# RESEARCH STATEMENT

My research is primarily focused on two interrelated areas: (1) smart city operations, encompassing both foundational and emerging business models mainly in transportation and logistics; and (2) analytics for social good, which addresses critical issues related to system robustness, social well-being, and sustainability. In the sections that follow, I will detail these themes, highlighting my contributions and outlining the implications of my work on the field.

## 1 Smart City Operations

In this section, I will elaborate on my first research theme, which tackles fundamental and emerging challenges in urban environments across both the private and public sectors. By integrating advanced analytical methods and robust optimization techniques, my work provides actionable insights that address key issues in smart city operations.

### Analytics for Emerging Last-mile Delivery Problems

The rapid expansion of e-commerce has compelled e-retailers and local businesses to enhance their logistics operations, striving to deliver goods promptly, reliably, and cost-effectively. This imperative has significantly heightened the importance of optimizing last-mile delivery systems, where improving both cost-effectiveness and service quality is crucial. My research contributes directly to these challenges through two key papers — the first tackles strategic-level workforce management issues, while the second addresses operational challenges in delivery routing decisions.

- Robust Workforce Management with Crowdsourced Delivery. With Chun Cheng and Melvyn Sim. Forthcoming in ***Operations Research***.
- Service Oriented Considerate Routing: Data, Predictions and Robust Decisions. With Zhixing Luo, Stanley Frederick W. T. Lim, Caihua Chen, and Melvyn Sim. Major revision in ***Management Science***.

In the first paper, we investigate how crowdsourced delivery platforms with both contracted and ad-hoc couriers can effectively manage their workforce to meet delivery demands amidst uncertainties. Our objective is to minimize the hiring costs of contracted couriers and the crowdsourcing costs of ad-hoc couriers while considering the uncertain availability and behavior of the latter. Due to the complication of calibrating these uncertainties through data-driven approaches, we instead introduce a basic reduced information model to estimate the upper bound of the crowdsourcing cost and a generalized reduced information model to obtain a tighter bound. Subsequently, we formulate a robust satisficing model associated with the generalized reduced information model and show that a binary search algorithm can tackle the model exactly by solving a modest number of convex optimization problems.

In the second study, we refine service-oriented routing by integrating punctuality considerations, specifically addressing the challenge of ensuring timely deliveries by couriers. Leveraging a comprehensive dataset from a cold chain logistics company, we demonstrate that traditional reliance on travel distance alone falls short in predicting accurate delivery times. Our analysis underscores the importance of incorporating couriers' individual attributes—such as fixed effects and workload—into predictive models to enhance performance. Our Service Oriented Routing model incorporates these elements along with couriers' location familiarity, using the novel *Courier Assigned Location Mismatch (CALM)* metric to optimize route assignments within familiar areas, thereby reducing courier stress and enhancing delivery efficiency. Our Service Oriented Considerate Routing model, which minimizes CALM under routing distance constraints, is shown essentially aligned with the robust satisficing strategy that mitigates the risks of predictive inaccuracies and model misspecifications. Additionally, we have developed and implemented both exact and heuristic algorithms to effectively solve our model. These algorithms demonstrate superior out-of-sample performance using our real-world dataset.

**Implications.** The first study illustrates the application of our novel 'reduced information model' which bypasses the complex human decision-making process employed by couriers, achieving robust strategic decisions with streamlined data requirements. The second study introduces an axiomatic metric to quantify couriers'

familiarity with locations, seamlessly integrating this metric into the downstream prediction and routing model to improve the decisions. Collectively, these studies provide profound policy implications for enhancing operational effectiveness and service quality to a broad array of urban logistics platforms, including delivery companies, ride-sharing services, and taxi applications.

## Data-driven Robust Transportation and Logistic Models

In the realm of transportation and logistics, uncertainty remains a pervasive challenge. The era of big data presents unprecedented opportunities to refine foundational models and enhance decision-making processes. In this context, my co-authors and I have developed two research papers spanning maritime, land, and air transport networks.

- Vessel Deployment with Limited Information: Distributionally Robust Chance Constrained Models. With Zhi Chen, Zhenzhen Zhang and Andrew Lim. ***Transportation Research Part B: Methodological* 161 (2022): 197-217**.
- Distributionally Robust Chance Constrained $p$-Hub Center Problem. With Zhi Chen and Zhenzhen Zhang. ***Informs Journal on Computing* 35(6) (2023): 1361-1382**.

In the first study, we tackle the intrinsic uncertainties of the liner shipping industry, where demand can fluctuate significantly due to geopolitical shifts and other unpredictable factors. We developed a distributionally robust chance-constrained model that efficiently hedges against demand uncertainty within a defined ambiguity set. This set considers distributions with identical means and first-order dispersion constrained within a polyhedral set. Our theoretical framework not only encompasses existing models as special cases but also demonstrates substantial improvements in managing uncertainty over traditional benchmarks.

