Citation: CHEN Xianlong, ZHANG Hua, MA Yilin, SONG Cheng, WEI He. Reflections on the Paradigm Shift in Travel Demand Models in the Big Data Era[J], Urban Transport of China, 2024 (4).

Reflections on the Paradigm Shift in Travel Demand Models in the Big Data Era

CHEN Xianlong[1,2], ZHANG Hua[3], MA Yilin[4], SONG Cheng[1,2], WEI He[5]

1. Guangzhou Transport Planning Research Institute Co., Ltd., Guangzhou 510030, China;
2. Guangdong Sustainable Transportation Engineering and Technology Research Center, Guangzhou 510030, China;
3. National Maglev Transportation Engineering R&D Center, Tongji University, Shanghai 201804, China;
4. Beijing Transport Institute, Beijing 100073, China;
5. Beijing Municipal Institute of City Planning & Design, Beijing 100045, China

Abstract: Big data has brought revolutionary changes to the developing basis of travel demand models. Reflections on the paradigm shift is not only an adaptation to changes in data conditions in the new era but also a necessary requirement for improving the accuracy of travel demand models. This paper summarizes and contemplates four paradigm shifts in travel demand models based on changes in foundational data conditions. These shifts include: enhancing the explanatory power of models on travel behavior by transitioning from mathematical optimization to causal inference; clarifying the physical meaning of models by moving from proportional factors to probabilistic sampling; achieving inheritance and iterative evolution of current demand by transitioning from holistic reconstruction to incremental models; and improving the precision of travel demand models by moving from finite constraint convergence to prior empirical justification. This paper points out that for both practical application and scientific research in transportation governance, model accuracy is the sole and highest standard for evaluating model quality. Emphasizing paradigm shifts is not meant to negate previous technical approaches in travel demand models but to innovate and develop through inheritance, thus enhancing the capacity of models to simulate and predict the real world. **DOI:** 10.13813/j.cn11-5141/u.2024.0018-en

Keywords: travel demand model; big data; causal inference; probabilistic sampling; incremental model; prior empirical evidence; paradigm shift

1 Development and evolution of traffic demand model

Since the Detroit Metropolitan Area Traffic Study (DMATS) [1] and Chicago Area Transport Study (CATS) [2] were conducted in the 1950s, aggregated models represented by the four-step model (FSM) have continuously developed and improved [3–4]; In the mid to late 1970s, the semi aggregated tour-based model (TBM) and non-aggregated activity-based model (ABM) were developed, as well as the later agent-based model (AgBM) and hybrid travel demand model. With the arrival of the big data era [5] in 2010, mobile signaling, Internet space-time location, road checkpoint monitoring and other data applications have become popular, and the data basis of the traffic demand model has undergone qualitative changes. Data-driven in-depth learning, enhanced learning and other artificial intelligence (AI) modeling methods [6–7] have been continuously applied to the development of the traffic demand model.

For over 70 years, the traffic demand model technology has continuously developed and advanced in form, and the explanation of the travel behavior has been continuously improved. In terms of the development paradigm, it can be divided into two categories: the travel activity modeling method based on feature investigation and the data-driven model with machine learning as the core (see Fig. 1).

From the perspective of travel activity modeling, although the interpretation of travel demand by different modeling methods varies from aggregated to semi-aggregated and then to non-aggregated, their essence is, on the basis of small sample feature survey data, the process to use mathematical optimization and statistical analysis methods to calibrate parameters and restore the whole, and model validation is still mainly based on the traffic flow observation data of the verification line. However, even if the error between the model results and the observation data of the verification line is small, whether the traffic demand model can reflect the real travel OD still needs further verification [8]. Although the travel activity modeling method may have limitations, it can provide a clearer explanation of the travel behavior.

Big data has brought the possibility of improving the traffic demand model. On the one hand, big data can enable the model to obtain more reliable input conditions and verification information; on the other hand, big data has given rise to the data-driven model (DDM) [9–10] and data fusion synthetic model [11] that primarily rely on long-term spatiotemporal position data as input. The data-driven model with machine learning as the core has improved the reliability of the status quo traffic demand model to a certain extent, but the interpretability of the travel behavior still needs to be

Received: 2023-11-14

strengthened. Meanwhile, data-driven or fusion models are limited by data quality, such as the granularity, continuity, and reliability of big data, which directly affect the accuracy of the model. In addition, apart from some ticketing or billing system data, there is little data that can fully record the various transportation activities. Therefore, in terms of traffic demand modeling, multi-dimensional data fusion is a necessary task for characterizing the trip chain, which answers the question of why traffic demand models are still needed in the case of big data [12].

