*) adaptada para variáveis de interesse binárias, na qual, ao invés da combinação linear de variáveis explicativas, utilizou-se uma regressão logística. Dessa forma, os autores atribuíram aos resíduos a dependência espacial da probabilidade de escolha do modo, já que abrangia as variáveis explicativas locais não incluídas no modelo. Essa aplicação, que foi a única desse tipo a ser encontrada, superaria as limitações da Krigagem Indicativa, que não consegue prever a escolha do modo em cenários hipotéticos/futuros, e da Krigagem Ordinária, que depende de outras técnicas para interpolar a alternativa modal, exibindo uma propagação de erros.

Observando os trabalhos em que se aplicam a Krigagem, nota-se que as variáveis de interesse são dadas, em alguns casos, em séries históricas. A Geoestatística também pode ser utilizada em banco de dados que contenham um aspecto temporal, em conjunto com o espacial, conforme mostrado a seguir.

### 3.4 Krigagem Espaço-temporal

A Krigagem Espaço-temporal se refere a um interpolador em que a matriz de covariância leva em consideração não somente os *lags* espaciais, mas também temporais. Apesar de vários trabalhos utilizarem dados *cross-section*, isto é, que abrangem apenas um período de tempo, em alguns casos nota-se a disponibilidade de informações exaustivas de demanda ao longo do tempo. Nesse âmbito, é primordial reconhecer o caráter dinâmico das variáveis de viagens e a potencial dependência entre valores presentes, passados e/ou futuros (autocorrelação serial / temporal), situação em que a Krigagem Espaço-temporal se mostra bastante conveniente.

Dentro do escopo desse artigo, foram encontrados, até o presente momento, poucos trabalhos que aplicam a Krigagem Espaço-temporal a uma variável de demanda. No intuito de prover uma taxonomia de

ambientes urbanos com base em sua dinâmica e em aspectos sociais, Yoon, Ravulaparthi e Goulias (2014) interpolaram a intensidade de atividade em células espaciais regulares em uma pequena cidade da Pensilvânia, EUA. Baseada em um diário de atividades, a variável interpolada correspondia à quantidade de pessoas desenvolvendo determinada atividade (compras, alimentação) no espaço e no tempo, com intervalos de uma hora. Uma vez que a amostra de respondentes correspondia, aproximadamente, a apenas 1% da população do município, a utilização da Krigagem Espaço-temporal, nesse caso, foi imprescindível para o conhecimento dos padrões espaço-temporais de atividades diárias em toda a cidade considerada.

Yang et al. (2018) aplicaram a Krigagem Espaço-temporal ao preenchimento de dados de volume de tráfego com base nas leituras de seis sensores instalados consecutivamente ao longo de uma via em Nashville, Tennessee, EUA. Um banco de dados em painel foi consolidado a partir das leituras dos sensores durante 33 dias, em intervalos de 30 segundos, e diferentes cenários de análise foram propostos aumentando-se gradualmente a porcentagem de dados ausentes. A comparação com outros dois métodos de preenchimento de dados (média histórica e *k-nearest neighborhood*), utilizando-se um dos sensores como estudo de caso, revelou que a Krigagem Espaço-temporal foi o método com melhor acurácia em dez dos onze cenários considerados. Essa abordagem tem a vantagem de incluir, na modelagem, o fato de que, além de o volume de tráfego do sensor escolhido para a análise depender do volume registrado nos sensores adjacentes, há também uma dependência entre os volumes de tráfego atual, passado e/ou futuro, registrados pelo sensor estudado.

### 3.5 Simulação Sequencial Gaussiana

Assim como a Krigagem Espaço-temporal, a Simulação Sequencial Gaussiana (SSG) também é uma ferramenta da Geoestatística cuja aplicação a variáveis de demanda é bastante recente e ainda incipiente. De forma simplificada, enquanto a Krigagem envolve a análise de apenas um cenário da variável de interesse, seja atual, hipotético ou futuro, a SSG realiza a aleatorização do fenômeno regionalizado (REMY; BOUCHER; WU, 2009). Nesse âmbito, ao invés de produzir somente uma superfície de valores estimados, essa técnica gera várias simulações dos resultados de Krigagem, a partir das quais é possível mapear medidas de incerteza, obter a média das realizações, intervalos de confiança e analisar cenários críticos.

Além disso, as simulações permitem gerar uma população sintética desagregada, a partir de uma base de dados agregada, que reproduz, com a maior similaridade possível, o perfil da população verdadeira (PIANUCCI et al., 2019), o que representa um importante avanço na busca por soluções à amostragem limitada das pesquisas domiciliares. Essa característica também faz com que a SSG seja capaz de lidar com mudanças de suporte.

Nesse contexto, a única aplicação da Simulação Sequencial Gaussiana a uma variável de demanda por transportes, encontrada até o presente momento, se deve ao estudo de Lindner e Pitombo (2019). As autoras simularam os dados de produção de viagens por transporte público nas 460 zonas de tráfego da região metropolitana de São Paulo, mudando o suporte para uma malha regular. Dessa forma, foi possível plotar mapas dos valores médios das realizações, variância, valores mínimos e máximos em cada quadrícula.

Após detalhados os modelos geoestatísticos já utilizados no tratamento da demanda por transportes, bem como suas vantagens e limitações, a seção a seguir sintetiza os comentários realizados na presente seção e na Seção 2, depreendendo, a partir deles, as lacunas de pesquisa ainda existentes dentro desse escopo. Ao leitor interessado no aprofundamento matemático dos interpoladores comentados nessa seção, as seguintes publicações são indicadas: para Krigagem Ordinária, Krigagem Indicativa e Krigagem Universal (CRESSIE, 1993; JOURNAL, 1983; MATHERON, 1971); para Krigagem Espaço-temporal (GRÄLER; PEBESMA; HEUVELINK, 2016); e para Simulação Sequencial Gaussiana (DEUTSCH; JOURNAL, 1998). Vários dos artigos citados nas seções 2 e 3 também oferecem uma demonstração objetiva do formalismo estatístico imbuído no estimador aplicado.

## 4 LACUNAS DE PESQUISA

O Quadro 1 consolida os trabalhos comentados nas seções 2 e 3, resumindo algumas de suas principais características.

Quadro 1 – Síntese dos trabalhos analisados. KI, KO, KU, RLS, RLM e VMT se referem, respectivamente, à Krigagem Indicativa, Krigagem Ordinária, Krigagem Universal, Regressão Linear Simples, Regressão Linear Múltipla e Milhas Percorridas por Veículo.