The second work addresses a critical challenge in network and logistic management: the $p$-hub center problem. This problem involves determining the optimal placement of a fixed number of hubs in a hub-and-spoke network to ensure that the maximum travel time between any two nodes does not surpass a specified threshold with a certain probability. The hub-and-spoke network serves as a fundamental building block for air transport and ground shipping systems, providing essential infrastructure for efficient routing and distribution. We introduced a novel, data-driven model utilizing the Wasserstein ambiguity set to derive tractable reformulations that are solvable with standard optimization software. Additionally, we have developed an innovative acceleration technique that significantly enhances computational efficiency, achieving speeds thousands of times faster than conventional methods. This advancement not only streamlines the computational process but also enhances the practical applicability of our model in dynamic and large-scale network settings.

**Implications.** For the liner shipping industry, our distributionally robust chance-constrained model offers a more reliable framework for managing demand uncertainty, which is crucial for optimizing operational costs and enhancing service reliability in an industry vulnerable to global economic shifts and geopolitical tensions. In terms of the $p$-hub center problem, our advancements extend beyond theoretical contributions to provide tangible benefits for network design, particularly in logistics and air transportation. Our strategy is particularly critical in scenarios where timely delivery is paramount, such as in emergency response logistics or in the global distribution chains of e-commerce giants. The solving acceleration technique we developed not only simplifies the computational process but also enables real-time decision-making, allowing companies to respond swiftly to changing conditions and maintain service excellence across their operations.

## Smart Urban Policing Resource Allocation

In the complex ecosystem of smart city operations, the integration of proactive policing strategies plays a critical role in enhancing public safety and reducing the economic and social costs associated with crime. As urban centers grow, the demand for innovative and efficient law enforcement methods that prevent crime rather than react to it becomes imperative. My research contributes to this field by introducing a novel model that optimizes the allocation of policing resources to not only deter criminal activities but also to strategically enhance the overall safety infrastructure of smart cities. The following paper is pivotal in advancing the role of data-driven and preventive policing measures, aligning with broader smart city goals to create safer, more resilient urban environments.

- Proactive Policing: A Resource Allocation Model for Crime Prevention with Deterrence Effect. With Long He and Xiaobo Li. Major revision in ***Operations Research***.

This paper advances the understanding of police resource allocation by proposing a proactive model focused on minimizing the overall impact and cost of potential crimes across multiple locations. Moving beyond the traditional reactive methods detailed in prior literature, our approach centers on crime prevention through strategic deterrence. We utilize a multinomial logit model to accurately estimate the impact of police presence on the geographic distribution of crime, highlighting two critical aspects of deterrence in proactive policing: crime control diffusion and crime displacement. These elements are crucial for addressing contemporary crime patterns from both criminological and economic perspectives. Further, our work explores the structural properties of the problem, drawing parallels with mixture-of-logits assortment optimization, which enhances the theoretical results of our approach. We have developed reformulations that allow our model to be implemented as mixed-integer linear or conic programs, facilitating easy application with standard optimization software. To demonstrate the practical impact of our model, we present a data-driven case study that evaluates the strategic placement of surveillance cameras in New York City, showcasing significant improvements in crime prevention.

**Implications.** This paper marks a transformative shift in urban safety strategies by introducing a proactive model for police resource allocation, distinct from the traditional reactive approaches prevalent in sectors like emergency medical services. By emphasizing crime prevention through strategic deterrence, our strategy provides crucial insights for smart city operations, enabling decision-makers to effectively prevent crime rather than simply respond to it. This pioneering approach not only enhances urban safety but also serves as a vital tool for policymakers aiming to build more secure and resilient smart cities.

## 2 Analytics for Social Good

This section highlights my research on utilizing advanced analytical techniques to tackle pressing societal challenges and foster solutions for social good. It delves into three critical areas related to my past and ongoing research: developing trustworthy machine learning models, optimizing resource allocation for enhanced social well-being, and designing sustainable network for net-zero data centers. Each subsection illustrates how data analytics and optimization can significantly advance social benefits, adapting to the needs of a rapidly evolving world.

### Trustable Machine Learning

Trustable machine learning involves creating robust and safe systems that users and stakeholders can rely on, particularly in critical applications such as autonomous vehicles and healthcare systems. This aspect of machine learning is crucial in operations research, as it ensures the security and integrity of models that are essential for promoting social good and harnessing the power of AI in operations management. In this vein, my research includes two notable publications. Both papers specifically tackle the vulnerabilities in deep learning models that process 3D data — an area of growing importance in machine learning research.

- On Isometry Robustness of Deep 3D Point Cloud Models under Adversarial Attacks. With Yuwei Wu, Caihua Chen, and Andrew Lim. ***CVPR2020*** (top AI conference).
- PointBA: Towards Backdoor Attacks in 3D Point Cloud. With Xinke Li, Zhirui Chen, Zekun Tong, Yabang Zhao, Andrew Lim and Joey Tianyi Zhou. ***ICCV2021*** (top AI conference).