From describing the starting and ending points, modes, and times of daily travel for people or vehicles based on small sample data, to describing the spatiotemporal trajectory of activities based on big data, the ability of data to characterize the travel behavior has significantly improved, enabling researchers to more fully extract the characteristics of tour activities and spatiotemporal distribution. A more accurate travel population can be obtained by shifting from modeling the travel of individuals (people or vehicles) in a small proportion to reproducing the complete trajectories of large-sample populations and vehicle activities. These feature data and overall data can serve as input conditions for modeling the travel behavior, and change the traditional modeling methods that use the mathematical optimization method and operational indicators such as the total travel volume and inspection line for parameter calibration and verification. This advantage takes results closer to the real population as a prior condition, examines the possible problem of the conventional traffic demand model theory, and explores the improvement direction. Therefore, relied on the advantages and changes brought by big data, this article explored in depth the possible changes that may be brought about by the corresponding traffic demand modeling method, which not only makes good use of the interpretability of travel activity modeling for the travel behavior, but also leverages the role of big data in improving the model accuracy.

2 From mathematical optimization to causal inference

2.1 Limitations in mathematical optimization models

The mathematical optimization method almost runs through the entire process of traffic demand modeling, such as the gravity model commonly used for travel distribution, the Logit model used for mode split, and the equilibrium model used for traffic allocation. However, mathematical optimization is not causal, and its interpretability for predicting the travel behavior, especially destination selection in travel distribution, is insufficient.

For example, the travel distribution gravity model [13] calibrates the impedance function parameter based on the travel generation/attraction volume, network travel impedance skim matrix, and target travel impedance distribution curve of each traffic zone as input conditions. The convergence criterion is that the travel generation/attraction volume and travel impedance distribution of each traffic zone meet the convergence conditions or reach the maximum iteration number. The formula for calculating the impedance factor is

$$f\left(t_{ij}\right)=a\times t_{ij}^{b}\times \mathrm{e}^{c\times t_{ij}}, \tag{1}$$

where $f(t_{ij})$ is the impedance factor, which is the impedance function between traffic zones i and j, and is usually represented by a function of the impedance index variable t_{ij} calculated based on a network model; a, b and c are model parameters, where a = b = 0 and a = c = 0 correspond to power and exponential functions, respectively. The formula for the travel OD matrix of the gravity model is

$$T_{ij}^{p}=P_{i}^{p}\times\frac{A_{j}^{p}\times f\left(t_{ij}\right)\times K_{ij}}{\sum_{l=1}^{m}A_{l}^{p}\times f\left(t_{il}\right)\times K_{il}}, \tag{2}$$

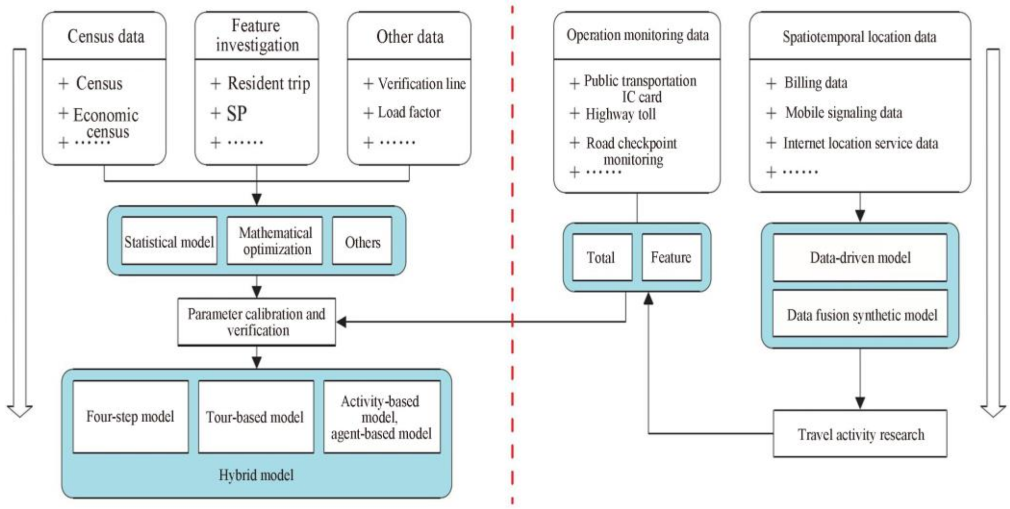

Fig. 1 Development paradigms of travel demand models

where i, j, and l are the traffic zone numbers; m is the total number of traffic zones, and T_{ij}^p is the travel volume generated by traffic zone i with the destination of traffic zone j and objective of p; P_i^p refers to the amount of travel generated in traffic zone i with the objective of p; A_j^p and A_l^p represent the travel attraction volume of traffic zones j and l, respectively, with the purpose of p; K_{ij} and K_{il} are optional adjustment coefficients, also known as K coefficient, used to represent the influence of variables other than travel impedance.