Fonte	Suporte	Variável(is)	Método utilizado (software)	Vantagens (+) e Limitações (-)
Lindner et al. (2016)	Zonas de tráfego	Viagens produzidas por transporte público	KU com apenas 1 covariável e RLS ( <i>GeoMS 1.0</i> )	(+) Análise espacial multivariada; projeções futuras. (-) Variável secundária deriva de uma Análise em Componentes Principais; Problema da Unidade de Área Modificável.
Yoon, Ravulaparthi e Goulas (2014)	Quadrículas regulares	Quantidade de pessoas desenvolvendo determinada atividade	Krigagem Espaço-temporal sem validação ( <i>R</i> )	(+) Mapeamento de como os espaços urbanos atraem diferentes interações sociais em diferentes momentos do dia. (-) Não considera a interdependência de atividades entre indivíduos; não inclui covariáveis nem comparação com outros métodos.
Rocha, Pitombo e Salgueiro (2016) e Rocha, Lindner e Pitombo (2017)	Quadrículas regulares	Viagens produzidas por transporte público	KU com apenas 1 covariável e validação ( <i>GeoMS 1.0</i> )	(+) Desagregação em malha regular; análise espacial multivariada. (-) Propagação de erros em razão de a variável dependente ser uma estimativa.
Rocha (2019)	Quadrículas regulares	Porcentagem de viagens por automóvel	KO ( <i>R Studio, SGeMS e GeoMS 10.1</i> )	(+) Desagregação em malha regular ótima; (-) Tempo de execução do algoritmo; baixa dependência espacial; ocorrência de semivariâncias negativas.
Lindner e Pitombo (2019)	Quadrículas regulares	Viagens produzidas por transporte público	Simulação Sequencial Gaussiana ( <i>R Studio e SGeMS 3.0</i> )	(+) Mapeamento de cenários críticos; cálculo de parâmetros de incerteza; mudança de suporte. (-) Lida apenas com o planejamento de curto prazo e variáveis normais.
Eom et al. (2006)	Trechos viários	Volume médio diário anual	KU e RLM com validação	(+) Comparação entre diferentes métodos de calibração dos parâmetros; desempenho superior da KU. (-) Variação dos erros de acordo com a densidade de postos de contagem.
Wang e Kockelman (2009)	Trechos viários	Volume médio diário anual	KU apenas com coordenadas e validação ( <i>ArcGIS</i> )	(+) Tentativa inicial de modelar a tendência de longo prazo a partir das coordenadas geográficas dos pontos. (-) Resultados razoavelmente confiáveis apenas para volumes de tráfego intermediários; ausência de covariáveis e da utilização de distâncias em rede; potencial propagação de erros em função de a variável dependente ser uma estimativa.
Chi e Zheng (2013)	Trechos viários	Pegada de carbono relativa ao transporte / Volume médio diário anual	KO com distâncias em rede ( <i>R 2.12.2</i> )	(+) Utilização de estimador robusto do semivariograma e de distâncias em rede. (-) Variação da porcentagem de erro de acordo com a magnitude dos valores reais; quantidade reduzida de pontos para o cálculo do semivariograma; ausência de covariáveis.
Selby e Kockelman (2013)	Trechos viários	Volume médio diário anual	KU, com distâncias em rede e euclidianas, GWR e modelo não espacial, com validação ( <i>MATLAB</i> )	(+) Utilização de transformação de Box-Cox e de estimador robusto; comparação entre modelos espaciais e não espacial; melhor desempenho da KU. (-) Limitação da Krigagem quanto à heterocedasticidade da variável de interesse; variação dos erros de acordo com a densidade de postos de contagem.
Shamo, Asa e Membah (2015)	Trechos viários	Volume médio diário anual	KS, KO e KU com validação ( <i>ArcGIS</i> )	(+) Comparação de combinações entre interpoladores e modelos teóricos de semivariograma. (-) Não modelagem da tendência de longo prazo a partir de covariáveis.
Klatko et al. (2017)	Trechos viários	VMT / Volume médio diário anual	KO com validação ( <i>ArcGIS</i> )	(+) Não requer informações adicionais; divisão das vias em classes. (-) Exige um inventário abrangente de vias locais.
Biswas et al. (2017)	Trechos viários	Velocidade dos veículos / Unidades de carros de passeio	KU e RLM com validação	(+) Transferabilidade dos resultados para outros países, seja pelos gráficos elaborados, seja por meio de replicação do experimento a novos dados; KU superior ao modelo não espacial; divisão dos veículos por categoria. (-) Apenas o volume de tráfego de cinco categorias de veículos é considerado como covariável.

(Continua)

(Conclusão)

Fonte	Suporte	Variável(is)	Método utilizado (software)	Vantagens (+) e Limitações (-)
Yang et al. (2018)	Trechos viários	Volume de tráfego	Krigagem Espaço-temporal com validação ( <i>R Studio</i> )	(+) Validação com aumento gradual da porcentagem de dados faltantes (análise de sensibilidade); método mais preciso e flexível que o de média histórica e <i>k-nearest neighborhood</i> ; o método não assume nenhuma distribuição de probabilidade para os dados. (-) Ausência de comparação com outros interpoladores geoestatísticos e modelos não espaciais.
Zhang e Wang (2014)	Estações de metrô	Número de embarques	KU, com distâncias euclidianas e em rede, e RLM sem validação ( <i>MATLAB</i> )	(+) Utilização de covariáveis e distâncias em rede; KU com distâncias em rede superior à KU tradicional e modelo não espacial. (-) Covariáveis explicam parte significativa da variação de embarques e deixam pouco para a função semivariograma.
Marques e Pitombo (2019)	Pontos de parada e trechos de linha de ônibus	Embarques, desembarques e carregamento	KO ( <i>ArcGIS 10.1</i> )	(+) Nível de agregação favorável; redução de custos com pesquisa sobe/desce. (-) Ausência de covariáveis, de validação e de comparação com outros métodos.
Marques (2019)	Pontos de parada e trechos de linha de ônibus	Embarques, desembarques e carregamento	KO com distâncias euclidianas e em rede ( <i>R Studio</i> )	(+) Nível de agregação favorável; utilização de distâncias em rede; redução de custos com pesquisa sobe/desce. (-) Ausência de covariáveis, de validação e de comparação com outros métodos; distâncias ao longo da rede muito semelhantes às distâncias em linha reta.
Pitombo et al. (2015)	Domicílios	Probabilidade de escolha do modo carro, de escolha de um modo não motorizado e de escolha do transporte público	KO com Árvores de Decisão e validação ( <i>GeoMS</i> )	(+) Estimativa da escolha modal com base na posição geográfica e atributos socioeconômicos; variável de interesse gerada a partir de uma técnica não paramétrica. (-) Propagação de erros; apenas planejamento de curto prazo.
Gomes et al. (2016a, 2016b)	Domicílios	Probabilidade de escolha do modo individual motorizado	KO com Árvores de Decisão e KI ( <i>GeoMS</i> )	(+) Estimativa da escolha modal com base na posição geográfica e atributos socioeconômicos; variável de interesse gerada a partir de uma técnica não paramétrica; comparação com método direto; melhor desempenho da KI. (-) Propagação de erros; apenas planejamento de curto prazo.
Lindner e Pitombo (2017)	Domicílios	Probabilidade de escolha do modo de transporte público	KO com regressão logística e KI, ambas com validação ( <i>GeoMS</i> )	(+) Planejamento de médio e longo prazo da KO com <i>logit</i> ; estimativa da escolha modal com base na posição geográfica e atributos socioeconômicos; comparação com método direto; desempenho semelhante da KO e KI. (-) Propagação de erros na KO.
Chica-Olmo, Rodríguez-López e Chillón (2018)	Indivíduos	Probabilidade de escolha do modo a pé	KU com tendência <i>logit</i> e sem validação	(+) Planejamento de médio e longo prazo; análise do efeito de covariáveis; não propagação de erros. (-) Ausência de validação e comparação com métodos clássicos não espaciais.

