Application of Web-based Data Mining in Road Network Evaluation in the PRC

Zilong Wang
PhD student, Rutgers University
EARD Intern, ADB

Supervisors:
Dr. Robert Guild, EARD
Dr. Chen Chen, SARD

Date: 22 September 2016

Outline

1. Introduction of Road Network
2. Summary of Provincial Road Network in the PRC
3. Effect of Road Network on Intercity Transport
4. Public Interest on Transportation and Development of Provincial Road Network
5. Influence of Transportation Research on Provincial Road Network
6. Conclusion and Recommendation

Transportation Infrastructure in the PRC

Development

- Construction of national expressway system in the PRC was initiated in 1984.
- Construct 3000 km of expressway annually.
- The PRC has the largest interstate expressway system in the world.

Influence

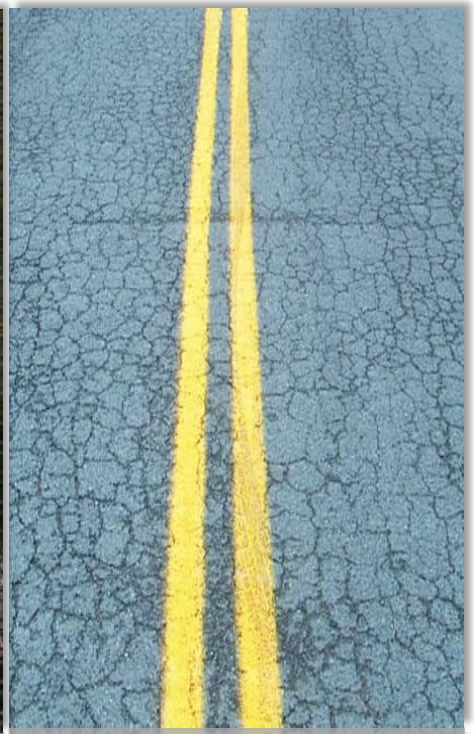
- Reduce trade costs.
- Increase market size in different regions.
- Promote the mobility of capital and labor.

Transport infrastructure is a key driver for economic growth and regional development during the twenty-year period of rapid growth in the PRC.

Pavement Distress

Longitudinal cracking

Transverse cracking



Fatigue cracking

Rutting

- Pavement ages as a result of repeated loading and environment.
- A number of structural and functional distress symptoms occur.
- Affect driving experience and the safety of travelers and freight vehicles.

Better Road vs More Road

US constructed interstate highway system since 1956

- By the 1970s, the majority of highways started showing ostensible signs of wear.
- Due to the increasing heavy traffic load, many highways showed sign of failure before reaching design life.

After the mid-1980s, the focus of highway agencies has shifted from expanding the system to preserving highways.

PRC constructed national expressway system 30 years after 1956

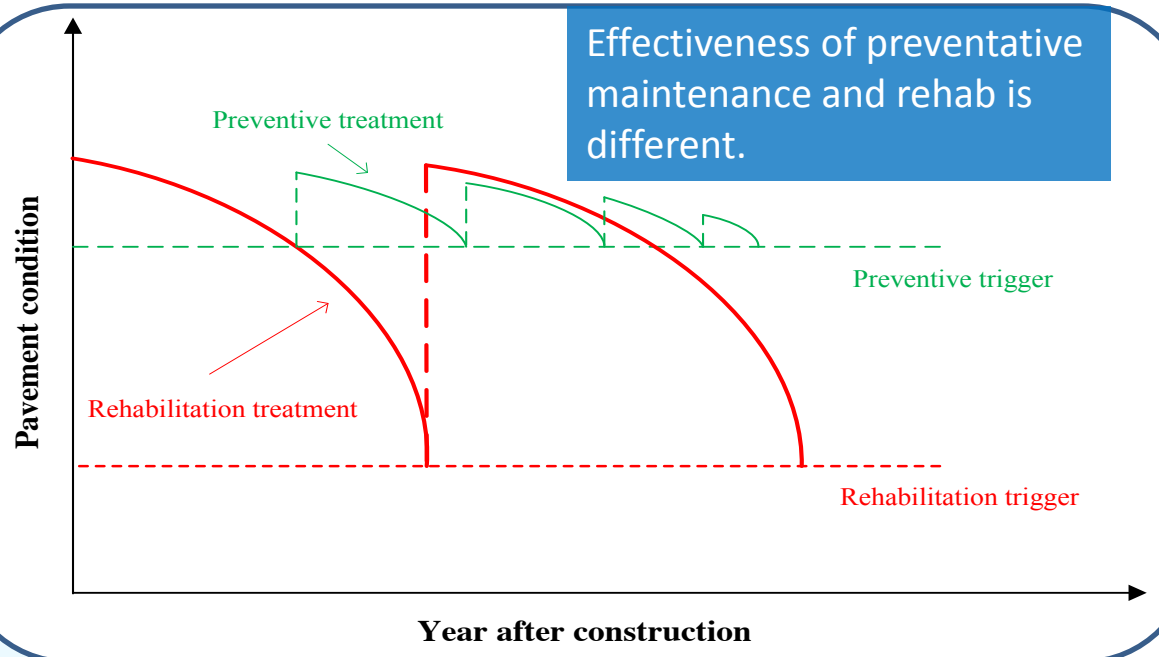
- Prevalence of overweight trucks.
- Poor material quality.
- Poor quality control during pavement construction.
- Highways undergo more damage than it was supposed to absorb and deteriorate rapidly.