It can be seen from the model expression that in addition to parameters a, b, and c, K coefficient also has a significant impact on the model results. More importantly, even if the parameter calibration results are very ideal, whether the model results can reflect the real situation of urban traffic still needs to be verified. Empirical studies have shown [3, 14–15] that it is almost impossible to obtain the high-quality travel distribution matrix through the conventional mathematical optimization method, but big data provides the possibility for verification.

2.2 Possibility of causal inference

Based on long-term massive spatiotemporal location data, urban traffic operation monitoring data, resident travel survey feature data, socio-economic data, and land use status and planning data, Ref. [3] constructed a travel activity model framework with the pre-positioning demand attribute from the perspective of the stability of traveler activities (see Fig. 2). The core idea of this framework is to use big data to explore the characteristics and pattern of the real urban travel activity as much as possible, and use personal attribute changes to achieve travel inference and prediction in the near and medium term based on replicating the current urban travel situation. In other words, the personal attribute and urban activity pattern are the basic prerequisites for causal inference and correlation analysis. Big data is used to explore the evolution process of the individual and group activity pattern as much as possible, making it possible to transform travel behavior modeling from mathematical optimization to causal inference. The specific steps of the model include:

1) The total data is inferred using population census, economic census, and location-based services (LBS) data, and the household member characteristic data is obtained through resident travel survey and population census. Population synthesis research is conducted based on total data and household member characteristic data to construct population distribution that conforms to individual and overall characteristics.

2) Based on mobile signaling or LBS data, the occupational-residential OD matrix is inferred to allocate the work locations to the employed population in each traffic zone, and the individual spatial location attributes are assigned.

3) Mobile signaling or LBS data is used to distinguish four main travel destinations of residential, work, daily life, and others to construct an activity matrix to discover the stability characteristics of the individual travel activity [16].

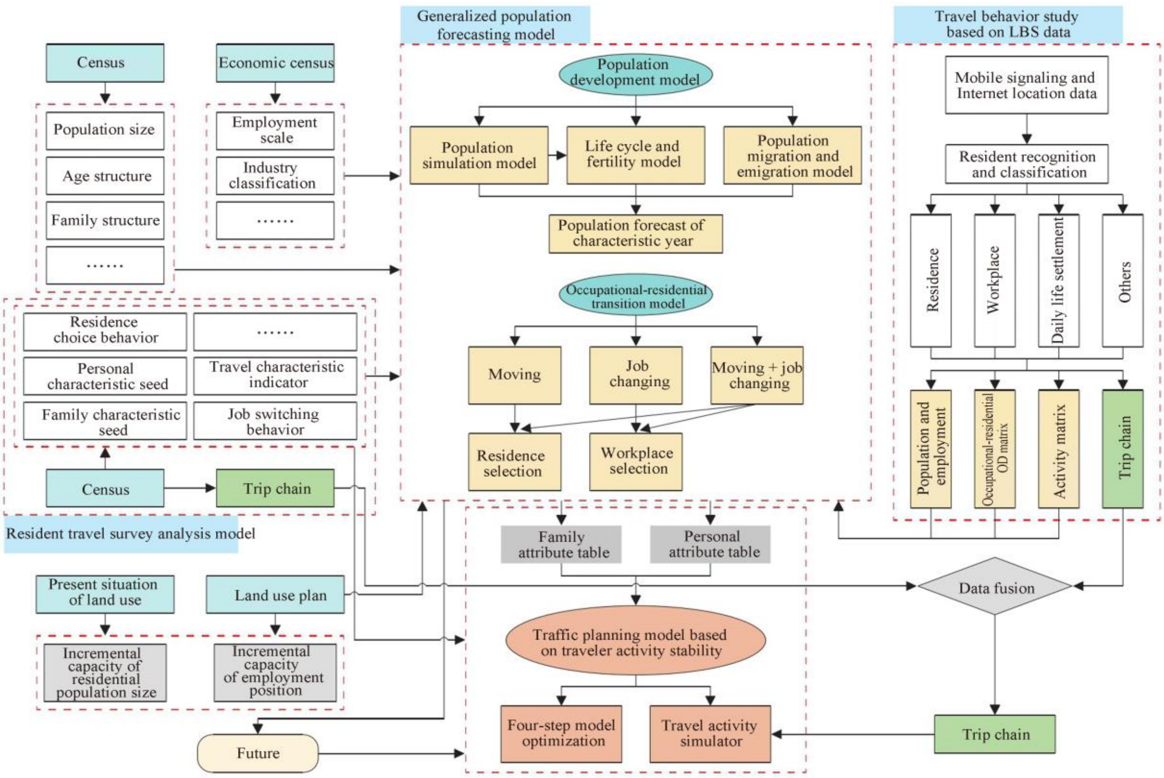

Fig. 2 Framework for travel activity models based on demand attribute preposition