Fonte: Os autores (2020).

Apesar de, na Geoestatística, a validação, seja ela cruzada ou não, ser realizada pelo teste do ponto fictício, no presente trabalho, considerou-se validação somente os casos em que os autores reservaram uma parte do banco de dados especificamente para esse fim. Quando todo o banco de dados é utilizado tanto na calibração do modelo quanto validação, tal validação foi chamada de cruzada. No Quadro 1, a coluna de método utilizado discrimina os trabalhos de acordo com a inclusão de uma etapa de validação ou ausência desta. Para os casos em que a validação foi cruzada, omitiu-se essa informação na célula.

O Quadro 1 revela que, de fato, a linha de pesquisa voltada a aplicações da Geoestatística à demanda por transportes é relativamente recente: o primeiro trabalho, dentre os analisados, remete ao ano de 2006. Porém, tendo em vista a grande variedade de estudos já realizados, nota-se uma evolução vertiginosa dessa linha. Além disso, os artigos citados limitam-se, em sua maioria, ao Brasil e Estados Unidos, com um representante na Espanha. Esses estudos de caso se concentram, no que se refere ao Brasil, nas etapas de

geração de viagens e escolha modal, enquanto nos EUA a ênfase é voltada ao volume de tráfego em trechos viários, ou seja, à fase de alocação de fluxo.

Embora a Geoestatística possua suposições que não são verificadas em todos os trabalhos, ressalta-se que a modelagem das etapas supracitadas é intrínseca à utilização de determinados suportes e técnicas. Nesse âmbito, recomenda-se continuar buscando soluções que convirjam com as restrições geoestatísticas, mas que não limitem as possibilidades de aplicações, já que, no caso da demanda por transportes, a ocorrência de dados escassos é recorrente.

A partir do Quadro 1, algumas lacunas de pesquisa podem ser depreendidas. Nota-se que a Krigagem Espaço-temporal foi utilizada apenas duas vezes entre os trabalhos abordados. É importante ressaltar que a coleta de dados de demanda ao longo do tempo não é algo incomum na engenharia de transportes: os postos de contagem de tráfego usualmente registram o número de veículos em intervalos de 5min. Além disso, a pesquisa de origem e destino, na qual baseiam-se os trabalhos de escolha modal e por zonas de tráfego, também dispõe de um relatório de atividades diárias praticadas pelo entrevistado no dia anterior à visita. Por sua vez, a coleta de informações de viagens a partir de *smartphones* também permite a criação de um banco de dados em painel. Tendo em vista que as informações de demanda ao longo do tempo apresentam, em menor ou maior grau, uma dependência temporal, a aplicação do interpolador espaço-temporal da Geoestatística se torna fortemente recomendável, contribuindo para o preenchimento de dados espaciais em séries históricas.

Essa sugestão pode ser estendida à Simulação Sequencial Gaussiana. O planejamento de transportes usualmente abrange a proposição de soluções, a problemas de mobilidade previamente diagnosticados, com base na análise de cenários hipotéticos. Nesse âmbito, a projeção da condição de transporte atual para o futuro leva em conta, no mínimo, três abordagens: realista, pessimista e otimista. Uma vez que a SSG é capaz de gerar vários cenários simulados, incluindo inúmeras combinações de resultados, sua aplicação a outras variáveis de demanda e estudos de caso se mostra bastante conveniente. Porém, ressalta-se que, tanto no caso da SSG quanto da Krigagem Espaço-temporal, a validação e comparação com outros métodos tradicionais são essenciais para o fortalecimento e consolidação da técnica.

Nesse contexto, observa-se a ausência de validação, ou realização apenas de validação cruzada, em vários trabalhos contidos no Quadro 1. Nos casos dos interpoladores univariados (Krigagem Indicativa e Krigagem Ordinária), uma justificativa plausível seria o fato de tais técnicas dependerem apenas da variável de interesse para a realização da modelagem. Dessa forma, quando se reserva uma parte do banco apenas para a validação propriamente dita, a perda de informação na calibração dos modelos pode impactar negativamente os resultados, o que se torna inconveniente se o propósito da investigação é destacar o potencial máximo do estimador. Entretanto, se o estudo não inclui uma etapa de validação, a capacidade do ferramental geoestatístico em tratar dados escassos não é devidamente comprovada.

O potencial de utilização da validação é aumentado por meio da Krigagem Universal, que se vale de informações acerca de variáveis explicativas, além da variável de interesse. O ideal seria validar o método a partir da remoção gradual de dados e verificar até que ponto o interpolador ainda consegue gerar estimativas aceitáveis. Nesse caso, poder-se-ia atestar a efetiva redução de custos proveniente da não necessidade de realização de um levantamento em todo o campo em que a variável de interesse atua, como, por exemplo, a pesquisa sobre/desce, da qual se originam as variáveis abordadas por Marques e Pitombo (2019) e Marques (2019). Validação com aumento gradual da porcentagem de dados ausentes foi observada, dentre os trabalhos do Quadro 1, apenas no estudo de Yang et al. (2018).

Porém, é preciso notar que, apesar de a Krigagem Universal ser um interpolador cuja validação permitiria resultados melhores que os da KO e KI, a necessidade de coleta de dados adicionais pode representar um *trade-off* nessa técnica. Dessa forma, cabe confrontar o estimador multivariado com os univariados sempre que possível, no intuito de questionar se os ganhos resultantes da KU compensam o custo adicional das informações auxiliares. A comparação da Geoestatística com outros métodos, espaciais e não espaciais, também é um artifício ausente em vários trabalhos do Quadro 1 e que se mostra de fundamental relevância para o avanço no estado da arte na linha de pesquisa sob análise.

Em contrapartida, a Krigagem Ordinária e Krigagem Indicativa conseguem gerar estimativas apenas para o cenário sob escrutínio, enquanto que a Krigagem Universal, além de permitir o conhecimento de como e quanto algumas variáveis externas afetam a demanda por transportes, ainda possibilita a projeção da variável



de interesse para cenários hipotéticos e futuros. Em outras palavras, a KU é capaz de contribuir para o planejamento de médio e longo prazo, enquanto a KO e KI se restringem ao curto prazo, no qual assume-se que não há variação significativa nos fatores intervenientes da demanda por transportes.