My first study delves into the robustness of 3D deep neural networks during the testing phase, highlighting a critical issue unique to 3D models — the isometric robustness. We are the first to discover that minor rotational manipulations can drastically compromise model integrity, achieving an attack success rate exceeding 95%. This vulnerability demonstrates strong transferability across various 3D network architectures. The second paper pioneers the exploration of backdoor attacks during the training phase. We have developed a comprehensive framework for introducing data-poisoned backdoor attacks in 3D deep learning environments. The triggers for these attacks are ingeniously mundane, incorporating everyday objects such as small balls or subtly altered point clouds, making them highly effective and challenging to detect. These two works not only advance our

understanding of robustness and security in 3D neural networks but also contribute to broader discussions in machine learning security, including the generative adversarial networks and transformers.

**Implications.** As operations research increasingly integrates with artificial intelligence, particularly large language models, the implications of our research become crucial. Our studies highlight significant robustness and integrity challenges in deep learning models, emphasizing the need for enhanced security measures as these technologies merge. Addressing these vulnerabilities is essential for ensuring the reliability of AI systems in critical applications, safeguarding the trust and effectiveness of AI-related operations research as it evolves.

## Resource Allocation for Social Well-being

Resource allocation in the public domain is critical for enhancing social well-being and ensuring equitable access to essential services. This encompasses a wide range of activities, from deploying policing resources to prevent crime to distributing food to those in need. Effective allocation of these resources not only addresses immediate needs but also builds long-term resilience in communities. In the realm of law enforcement, strategic resource allocation can deter criminal activities and promote safer neighborhoods. Similarly, in the context of food security, efficient logistics and distribution strategies are vital for food banks, contributing to community resilience and sustainability by reducing food waste and ensuring that surplus food reaches those who need it most instead of going to landfills.

- Proactive Policing: A Resource Allocation Model for Crime Prevention with Deterrence Effect. With Long He and Xiaobo Li. Major revision in ***Operations Research***.
- Joint Routing and Resource Allocation for Nonprofit Operations under Uncertainty. With Sheng Liu and Mengling Zhang. Work in progress.

The first study has been elaborated before thus we only highlight its societal implication below. In the second study, we address the crucial role of food banks in nonprofit operations, focusing on optimizing the distribution of resources to those in need. Our research develops a joint routing and sequential resource allocation model for food banks, formulated as a multistage robust optimization. The goal is to maximize the (worst-case) minimum fill rate across all food banks, ensuring efficient and equitable resource distribution among locations. Our findings reveal that an optimal policy can be established based on a fixed fill rate, which inspires further improvements through an adaptive robust linear decision rule. Utilizing real data from the food banks in London, our model demonstrates a marked improvement over existing strategies, offering a more effective approach to managing food bank logistics.

**Implications.** The first paper introduces a proactive policing model that employs deterrence and preemptive strategies, significantly improving public safety and advancing social well-being by preventing crimes before they happen. This innovative approach demonstrates how data analytics and optimization can transform traditional practices, offering a new paradigm for resource allocation in law enforcement. The second paper enhances the efficiency of food banks through an robust resource allocation model that ensures equitable distribution. By applying data-driven strategies to address logistical challenges, this work highlights the vast potential of analytics to improve social welfare.

## Network Design of Net-Zero Data Centers

Data centers are major electricity consumers, underscoring the urgent need for sustainable operations. Tech giants like AWS, Google, Meta, and Microsoft are pursuing net-zero emissions, challenging due to the variability and intermittency of renewable energy sources. My recent research addresses these issues, aiming to develop resilient and sustainable network solutions that align with these environmental targets.

- Network Design of Net-Zero Data Centers for Sustainable and Efficient Operations. With Sheng Liu, Ho-yin Mark, and Shixuan Zhang. Work in progression.

Our research introduces the first comprehensive network design model for data centers aiming for net-zero emissions, specifically crafted to meet the significant energy demands of modern computational technologies. This model addresses key challenges in the rapidly evolving field of sustainable data centers, including optimal

setup, strategic job allocation, and the incorporation of renewable energy sources. The fluctuating nature of green energy sources like solar and wind introduces complexity, especially in matching energy supply with the variable computational loads of data centers. At the core of our approach is the Stochastic Dual Dynamic Programming method, which excels in managing the uncertainties associated with renewable energy production and the demands of data processing. Our strategy not only helps control the carbon emissions of data centers but also guarantees their ability to consistently fulfill extensive computational needs.

**Implications.** The implications of our work are particularly significant for data centers involved in processing large language models, which require immense computational resources and can cost millions of dollars per training session. By implementing our net-zero emission model, these data centers can drastically reduce their environmental impact while managing costs more effectively. This strategic approach not only supports the technological advancement of AI but also aligns with global environmental goals, demonstrating a scalable solution for energy-intensive industries worldwide.