4) Due to the fact that one's place of residence and workplace determine the starting and ending points of commuting, causal inference of commuting activities can be carried out using individual occupational and residential attributes. Correlation analysis of accidental travel activity prediction is carried out using the stability of the group activity pattern, that is, destinations for travel in the same region and with the same purpose have similarities, thus forming the stability of the group travel activity.

5) Finally, based on changes in land use, the spatiotemporal characteristics of the individual attribute are deduced to further predict future travel demand.

China's mega cities have entered a stage of stock renewal and development, where the individual choice and spatiotemporal activity pattern of existing residents have a dominant impact on the future activity pattern. Therefore, the real situation of a city is fully explored and development deduction is conducted to improve the accuracy of the traffic demand model, which is of strategic significance for the next stage of China's urban transportation governance and refined management. Meanwhile, data availability, big data modeling method, and experimental research have shown that the travel activity model framework with the pre-positioning demand attribute has strong feasibility.

3 From proportional factors to proportional sampling

The concept of "real number solution", which contradicts the physical world, almost runs through the entire process of solving the traffic demand model. In the conventional traffic demand model method, the travel generation rate, demand spatial distribution, traffic mode selection, and traffic volume allocation are all real number solutions. However, in the real physical world, there will not be 0.5 trips or 0.5 vehicles, which means that the processing method of a model is limited from the perspective of the physical world. Another problem arising from the real number solution is the fragmentation of traffic demand [14], which poses new obstacles for subsequent microscopic traffic simulation. How to aggregate vehicles with decimal places into integers is one of the difficulties in microscopic traffic simulation.

3.1 Probability problem of mode split

The ABM modeling method can achieve the integer solution for travel generation and distribution, but the results obtained in mode split and traffic allocation are still real solutions. Taking mode split as an example, the simplest multinormal logit model [17] (MNL) is adopted, and the probability of selecting traffic mode m is

$$P_m = \frac{e^{U_m}}{\sum_{i=1}^{n} e^{U_i}}, \tag{3}$$

where U_m and U_i are the utility function of traffic modes m and i, respectively.

The traffic mode selection model in the classic textbook [13] defines P_m as the choice probabilities of a traffic mode. In the actual model operation process, P_m is multiplied by the total traffic demand to obtain the traffic volume of each traffic mode, which transforms the concept of probability into proportion.

Based on the definitions of probability in three dictionaries, Webster, Oxford, and Longman, it can be concluded that the original meaning of probability is the possibility of occurrence. In the process of choosing a travel mode, even if the probability of a certain travel mode is only 1%, if the traveler chooses that mode, the result will be 1, and of course, it is more likely to be 0. But in the logic of proportional calculation, if the probability is 1%, the result can only be 0.01. In the conventional model operation process, probability is processed proportionally to solve the stability problem of model solution, but it also brings about the contradiction between the non-integer solution of the model and the real physical world.

3.2 Probability problem of traffic allocation

The contradiction of proportional or probabilistic problems also arises in the traffic allocation algorithm. For example, the famous Dial allocation algorithm [18–19] calculates the probability of path selection as a proportion, resulting in the fragmented path traffic demand that contradicts the real physical world. Lin et al. [20] proposed a multiple vehicle traffic allocation method based on the integer solution, which solves the problem of fragmented traffic volume. The theoretical basis of balanced traffic allocation is utility maximization, which refers to rational choice (assumption of economic man), and considers the network equilibrium constraint on this basis. The irrational behavior is common in reality, and the transportation system is more likely to be in an unbalanced state, so the equilibrium state pursued by equilibrium allocation lacks basis in reality. T. de la Barra [21] used the enumeration path method for traffic allocation in the integrated transportation and land use software TRANUS. This method reduces the computational load to a certain extent, but still treats probability as a proportion for calculation, and the results obtained are still real number solutions.