É importante enfatizar, porém, que mesmo que a KU tenha sido utilizada, com sucesso, em vários dos trabalhos presentes no Quadro 1, sua modelagem ainda pode ser melhorada. Notou-se, em alguns estudos (SELBY; KOCKELMAN, 2013; ZHANG; WANG, 2014) que a porcentagem de variação, da variável de interesse, explicada pelas várias explicativas era bastante elevada. Na KU, a estrutura espacial do modelo está atribuída aos resíduos, sobre os quais a função variograma é calculada. Se apenas uma pequena parte da variação da variável de interesse é imputada à estrutura espacial, pouca ou nenhuma melhoria pode ser proporcionada pela KU em comparação com a regressão linear clássica. Nesse contexto, torna-se necessário o desenvolvimento de métodos que consigam prever quantos e quais são os melhores preditores a serem incluídos na Krigagem Universal, como forma de otimizar o desempenho de suas estimativas e considerando o *trade-off* entre poder preditivo e número de variáveis explicativas, já comentado anteriormente.

Ressalta-se, também, que a KO e KU tradicional pressupõem normalidade das variáveis de interesse. Tendo em vista que as informações de demanda por transportes ocorrem, usualmente, na forma de contagens e apresentam assimetria, soluções para aproximar tais dados da distribuição normal incluem a transformação de Box-Cox, utilizada por Eom et al. (2006), Selby e Kockelman (2013), Marques (2019) e Marques e Pitombo (2019). Entretanto, não se pode afirmar que esse artifício seja suficiente para melhorar as estimativas, uma vez que, no cálculo das métricas de aderência, os valores reais da variável de interesse são comparados aos valores estimados já submetidos à transformação inversa, cujos resultados acabam sendo semelhantes aos que seriam obtidos caso a transformação não tivesse sido usada.

Uma outra forma de suavizar a assimetria dos dados de contagem seria por meio da aplicação de estimadores robustos do variograma, como o proposto por Cressie e Hawkins (1980), apropriado para os casos em que a distribuição é normal na região central, mas mais achatada que a normal nas caudas. Dos trabalhos citados, apenas Selby e Kockelman (2013) e Chi e Zheng (2013) utilizaram o estimador supracitado. Surge, dessa forma, a oportunidade de atestar o melhor desempenho das estimativas provenientes de estimadores robustos do variograma a fim de popularizar a utilização destes. A elaboração de métodos que consigam imbuir a natureza dos dados de contagem nos interpoladores geoestatísticos, assim como a KI foi criada para lidar com variáveis binárias, também é oportuno.

No que se refere à demanda ao longo da rede de transporte (sistema viário, linha de metrô ou de ônibus), o melhor desempenho dos interpoladores com a utilização de distâncias em rede, em detrimento das euclidianas, ainda não alcançou uma conclusão consensual. No caso de Marques (2019), o autor atribuiu a semelhança dos resultados das duas abordagens de distância, entre outros fatores, ao fato de a linha de ônibus utilizada no estudo ter um traçado mais ou menos retilíneo, o que aproximou as distâncias em rede das distâncias em linha reta. Dessa forma, a aplicação desse método a linhas com traçado mais curvilíneo e/ou irregular, e posterior comparação com o estimador tradicional é indicada.

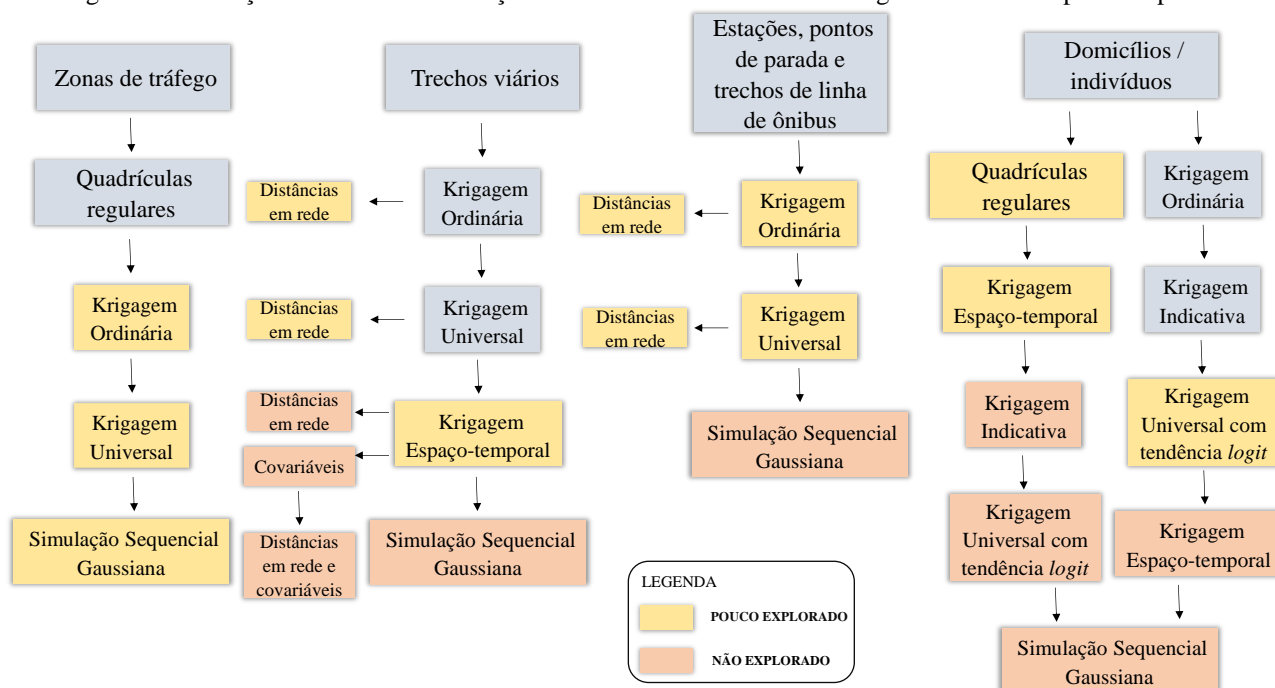
Cabe comentar, porém, que, conforme apontado por Selby e Kockelman (2013), a utilização de distâncias não euclidianas pode comprometer a obtenção de uma matriz de covariância positiva, necessária ao cômputo das estimativas de Krigagem. No intuito de contornar esse problema, Ver Hoef (2018) elaborou rotinas computacionais para a verificação prévia, a partir do arranjo espacial da rede de transporte, da possível não ocorrência do fenômeno supracitado. Esse autor também disponibilizou um método para a resolução desse obstáculo, caso ele fosse confirmado. Entretanto, dentre os trabalhos que utilizaram distâncias em rede, apenas o de Marques (2019), por ser o mais recente, se valeu da inspeção proposta por Ver Hoef (2018), fato que aponta a urgência de adequação do método a essa atualização nas próximas pesquisas a serem realizadas nesse âmbito.