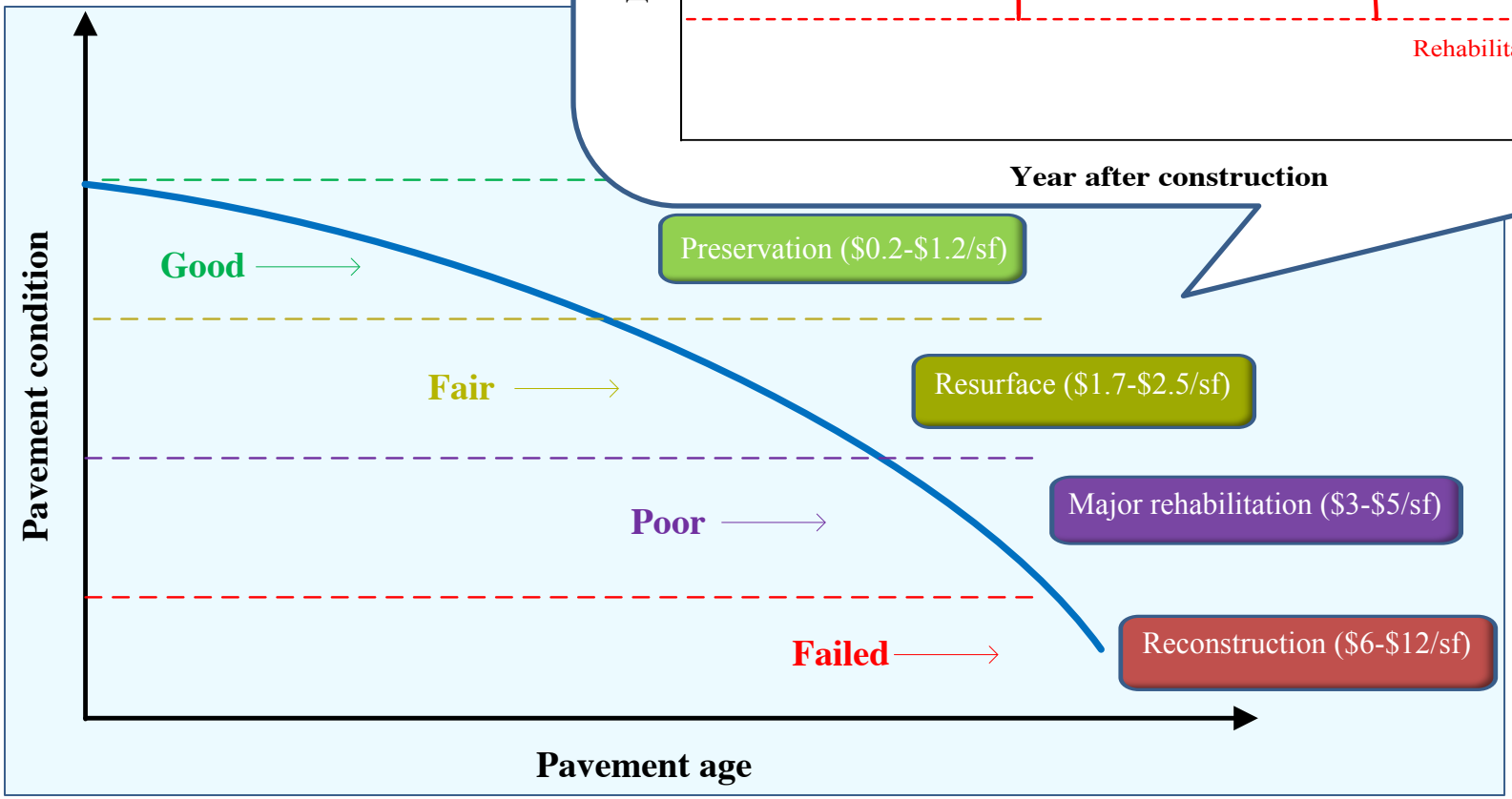
Still focuses on the major building and expansion of national expressway system.