The path-based utilization probability distribution and the application of the proportional sampling method for traffic allocation is a new technological path worth exploring and trying. For example, there are three paths between a certain OD pair, and the choice probability of each path based on utility calculation is 0.781 9, 0.200 6, and 0.017 5, respectively. Tab. 1 shows the results of 20 traffic allocations based on the proportional sampling method under three travel volume scenarios. The results show that the volatility of traffic allocation results is greater when the travel volume is small, and as the travel volume increases, the allocation results tend to approach the input probability distribution. However, in reality, the travel demand between fine-grained traffic zones is usually small. If it is specific to an individual, then the travel demand is in the unit of 1, and traffic allocation based on the proportional sampling method is difficult to achieve the results solved by the proportional algorithm. In addition, the uncertainty brought by the proportional sampling calculation of small-scale demand can lead to unstable (non-reproducible) model results. In contrast, the general traffic allocation algorithm mainly uses equilibrium as the convergence criterion, and whether it conforms to the real traffic operating conditions still needs further verification. Therefore, the transition from the proportional factors to proportional sampling remains a modeling challenge and requires further research.

4 From holistic reconstruction to incremental models

4.1 Necessity of incremental models

In addition to the data-driven model with machine learning as the core, the existing traffic demand modeling method is an holistic construction method mainly based on feature seeds and constraint optimization, that is, the feature seed obtained from current situation survey is used and the overall indicator of transportation operation is taken as the constraint condition, and the mathematical optimization method is used to construct the aggregated or non-aggregated overall travel activity dataset. The characteristic of this method is that all participants enter each selection process, which means that every iteration of the model calculation, i.e., the spatiotemporal distribution of all travel activities and traffic mode selection will be recalculated. However, the real urban system does not operate according to this logic, and travelers have stable attributes and activities.

The stability of the attribute is mainly reflected in the fact that citizens have relatively stable places of residence and work. Fig. 3 shows the distribution of current residence duration and current employment duration of 10 000 householders in Guangzhou in 2019. The proportion of population with medium to long fixed residence is relatively high, and the older the age, the higher and more stable the proportion of self-owned housing. For the employed population, the proportion of those who have not changed jobs for more than 10 years is 27.4%, and that of those who have been employed in their current positions for more than 5 years is as high as 56.8%. In addition, from the perspective of the willingness to change the jobs (see Fig. 4), the shorter the employment duration of the current position, the stronger the willingness, and the group with 3 years of work experience has the strongest willingness. As the duration of continuous employment in the same position increases, the willingness to change jobs gradually decreases. The proportion of people who have been working in their current position for 10–20 years and have a willingness to change jobs is 4.6%, and that of people who have been working for more than 20 years and have a willingness to change jobs is 1.2%, indicating that the employed group has a high degree of stability. The stability of the residential and work place, as well as the willingness to migrate, indicate that urban development is a gradual and iterative evolution. Constructing an incremental model of local changes is more in line with the laws of reality than a holistic reconstruction model that optimizes every calculation.

Tab. 1 Illustration of transportation distribution results based on probabilistic sampling methods

No.	Travel amount of 3			Travel amount of 10			Travel amount of 100		
	Path A	Path B	Path C	Path A	Path B	Path C	Path A	Path B	Path C
1	3	0	0	8	2	0	83	16	1
2	3	0	0	6	4	0	78	21	1
3	1	2	0	7	3	0	74	25	1
4	2	1	0	9	1	0	82	18	0
5	2	1	0	10	0	0	75	20	5
6	2	1	0	9	0	1	77	22	1
7	1	2	0	8	2	0	83	15	2
8	2	1	0	8	2	0	78	20	2
9	3	0	0	9	1	0	78	22	0
10	2	1	0	9	1	0	76	23	1
11	3	0	0	8	2	0	80	19	1
12	3	0	0	6	4	0	74	22	4
13	2	1	0	9	1	0	71	25	4
14	3	0	0	10	0	0	77	21	2
15	3	0	0	5	4	1	76	20	4
16	3	0	0	7	3	0	73	24	3
17	3	0	0	7	3	0	75	22	3
18	2	0	1	8	2	0	74	26	0
19	3	0	0	9	1	0	74	25	1
20	3	0	0	9	1	0	81	17	2

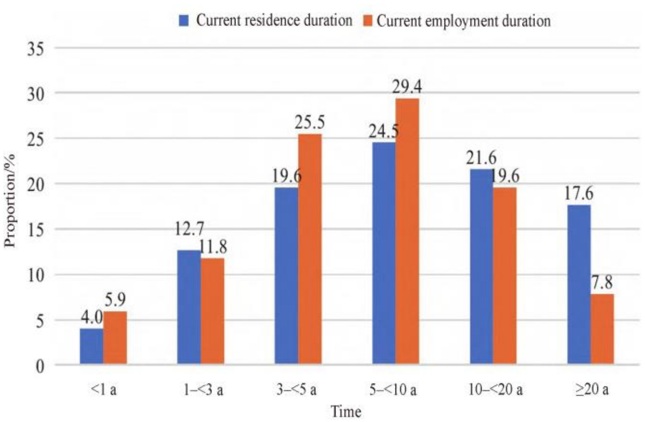

Fig. 3 Distribution of current residence duration and current employment duration