Chama-se a atenção, também, para a utilização da Krigagem Universal conforme proposto por Chica-Olmo, Rodríguez-López e Chillón (2018). Tendo em vista que a variável de interesse do estudo era binária (escolha modal), ao invés da equação linear usualmente imbuída no estimador de KU, os autores aplicaram a função *logit*. Essa abordagem deve ser mais aprofundada no que diz respeito ao seu processo de calibração, bem como comparada aos outros tratamentos de escolha modal no âmbito da Geoestatística: Krigagem Indicativa e Krigagem Ordinária com Árvores de Decisão ou regressão logística. O confronto do desempenho

dessa abordagem com a do modelo sequencial de planejamento, que utiliza uma regressão logística não espacial, também se faz necessária a uma melhor compreensão de sua contribuição. Além disso, o cálculo de métricas de erros, baseado em uma amostra de validação, não foi incluído no trabalho supracitado e representa uma importante ferramenta para o ateste da precisão e confiabilidade do modelo. Se confirmada a sua aplicabilidade e eficácia, a utilização da tendência *logit* nas aplicações de Krigagem Universal à escolha modal poderá servir como um importante instrumento para a previsão da escolha modal em outros cenários além do analisado, o que não é possível na Krigagem Indicativa, e sem depender de tratamentos anteriores que podem levar à propagação de erros, como é o caso da Krigagem Ordinária.

Por fim, além de sugerir a aplicação da Krigagem Espaço-temporal aos dados em painel de demanda por transportes, recomenda-se, também, a inclusão de variáveis explicativas a esse interpolador, cuja formulação original exige apenas o valor da variável de interesse no espaço e no tempo. Assim como a Krigagem Universal demonstra, comumente, um desempenho melhor que o da Krigagem Ordinária, um interpolador espaço-temporal que contabilize a influência de variáveis externas pode trazer uma contribuição significativa à modelagem geoestatística da demanda por transportes. Ainda no que se refere a variáveis externas, encoraja-se a busca por interpoladores que consigam discriminar a heterogeneidade espacial dos parâmetros estimados para as covariáveis, que, em vários casos, variam no âmbito local. Sintetizando e complementando os comentários realizados na presente seção, a Figura 3 mostra a evolução da linha de pesquisa voltada a aplicações da Geoestatística na modelagem da demanda por transportes. Essa figura resume os avanços na utilização do suporte mais apropriado, bem como o aprofundamento dos interpoladores, destacando os casos que foram pouco ou ainda não explorados.

Figura 3 – Evolução resumida da interseção de Geoestatística com a modelagem da demanda por transportes.



Fonte: Os autores (2020).

## 5 DESAFIOS FUTUROS

A partir das lacunas de pesquisa consolidadas na Seção 4, alguns desafios de pesquisa, sintetizados na forma de questionamentos, podem ser apreendidos:

- Que aprimoramentos a modelagem geoestatística pode trazer para o planejamento urbano e do sistema de transportes, tanto no aspecto metodológico quanto prático?
- De que forma é possível otimizar a seleção de preditores para a modelagem geoestatística multivariada?
- Em se tratando de variáveis de demanda que ocorrem ao longo da rede viária, a utilização de

distâncias em rede pode, de fato, trazer resultados melhores que as distâncias em linha reta? Se sim, sob que circunstâncias e qual a dimensão dos ganhos?

- d) Como aplicar a Geoestatística ao caso de linhas de ônibus com trechos que se sobrepõem?
- e) A Krigagem Universal, com equação de tendência *logit*, pode realmente ser utilizada com sucesso para modelar variáveis de escolha modal?
- f) Considerar a não estacionariedade dos parâmetros estimados na modelagem geoestatística multivariada pode melhorar as estimativas de demanda?
- g) Qual a porcentagem mínima de pontos de parada, a serem amostrados em uma pesquisa de embarque e desembarque, suficiente para gerar estimativas aceitáveis de demanda nos pontos não amostrados?
- h) A Krigagem Espaço-temporal uni e multivariada pode ser aplicada com sucesso a banco de dados longitudinais de demanda por transportes? Se sim, esses modelos são competitivos frente às regressões tradicionais para dados em painel?
- i) De que forma o uso da Simulação Sequencial Gaussiana pode ser útil para o planejamento da operação do sistema de transporte público?

Espera-se que respostas a essas perguntas componham o escopo de pesquisas futuras voltadas à aplicação da Geoestatística na modelagem da demanda por transportes.

## 6 CONCLUSÕES E CONSIDERAÇÕES FINAIS

O planejamento de transportes depende, usualmente, da modelagem de variáveis de demanda por transportes, as quais são de difícil coleta, ou seja, exigem elevados recursos financeiros para serem catalogadas. Tendo em vista a dependência espacial normalmente presente nessas variáveis, na segunda metade dos anos 2000, aplicações da Geoestatística à demanda por transportes começaram a ser encontradas na literatura científica, as quais se valem da capacidade do ferramental geoestatístico em utilizar o máximo de informação acerca de uma variável espacial na estimativa de seus valores em locais não amostrados.

Nesse contexto, os trabalhos, que se concentram em três das quatro etapas do modelo sequencial (geração de viagens, escolha modal e alocação de fluxo), foram analisados com base em duas características principais: o suporte geográfico e o modelo geoestatístico utilizados. No que se refere ao suporte, as pesquisas realizadas no âmbito de pontos de parada, trechos de linhas de ônibus, estações de metrô e segmentos viários são as que mais se destacam, em se tratando da adequabilidade do nível de agregação das variáveis e da convergência com as suposições da Geoestatística. A utilização de grades regulares otimizadas, seja por algoritmos genéticos ou Simulação Sequencial Gaussiana, em substituição à informação original dada em zonas de tráfego e domicílios, é um dos avanços na busca pelo suporte geográfico mais apropriado a aplicações geoestatísticas à demanda por transportes.

O tipo de suporte está diretamente associado à contribuição do estudo na modelagem da demanda por transportes. Os trabalhos realizados no âmbito de zonas de tráfego, representadas por seu centroide, buscam incluir a dependência espacial das observações de viagens no processo de estimativa da variável de interesse, tendo em visto que tal característica não é contabilizada nos modelos clássicos de regressão. Essa contribuição se estende a todos os demais suportes, porém, no caso das malhas regulares, o processo de desagregação dos dados de demanda, quando provenientes de zonas de tráfego, ou agregação, quando oriunda de domicílios, soma, à já comentada contribuição, um avanço na procura de uma unidade de agregação que atenda os pressupostos da Geoestatística e ainda demonstre resultados satisfatórios. De uma forma geral, esses estudos cobrem a etapa de geração de viagens do modelo sequencial de planejamento.

Como representantes da etapa de alocação de fluxo, as aplicações da Geoestatística à modelagem do volume médio diário anual se constituem em uma importante contribuição ao conhecimento da demanda de veículos em trechos para os quais não se dispõe desse dado. Cabe lembrar que a análise de nível de serviço, projetos de pavimentos, de segurança viária, manutenção rodoviária, entre outros, usualmente dependem do volume de tráfego, entretanto, é inviável a instalação de contadores em todos os trechos de uma malha rodoviária. Em se tratando de uma linha de ônibus, a estimativa do carregamento de passageiros em trechos

não amostrados apresenta uma forte contribuição à identificação da ocupação crítica do veículo, comumente utilizada para o dimensionamento da frota de ônibus.