Pavement Management under Limited Budgets

Proper use of maintaining strategy can extend road life and maximize limited budgets.

Effectiveness of preventative maintenance and rehab is different.





Big Data in Transportation

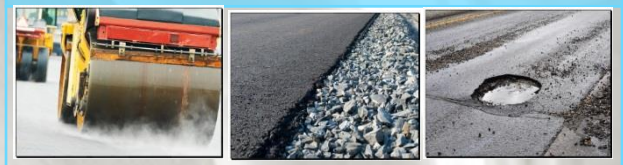
Material

- Material type
- Air voids
- Asphalt contents
- Gradation
- Modulus
- Material application rate.

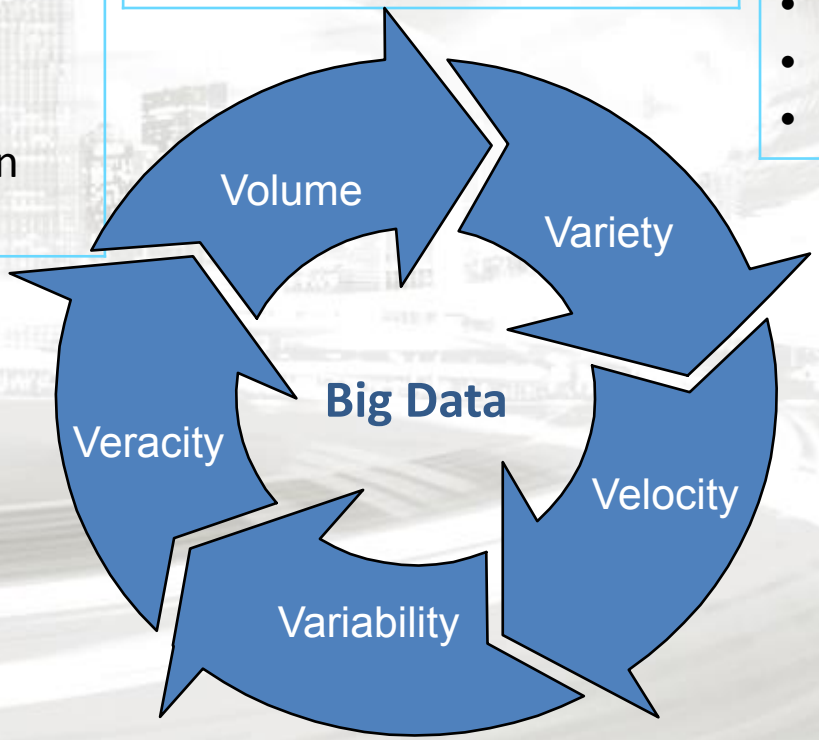
Traffic

- Twitter messages
- Geotagged photos
- mobile phone usage
- Road accident reports

Pavement



- Location
- Cost
- Thickness
- Cracking
- Temperature



Better Understanding of Provincial Road Network Can Prioritize the Investment

Objectives

- Investigate transport network condition for the 31 administrative divisions in the PRC.
- Identify the factors that indicate or affect provincial road network.