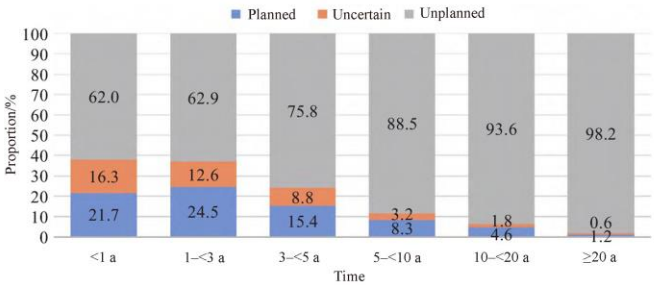

Fig. 4 Distribution of job change intentions among groups with different years of employment

4.2 Development and future prospects of incremental model

The incremental model is not a new concept, and Manheim [22] proposed the issue of prediction based on a given fixed base point. Abraham et al. [23] proposed an improved four-stage model based on increment. Bates et al. [24] proposed a nested incremental logit model based on Kumar's early work, and mathematically demonstrated how to apply this model to predict changes in mode selection. The *Transport Analysis Guidance* [25] released by the UK Department for Transport describes the incremental model. The appendix document of the Guidance explains the implementation method of various models such as the multivariate logit model, double constrained gravity model, composite utility model, and conditional probability model that consider the incremental factor. Overall, the incremental model is a change model based on the current situation, with the core of constructing a high-quality current situation as a reference frame. However, most studies and methods for constructing the current situation are mainly based on the holistic reconstruction of features, and it is difficult to achieve high-quality modeling of real urban travel based on the mathematical optimization method. Therefore, constructing the overall real urban travel has become the key to an incremental model.

The key technology of the transportation planning model based on the stability of the traveler activity [16] relies on mobile signaling data to mine the traveler attribute and construct the overall urban travel. Combined with the extended population synthesis model, it can achieve the prediction of urban population and employment positions from the current situation to the future, and conduct empirical research using the activity-based model. The results show that it achieves a high-quality replica of the current urban travel activity, and relying on changes in the traveler attribute to infer realizes the inheritance and continuation of traffic demand forecasting in time and space. It indicates that the incremental model structure can be further extended to the activity-based model.

Chinese cities have entered a stage of stock development, with more stable urban activities. High-quality mining of the current situation determines to some extent the accuracy of the short-term prediction and the rationality of medium- and long-term decision-making. In addition, high coverage mobile signaling, Internet LBS, and road checkpoint monitoring data, as well as high-quality traffic operation monitoring data (such as public transport IC card, taxi GPS, online car hailing order, and track data) have been widely used in urban mobility modeling and have gained rich experience. Based on these data foundations and technical methods, the higher-quality benchmark feature year model can be constructed. It can be foreseen that relying on high-quality replica of the current urban travel activity to construct an incremental model for prediction and deduction will provide a more accurate quantitative analysis platform for transportation planning and governance at the level of operational management.

5 From finite constraint convergence to prior empirical justification

The traffic demand model has a strong dependence on the mathematical optimization solution, and the constraints for the optimization solution are generally known traffic operation statistics and monitoring data or the overall distribution inferred through sampling survey. Before the arrival of the big data era, it is often unknown whether the constraints are complete or sufficient, which makes the optimal solution that meets the constraints not necessarily the real solution required for traffic demand modeling. The fundamental transformation brought about by the era of big data is the possibility to fully grasp the real operating status indicators, such as accurate travel OD and the activity trajectory, providing a better perspective for better understanding the reliability of the mathematical optimization solution based on the finite constraint condition, for it can better reflect on and examine the problem of the existing modeling method and find ways to improve it.

The data-driven analysis method [14–15] has demonstrated the limitations of the mathematical optimization method based on the finite constraint condition, which is an empirical study from a prior perspective, that is, to judge the problem of existing methods from the real and overall perspective. Under the premise of a given population, a simulation study on the sampling and expansion problem of resident travel survey from a prior perspective [26] used the simulation method to analyze the seed characteristics and population differences of resident travel survey sampling with different sampling rates, and further verified the expansion results with the real population. It found that the method of constructing and restoring the population based on seed characteristics using the multi-attribute weighted expansion method has obvious limitations and may even introduce new biases. Multi-attribute weighted sampling is a classic method in the era of small data, and resident travel survey sampling is a typical representative of solving finite constraint conditions.