Ainda no que se refere ao volume de passageiros ao longo de uma linha de ônibus, a Geoestatística tem a função primordial de estimar o número de embarques e desembarques por ponto de parada. Nesse caso, a maioria das cidades não dispõem de contadores automáticos e tais informações, tão importantes para o planejamento da rede de transporte público, dependeriam de pesquisas de campo bastante onerosas. No intuito de superar esse obstáculo, a Geoestatística permite facultar, aos municípios, a realização de tais levantamentos em toda a rede de ônibus, mas ainda garantiria resultados aproximados aos da pesquisa completa.

A abordagem domiciliar, que se concentra, de uma forma geral, na etapa de escolha modal, se justifica na impossibilidade de realizar uma pesquisa domiciliar que alcance toda a população de determinada área urbana. Por outro lado, o conhecimento da forma com que a utilização dos variados modos de transporte varia ao longo da cidade é de primordial relevância para a elaboração de políticas e projetos públicos relacionados à mobilidade. Apesar dos vários tipos de suporte já utilizados no contexto da demanda por transportes, a aplicação da Geoestatística a novas unidades de agregação é encorajada, a fim de estimular a contínua evolução e abrangência da linha de pesquisa sob análise.

Por sua vez, os modelos utilizados estão condicionados, entre outros fatores, ao tipo de variável de interesse e propósito do estudo. Na escolha modal, a variável dependente é categórica, o que exige a aplicação de Krigagem Indicativa ou Krigagem Ordinária (com dados derivados de um tratamento anterior). Uma abordagem de Krigagem Universal à seleção do modo de transporte também foi encontrada. De uma forma geral, a KU, por incluir variáveis explicativas em sua formulação, consegue alcançar resultados melhores que os anteriores e apresenta contribuições ao planejamento de médio e longo prazo, sendo o modelo mais utilizado, juntamente com a Krigagem Ordinária, em aplicações a variáveis de demanda das três etapas supracitadas.

A popularidade da Krigagem Ordinária provém do fato de ser um interpolador univariado, isto é, que se ajusta aos casos nos quais a única informação disponível se refere à variável de interesse, e por assumir que a média dos dados é desconhecida e varia no âmbito local. Nesse contexto, porém, é necessário supor que os fatores intervenientes da demanda por transportes não variam ou que demonstram apenas uma mudança desprezível no período de tempo para o qual os resultados são válidos. Dessa forma, as aplicações de KO se restringem apenas ao planejamento de curto prazo.

Por sua vez, a Krigagem Espaço-temporal, cujas aplicações encontradas também não incluem variáveis explicativas, acrescenta, à KO, o aspecto temporal presente nos bancos de dados em painel, contribuindo para uma modelagem mais aprimorada das variáveis espaciais que também possuem dependência temporal. Apesar de várias variáveis de demanda por transportes sofrerem significativas flutuações ao longo do tempo, a dificuldade de se obter, para muitas delas, um banco de dados em painel, impossibilita a utilização da Krigagem Espaço-temporal. Contudo, quando tais informações estão disponíveis, o aspecto temporal não deve ser negligenciado. Por fim, cabe lembrar que a Geoestatística, por gerar uma superfície contínua de valores estimados, além de propiciar o conhecimento da variável em pontos não amostrados, ainda permite a visualização da variação espacial da variável de interesse, facilitando a identificação precisa de regiões críticas, padrões espaciais e, no caso dos interpoladores multivariados, a associação desses padrões com vários fatores intervenientes.

Em síntese, lacunas de pesquisa foram identificadas nas etapas de validação dos modelos, comparação com outras abordagens espaciais e não espaciais, utilização de distâncias em rede, aplicações da Krigagem Universal a variáveis de escolha modal e seleção de preditores para a KU. Atenção especial deverá ser dada à Simulação Sequencial Gaussiana e à Krigagem Espaço-temporal, abordagens que devem ditar a evolução da linha de pesquisa sob análise nos próximos anos.

Alguns modelos geoestatísticos, ainda não encontrados no âmbito da modelagem da demanda por transportes, também merecem destaque e podem/devem compor o novo estágio evolutivo dessa linha de pesquisa. A Simulação Sequencial por Indicador seria importante para a obtenção de diversos cenários em se tratando de problemas dicotômicos (ou relativos a variáveis contínuas divididas em classes) efetivamente úteis na área de Transportes, tais como aqueles referentes a variáveis de escolha modal. A Co-krigagem também pode ser um campo de aplicação futura, no intuito de realizar comparações com outros interpoladores

multivariados como a Krigagem Universal, por exemplo. Krigagem em Blocos e Simulação Sequencial Direta são outras técnicas cuja exploração também é oportuna.

Além disso, com a evolução da tecnologia, a implantação dos algoritmos da Geoestatística em programas de computador tornou a aplicação de tais interpoladores exequível, de forma que já não há mais justificativas para ignorá-los. Nesse contexto, o Quadro 1 revela ainda que, além de um *software* fechado, como o ArcGIS, a Geoestatística tem sido encontrada em outros programas, livres e abertos, a exemplo da interface de programação R. A facilidade de acesso a recursos computacionais para a utilização da Geoestatística contribui para consolidar, definitivamente, o espaço dessa ferramenta na modelagem da demanda por transportes, com possibilidade de contínua expansão.

## Agradecimentos

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Brasil - Processo 2019 / 12054-4) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasil - Processo 304345 / 2019-9).

## Contribuição dos Autores

O primeiro autor (Samuel de França Marques) foi responsável pela Conceptualização, Investigação, Visualização, Redação - minuta inicial e Redação - revisão e edição; a segunda autora (Cira Souza Pitombo) foi responsável pela Conceptualização, Supervisão e Redação - revisão e edição.

## Conflitos de Interesse

Os autores declaram que não há conflitos de interesse.

## Referências

- BISWAS, S.; CHAKRABORTY S.; CHANDRA, S.; GHOSH, I. Kriging-based approach for estimation of vehicular speed and passenger car units on an urban arterial. **Journal of Transportation Engineering, Part A: Systems**, v. 143, n. 3, p. 4016013, 2017. American Society of Civil Engineers. DOI: 10.1061/JTEPBS.0000031.
- CARVALHO, S. D. P. C. E.; RODRIGUEZ, L. C. E.; SILVA, L. D.; CARVALHO, L. M. T.; CALEGARIO, N.; LIMA, M. P.; SILVA, C. A.; MENDONÇA, A. R.; NICOLETTI, M. F. Predição do volume de árvores integrando Lidar e Geoestatística. **Scientia Forestalis/Forest Sciences**, v. 43, n. 107, p. 627–637, 2015.
- CHI, G.; ZHENG, Y. Estimating transport footprint along highways at local levels: a combination of network analysis and kriging methods. **International Journal of Sustainable Transportation**, v. 7, n. 3, p. 261–273, 2013. DOI: 10.1080/15568318.2013.710150.
- CHICA-OLMO, J.; RODRÍGUEZ-LÓPEZ, C.; CHILLÓN, P. Effect of distance from home to school and spatial dependence between homes on mode of commuting to school. **Journal of Transport Geography**, v. 72, p. 1–12, 2018. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0966692316307013>>. Acesso em: jul. 2020.
- CRESSIE, N. A. C. **Statistics for spatial data**. John Wiley & Sons, Inc., 1993.
- CRESSIE, N.; HAWKINS, D. M. Robust estimation of the variogram: I. **Journal of the International Association for Mathematical Geology**, v. 12, n. 2, p. 115–125, 1980. DOI: 10.1007/BF01035243.
- DEUTSCH, C. V.; JOURNAL, A. G. **GSLIB: Geostatistical Software Library and User's Guide**, Oxford University Press, 1998.
- EOM, J. K.; PARK, M. S.; HEO, T.-Y.; HUNTSINGER, L. F. Improving the prediction of annual average daily traffic for nonfreeway facilities by applying a spatial statistical method. **Transportation Research Record**, , n. 1968, p. 22–29, 2006. DOI: 10.1177/0361198106196800103.