Methodology: data mining

- Web-based mining: discovering patterns stored on the Internet.
- Search interesting patterns from vast amount of data.
- Typical techniques include regression analysis, classification, and clustering.

Internet resources

Baidu 指数

gaode.com

Road Network in the PRC

31 administrative divisions

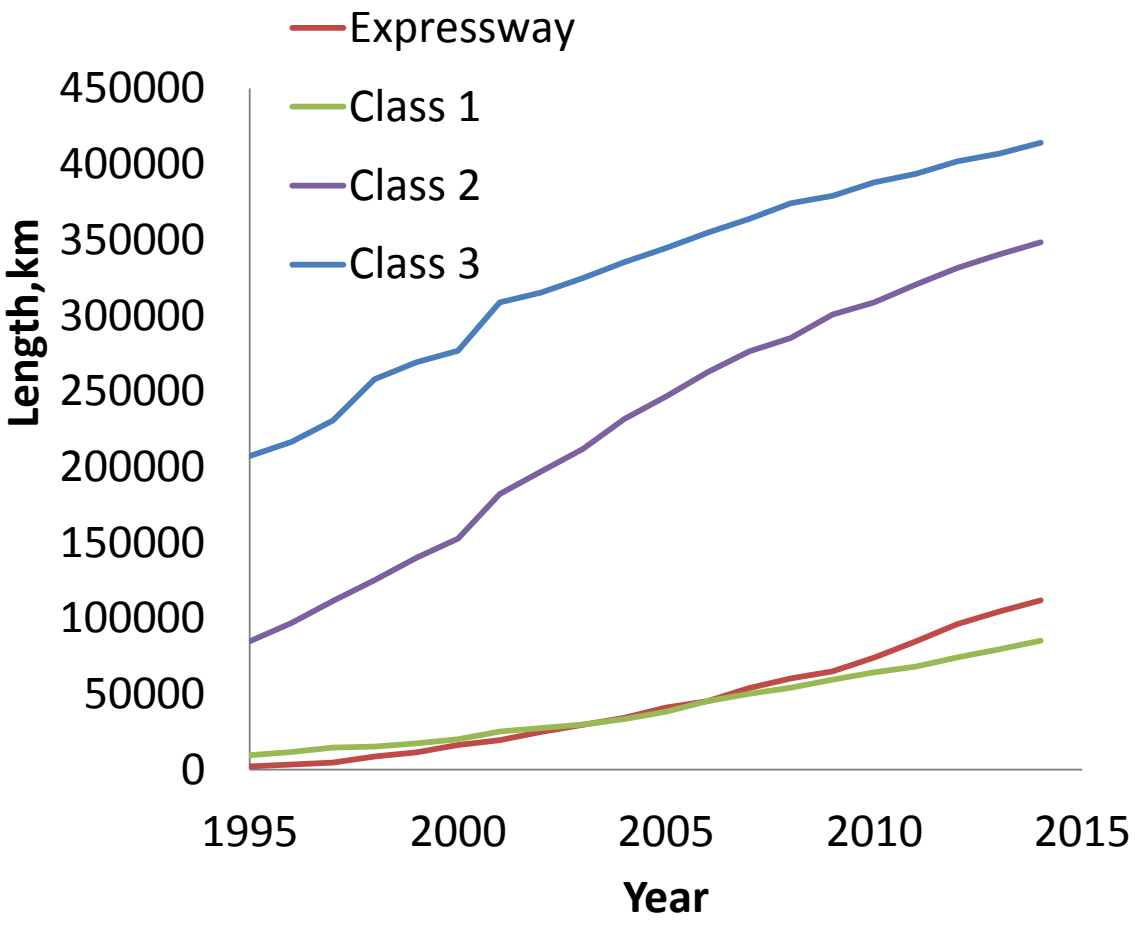
4 municipalities:
Beijing, Shanghai, Tianjin, and Chongqing.

5 autonomous regions:
Inner Mongolia, Guangxi, Ningxia, Tibet, and Xinjiang.

22 provinces.

Road type	Average annual daily traffic (AADT)
Expressway	>55000
Class 1 roads	15000–55000
Class 2 roads	5000–15000
Class 3 roads	2000–5000
Class 4 roads	400–2000

Road length in the PRC



(Source: CEIC, a Euromoney Institutional Investor Company)

2015 Road Performance Inspection

Background

From Sep. to Dec. in 2015, Ministry of Transport (MOT) examined maintenance practice and graded road maintenance performance in every province after the execution of 12th Fiscal Year Plan.