Empirical evidence shows that there can be new perspectives in the era of big data. Ref. [8] conducted empirical research on large and small transportation networks using the simulation experiment and found that the correlation analysis between the common road observation point survey value and simulation value, and GEH [4] value test cannot guarantee that matrix estimation can restore the true population. The study further indicates that the quality of the initial input matrix is the key to the usability of the matrix estimation technique. The initial matrix constructed based on the gravity model is taken as input, even in the case where the goodness of fit between the observed and simulated values of the road segment reaches 1, the estimated OD matrix obtained may be far from the true OD matrix. Strictly speaking, when the constructed OD matrix based on finite constraint conditions is taken as the input condition, the matrix estimation technique is almost ineffective, and in reality, the matrix estimation technique is the most common method in the field of traffic demand forecasting. These studies indicate that using a prior perspective can effectively identify key technical issues in existing traffic demand models, and demonstrate the importance of fully tapping into current urban travel activities.

The problems identified in the above research reveal a fundamental fact that the convergence or stability of the solution is only the optimization of finite constraint conditions and cannot represent objective truth (the real physical world). The increasingly mature technology for collecting and analyzing long-term and large-scale spatiotemporal location data allows us to obtain current results that are closer to the true values, and provides a prior perspective for the validation of the traffic demand model. Of course, the examination and verification of prior perspectives can be extended to a wider range of fields, such as the restoration of full OD from large-scale trajectory data [27] and the verification of traffic allocation based on GPS trajectory path data [28].

6 Summary and discussion

Traffic demand modeling is a core topic in transportation science research. The big data era has brought revolutionary changes to the development of the traffic demand model, and raised the need for further updates and innovations to improve the explanatory power of the real development of the transportation system and the guidance ability for future development. As proposed by Boyce et al. [29], "Later generations should focus on criticism of the nature and usage of the models, technical and political challenges, and the underlying principles of prediction. Innovations in the method stem from the problem-oriented response, a desire for the higher precision model, or the discovery of new ideas. It is essential to doubt the traditional explanation, the credibility of the model and its predictive assumption, and the inherent uncertainty of the prediction, and it is meaningful to understand when, where, and why to propose modifications to a theory or method." In addition, the irrational choice behavior [30] and uncertainty [31] can also have a significant impact on the model results. Although some argue that a model that can correctly respond to variable trends is a good model, the model accuracy is the only and highest standard for testing model quality in terms of the practical application and scientific research of traffic management. Therefore, based on the understanding of the development and practical work of the traffic demand model, this article proposed paradigm shifts from mathematical optimization to causal inference, from the proportional factors to probabilistic sampling, from holistic reconstruction to the incremental model, and from finite constraint convergence to prior empirical justification. Examining the technological paradigm is not to negate previous technical approaches, but to improve and develop them based on understanding and discussions.

References

[1] CARROLL J D. Detroit metropolitan area traffic study part II: future traffic and a long range expressway plan[R]. Detroit: Michigan State Highway Department, 1956.
[2] CARROLL J D. Chicago area transportation study final report volume III: transportation plan[R]. Chicago: Chicago Area Transportation Study, 1962.
[3] CHEN X L, ZHANG H, CHEN X H, et al. Urban activity models based on spatiotemporal inference[J]. Urban transport of China, 2023, 21(1): 60–68. (in Chinese)
[4] CHEN X L. A review on urban transportation models in China[J]. Urban Transport of China, 2016, 14(2): 17–21. (in Chinese)
[5] MAYER-SCHÖNBERGER V, CUKIER K. Big data: a revolution that will transform how we live, work, and think[M]. ZHOU T, et al, translated. Zhejiang: Zhejiang People's Publishing House, 2012. (in Chinese)
[6] GILPIN L H, BAU D, YUAN B Z, et al. Explaining explanations: an overview of interpretability of machine learning[C]//IEEE.2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). US: IEEE, 2018: 80–89.
[7] LINARDATOS P, PAPASTEFANOPOULOS V, KOTSIANTIS S. Explainable AI: a review of machine learning interpretability methods[J]. Entropy, 2020, 23(1): 18.
[8] CHEN X L, ZHANG H, MA Y L, et al. Can matrix estimation reveal the actual OD matrix effectively? empirical evidence based on simulation experiments[J]. Journal of transportation engineering, 2023, 23(6): 67–78. (in Chinese)
[9] ANDA C, ERATH A, FOURIE P J. Transport modelling in the age of big data[J]. International journal of urban sciences, 2017, 21(s1): 19–42.
[10] BREDERODE L, POTS M, FRANSEN R, et al. Big Data fusion and parametrization for strategic transport demand models[C]//IEEE. 2019 6th international conference on models and technologies for Intelligent Transportation Systems (MT–ITS). USA: IEEE,2019: 1–8.
[11] ANDA C, MEDINA S A O, AXHAUSEN K W. Synthesising digital twin travellers: individual travel demand from aggregated mobile phone data[J]. Transportation research part C: emerging technologies, 2021, 128: 103118.
[12] MALLESON N, BIRKIN M, BIRKS D, et al. Agent-based modelling for urban analytics: state of the art and challenges[J]. AI communications, 2022, 35(4): 393–406.
[13] DE DIOS ORTÚZAR J, WILLUMSEN L G. Modelling transport[M]. Chichester: John Wiley & Sons, Ltd, 2010.
[14] SHENG Z Q, WU Z X. Trip spatial distribution model evaluation and improvement based on commuting big data[J]. Urban transport of China, 2023, 21(1): 69–73. (in Chinese)