- GOMES, M. M.; PITOMBO, C. S.; PIRDAVANI, A.; BRIJS, T. Geostatistical approach to estimate car occupant fatalities in traffic accidents. **Revista Brasileira de Cartografia**, v. 70, n. 4, p. 1231–1256, 2018.
- GOMES, V. A.; PITOMBO, C. S.; ROCHA, S. S.; SALGUEIRO, A. R. Kriging geostatistical methods for travel mode choice: a spatial data analysis to travel demand forecasting. **Open Journal of Statistics**, v. 06, n. 03, p. 514–527, 2016a. DOI: 10.4236/ojs.2016.63044.
- GOMES, V. A.; PITOMBO, C. S.; ROCHA, S. S.; SALGUEIRO, A. R. Estudo comparativo do uso de krigagem para estimação da escolha do modo particular motorizado. **Revista Brasileira de Cartografia**, v. 68, n. 4 SE-Artigos, 2016b. Disponível em: <<http://www.seer.ufu.br/index.php/revistabrasileiracartografia/article/view/44279>>. Acesso em: jul. 2020.
- GOOVAERTS, P. Kriging and Semivariogram Deconvolution in the Presence of Irregular Geographical Units. **Mathematical geology**, v. 40, n. 1, p. 101–128, 2008. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2518693/>>. Acesso em: jul. 2020.
- GOOVAERTS, P. Medical geography: A promising field of application for geostatistics. **Mathematical Geosciences**, v. 41, p. 243-264, 2009a.
- GOOVAERTS, P. Combining area-based and individual-level data in the geostatistical mapping of late-stage cancer incidence. **Spatial and Spatio-temporal Epidemiology**, v. 1, n. 1, p. 61-71, 2009b.
- GRÄLER, B.; PEBESMA, E.; HEUVELINK, G. Spatio-temporal interpolation using gstat. **RFID Journal**, v. 8, n. 1, p. 204-218, 2016. Disponível em: <<https://cran.r-project.org/web/packages/gstat/vignettes/spatio-temporal-kriging.pdf>>. Acesso em: jul. 2020.
- GUNDOGDU, I. B.; FACULTY, E. Risk governance for traffic accidents by Geostatistical Analyst methods. **International Journal of Research in Engineering and Science (IJRES)**, v. 2, n. 9, p. 35–40, 2014.
- HENGL, T.; HEUVELINK, G. B. M.; ROSSITER, D. G. About regression-kriging: from equations to case studies. **Computers & Geosciences**, v. 33, n. 10, p. 1301–1315, 2007. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0098300407001008>>. Acesso em: jul. 2020.
- VER HOEF, J. M. Kriging models for linear networks and non-Euclidean distances: Cautions and solutions. **Methods in Ecology and Evolution**, v. 9, n. 6, p. 1600–1613, 2018. John Wiley & Sons, Ltd (10.1111). DOI: 10.1111/2041-210X.12979.
- JOURNEL, A. G. Nonparametric estimation of spatial distributions. **Journal of the International Association for Mathematical Geology**, v. 15, p. 445-468, 1983. DOI: 10.1007/BF01031292.
- JOURNEL, A. G.; HUIJBREGTS, C. J. **Mining geostatistics**. Academic press, 1978.
- KERRY, R.; GOOVAERTS, P.; GIMÉNEZ, D.; OUDEMANS, P.; MUÑIZ, E. Investigating geostatistical methods to model within-field yield variability of cranberries for potential management zones. **Precision Agriculture**, vol. 17, p. 247-273, 2016.
- KLATKO, T. J.; USMAN, S. T.; MATTHEW, V.; et al. Addressing the local-road VMT estimation problem using spatial interpolation techniques. **Journal of Transportation Engineering, Part A: Systems**, v. 143, n. 8, p. 4017038, 2017. American Society of Civil Engineers. DOI: 10.1061/JTEPBS.0000064.
- KRIGE, D. G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Southern African Institute of Mining and Metallurgy**, v. 52, n. 6, p. 119–139, 1951. Southern African Institute of Mining and Metallurgy.
- LINDNER, A.; PITOMBO, C. S. A conjoint approach of spatial statistics and a traditional method for travel mode choice issues. **Journal of Geovisualization and Spatial Analysis**, v. 2, n. 1, 2017. DOI: 10.1007/s41651-017-0008-0.
- LINDNER, A.; PITOMBO, C. S. Sequential Gaussian simulation as a promising tool in travel demand modeling. **Journal of Geovisualization and Spatial Analysis**, v. 3, n. 2, p. 15, 2019. DOI: 10.1007/s41651-019-0038-x.
- LINDNER, A.; PITOMBO, C. S.; ROCHA, S. S.; QUINTANILHA, J. A. Estimation of transit trip production using Factorial Kriging with External Drift: an aggregated data case study. **Geo-spatial Information**