Task	Score	Road	Distress type	Data source	Inspection sample
Road network performance	**65%**	Expressway	Roughness, Distress ratio, Rutting	Inspection data	**>25%**
		Non-expressway	Roughness, Distress ratio	**60%** historical data+**40%** inspection data	**>5%**
Agency management capability	**35%**	**Focus**: road administration, road maintenance, road policy execution, toll management, and technical support.			

Summary of 21 Administrative Divisions

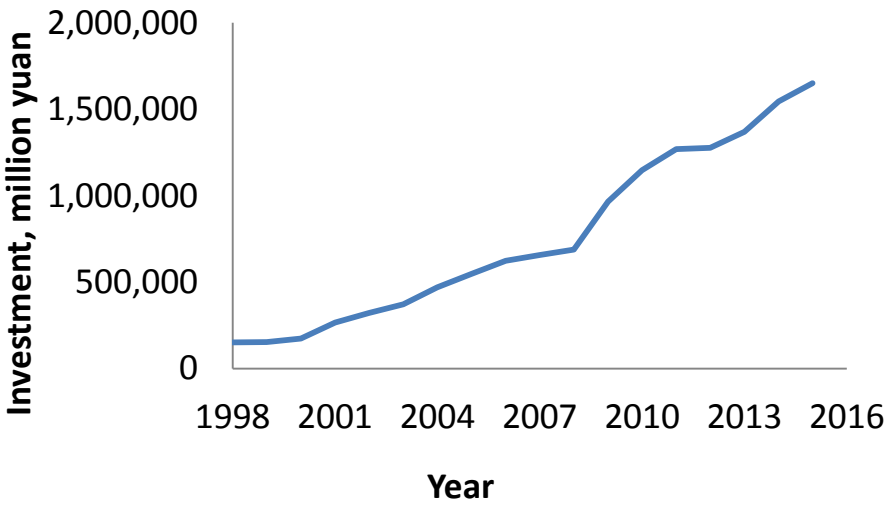
Road ranking	Division	GDP/capita (2014), yuan	Population (2014)	Land area, km²	Road length (2014)
1	Jiangsu	81,874	79,498	102,658	149,845
2	Beijing	99,995	21,332	16,411	21,816
3	Shanghai	97,343	24,204	6,219	12,945
4	Liaoning	65,198	43,907	148,000	100,854
5	Shandong	60,879	97,614	157,126	258,442
6	Fujian	63,472	37,900	120,000	82,907
7	Tianjin	105,202	14,945	11,920	16,110
8	Chongqing	47,859	29,807	82,400	98,680
9	Anhui	34,425	60,564	139,427	169,639
10	Zhejiang	72,967	55,030	104,141	113,730
11	Guangdong	63,452	106,840	184,800	197,131
12	Hunan	40,287	67,139	211,829	211,279
13	Shaanxi	46,928	37,696	205,800	151,189
14	Jiangxi	34,660	45,322	166,900	128,262
15	Hebei	39,984	73,582	190,000	172,891
16	Henan	34,808	100,377	167,000	197,624
17	Hubei	47,124	58,075	187,400	224,184
18	Sichuan	35,128	81,236	488,000	257,027
19	Jilin	50,162	27,518	187,000	88,667
20	Shanxi	35,064	36,389	156,000	137,094
21	Inner Mongolia	71,044	25,012	1,183,000	160,123

Economic and geographic factors may have relation to road performance.