[15] GUO J F, LI X, BAI S G, et al. A deep learning based urban travel pattern library for universal scene[J]. Urban transport of China, 2023, 21(1): 74–85. (in Chinese)
[16] CHEN X L. Key Technologies of transportation planning model based on stability of traveler's activity[D]. Shanghai: Tongji University, 2023. (in Chinese)
[17] BEN-AKIVA M, LERMAN S R. Discrete choice analysis: theory and application to travel demand[M]. Cambridge, Massachusetts: MIT Press, 1985: 390.
[18] DIAL R B. A probabilistic multipath traffic assignment model which obviates path enumeration[J]. Transportation research, 1971, 5(2): 83–111.
[19] FLORIAN M, FOX B. On the probabilistic origin of dial's multipath traffic assignment model[J]. Transportation research, 1976, 10(5): 339–341.
[20] LIN X, LECLERCQ L, TAVASSZY L, et al. Formulating multi-class user equilibrium using mixed-integer linear programming[J]. EURO journal on transportation and logistics, 2022, 11: 100097.
[21] DE LA BARRA T. Integrated land use and transport modelling: decision chains and hierarchies[M]. Cambridge: Cambridge University Press, 1989.
[22] MANHEIM M L. Fundamentals of transportation systems analysis. Volume 1: basic concepts[M]. Cambridge, Massachusetts: M. I. T. Press, 1979.
[23] ABRAHAM H, SHAW N, WILLUMSEN L G. A micro-based incremental four stage transportation model for London[C]//PTRC Education and Research Services Limited. Proceedings of seminar E held at the PTRC European transport, highways and planning20th summer annual meeting. London: PTRC Education and Research Services Limited, 1992: 261–273.
[24] BATES J, ASHLEY D, HYMAN G. The nested incremental logit model: theory and application to modal choice-transportation planning methods[C]//PTRC Education and Research Services Limited. Proceedings of seminar C held at the PTRC summer annual meeting, bath university England. London: PTRC Education and Research Services Limited, 1987: 75–97.
[25] Department for Transport. Transport analysis guidance (TAG)[R]. London: Department for Transport, 2022.
[26] CHEN X L, MA Y L, SONG C, et al. A simulation study on the sampling and expansion problems of resident travel surveys from a prior perspective[C]//Urban Transportation Planning Professional Committee of the Chinese Urban Planning Society. Resilient Transportation Quality and Service: Abstract of the paper from the 2023 Annual Conference on China Urban Transportation Planning. Beijing: China Architecture & Building Press, 2023: 15–16. (in Chinese)
[27] LIN P Q, YU Z, WANG Y Z, et al. Method, device, and medium for estimating road traffic density based on vehicle trajectory data: CN202211018930.0[P]. 2022-12-09. (in Chinese)
[28] SONG X Y, YANG Z, WANG T, et al. Dynamic traffic assignment model based on GPS data and Point of Interest (POI) in Shanghai[J]. Sensors, 2021, 21(21): 7341.
[29] BOYCE D, WILIAMS H. Forecasting urban travel: past, present and future[M]. Northampton, Massachusetts: Edward Elgar Publishing, 2015.
[30] VASCONCELOS M, MONTEIRO T, KACELNIK A. Irrational choice and the value of information[R/OL]. (2015-09-09) [2023-02-05]. https: //www.nature.com/articles/srep13874.pdf.
[31] RASOULI S, TIMMERMANS H. Uncertainty in travel demand forecasting models: literature review and research agenda[J]. Transportation letters, 2012, 4(1): 55–73.