- Science**, v. 19, n. 4, p. 245–254, 2016. Taylor & Francis. DOI.: 10.1080/10095020.2016.1260811.
- MAJUMDAR, A.; NOLAND, R. B.; OCHIENG, W. Y. A spatial and temporal analysis of safety-belt usage and safety-belt laws. **Accident Analysis & Prevention**, v. 36, n. 4, p. 551–560, 2004. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0001457503000617>>. Acesso em: jul. 2020.
- MANEPALLI, U. R. R.; BHAM, G. H. Crash prediction: evaluation of empirical Bayes and kriging Methods. In: International Conference on Road Safety and Simulation, 3, 2011, Indianapolis, USA. **Anais... Transportation Research Board**, 2011.
- MARQUES, S. F. **Estimativa do volume de passageiros ao longo de uma linha de transporte público por ônibus a partir da Geoestatística**. 98 p. Dissertação (Mestrado em Engenharia de Transportes) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2019. DOI.: 10.11606/D.18.2019.tde-26042019-110232.
- MARQUES, S. F.; PITOMBO, C. S. Estimativa do volume de passageiros ao longo de uma linha de transporte público por ônibus a partir da Geoestatística. **Transportes**, v. 27, n. 3, 2019. DOI: 10.14295/transportes.v27i3.2007.
- MATHERON, G. Principles of geostatistics. **Economic Geology**, v. 58, n. 8, p. 1246–1266, 1963.
- MATHERON, G. **The theory of regionalized variables and its applications**. Paris: Les Cahiers du Centre de Morphologie Mathématique in Fontainebleu, 1971.
- OFFOR, K. J.; VACI, L.; MIHAYLOVA, L. S. Traffic estimation for large urban road network with high missing data ratio. **Sensors**, v. 19, n. 12, p. 2813, 2019. DOI: 10.3390/s19122813.
- ORTÚZAR, J. DE D.; WILLUMSEN, L. G. **Modelling Transport**. John Wiley & Sons, 2011.
- PIANUCCI, M. N.; PITOMBO, C. S.; CUNHA, A. L.; SEGANTINE, P. C. L. Previsão da demanda por viagens domiciliares através de método sequencial baseado em população sintética e redes neurais artificiais. **Transportes**, v. 27, n. 4, p. 1–23, 2019. DOI: 10.14295/transportes.v27i4.1406.
- PITOMBO, C. S.; SALGUEIRO, A. R.; DA COSTA, A. S. G.; ISLER, C. A. A two-step method for mode choice estimation with socioeconomic and spatial information. **Spatial Statistics**, v. 11, p. 45–64, 2015. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2211675314000645>>. Acesso em: jul. 2020.
- REMY, N.; BOUCHER, A.; WU, J. **Applied geostatistics with SGeMS: a user's guide**. Cambridge University Press, 2009.
- ROCHA, S. S. **Uso de Algoritmos Genéticos para otimização de modelagem geoestatística aplicada à demanda por transportes**, 2019. Escola de Engenharia de São Carlos - Universidade de São Paulo. DOI.: 10.11606/T.18.2019.tde-23052019-114238.
- ROCHA, S. S.; LINDNER, A.; PITOMBO, C. S. Proposal of a geostatistical procedure for transportation planning field. **Boletim de Ciências Geodésicas**, v. 23, p. 636–653, 2017. DOI: 10.1590/s1982-21702017000400042.
- ROCHA, S. S.; PITOMBO, C. S.; SALGUEIRO, A. R. Interpolação espacial de viagens urbanas por transporte coletivo através de amostra sistemática desagregada artificial. **Revista Brasileira de Cartografia**, v. 68, n. 4, 2016. Disponível em: <<http://www.seer.ufu.br/index.php/revistabrasileiracartografia/article/view/44277>>. Acesso em: jul. 2020.
- SELBY, B.; KOCKELMAN, K. M. Spatial prediction of traffic levels in unmeasured locations: Applications of universal kriging and geographically weighted regression. **Journal of Transport Geography**, v. 29, p. 24–32, 2013. DOI.: 10.1016/j.jtrangeo.2012.12.009.
- SHAMO, B.; ASA, E.; MEMBAH, J. Linear spatial interpolation and analysis of annual average daily traffic data. **Journal of Computing in Civil Engineering**, v. 29, n. 1, p. 4014022, 2015. American Society of Civil Engineers. DOI.: 10.1061/(ASCE)CP.1943-5487.0000281.
- STELZENMÜLLER, V.; EHRICH, S.; ZAUKE, G. P. Impact of additional small-scale survey data on the geostatistical analyses of demersal fish species in the North Sea. **Scientia Marina**, vol. 69, n. 4, p. 587-

602, 2005.

- WANG, S.; MAO, G. Missing data estimation for traffic volume by searching an optimum closed cut in urban networks. **IEEE Transactions on Intelligent Transportation Systems**, v. 20, n. 1, p. 75–86, 2019. DOI.: 10.1109/TITS.2018.2801808.
- WANG, X.; KOCKELMAN, K. Forecasting network data. **Transportation Research Record: Journal of the Transportation Research Board**, v. 2105, p. 100–108, 2009. Transportation Research Board. DOI.: 10.3141/2105-13.
- YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística: conceitos e aplicações**. Oficina de textos, 2015.
- YANG, H.; YANG, J.; HAN, L. D.; LIU, X.; PU, L.; CHIN, S.; HWANG, H. A Kriging based spatiotemporal approach for traffic volume data imputation. **PloS one**, v. 13, n. 4, p. e0195957, 2018. Public Library of Science. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/29664928>>. Acesso em: jul. 2020.
- YOON, S. Y.; RAVULAPARTHY, S. K.; GOULIAS, K. G. Dynamic diurnal social taxonomy of urban environments using data from a geocoded time use activity-travel diary and point-based business establishment inventory. **Transportation Research Part A: Policy and Practice**, v. 68, p. 3–17, 2014. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0965856414000056>>. Acesso em: jul. 2020
- ZHANG, D.; WANG, X. C. Transit ridership estimation with network Kriging: A case study of Second Avenue Subway, NYC. **Journal of Transport Geography**, v. 41, p. 107–115, 2014. DOI.: 10.1016/j.jtrangeo.2014.08.021.
- ZOU, H.; YUE, Y.; LI, Q.; YEH, A. G. O. An improved distance metric for the interpolation of link-based traffic data using kriging: a case study of a large-scale urban road network. **International Journal of Geographical Information Science**, v. 26, n. 4, p. 667–689, 2012. Taylor & Francis. DOI.: 10.1080/13658816.2011.609488.

## Biografia dos autores



Samuel de França Marques é natural de Porto Nacional–TO. Formado em Engenharia Civil pela Universidade Federal do Tocantins (2016), é mestre em ciências (2019) e, atualmente, doutorando no Programa de Pós-Graduação em Engenharia de Transportes da Escola de Engenharia de São Carlos, Universidade de São Paulo. Reside na cidade de São Carlos–SP, onde se dedica à aplicação da estatística espacial na modelagem/estimativa do volume de passageiros ao longo de linhas de ônibus. Em 2018, foi agraciado com o Prêmio ANPET de Produção Científica. Tem interesse nas áreas de modelos e técnicas para o planejamento de transportes e segurança viária.



Professora Associada do Departamento de Engenharia de Transportes da Escola de Engenharia de São Carlos da Universidade de São Paulo. Fez o seu pós-doutoramento na Universidade Técnica de Lisboa, Portugal (2007-2009), e no Laboratório de Transportes e Mobilidade, Escola de Arquitetura, Engenharia Civil e Ambiental, EPFL, Lausanne, Suíça (2016-2017). Seus interesses de pesquisa incluem: Geoestatística aplicada ao planejamento de transportes, novas técnicas de coleta de dados para previsão da demanda por viagens e modelagem de acidentes de trânsito. Orientou 18 alunos de pós-graduação e escreveu 36 publicações em periódicos internacionais e nacionais.



Esta obra está licenciada com uma Licença [Creative Commons Atribuição 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/) – CC BY. Esta licença permite que outros distribuam, remixem, adaptem e criem a partir do seu trabalho, mesmo para fins comerciais, desde que lhe atribuam o devido crédito pela criação original.