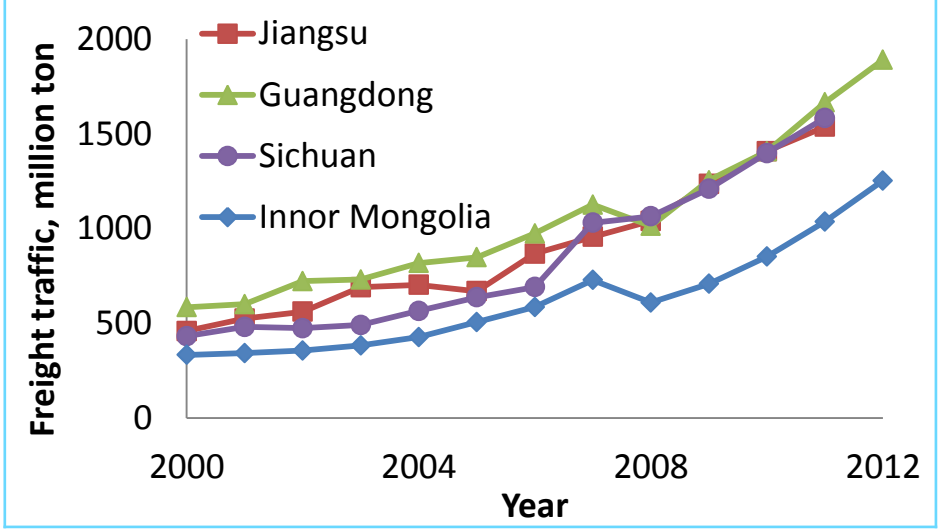
Source: National Bureau of Statistics of China

Development of Transport Infrastructure in Different Provinces

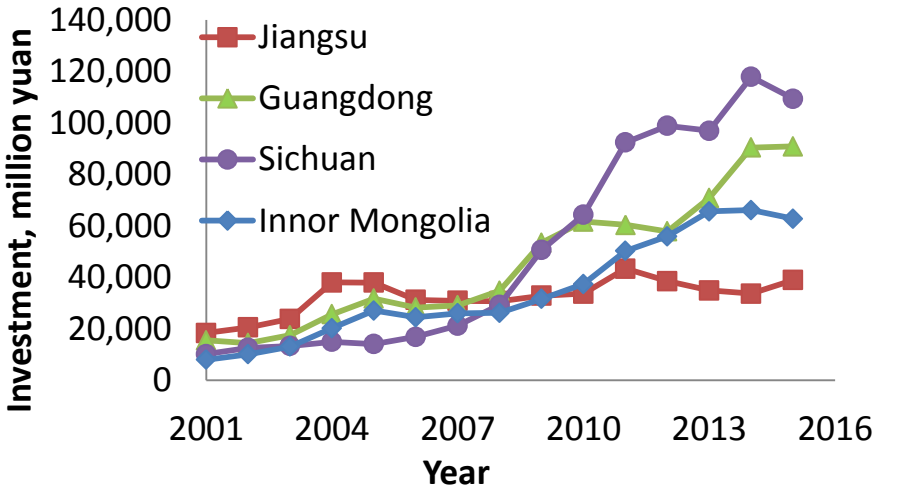
Highway investment in the PRC



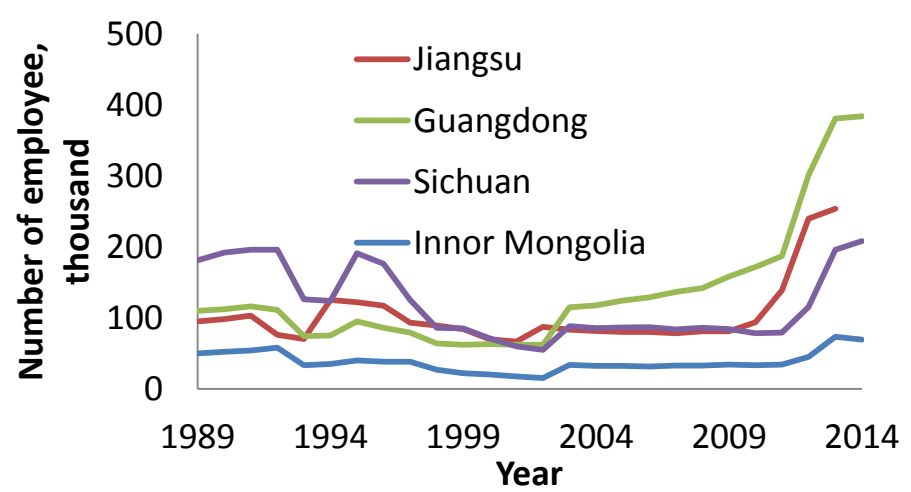
Freight transport in different provinces



Highway investment in 4 provinces



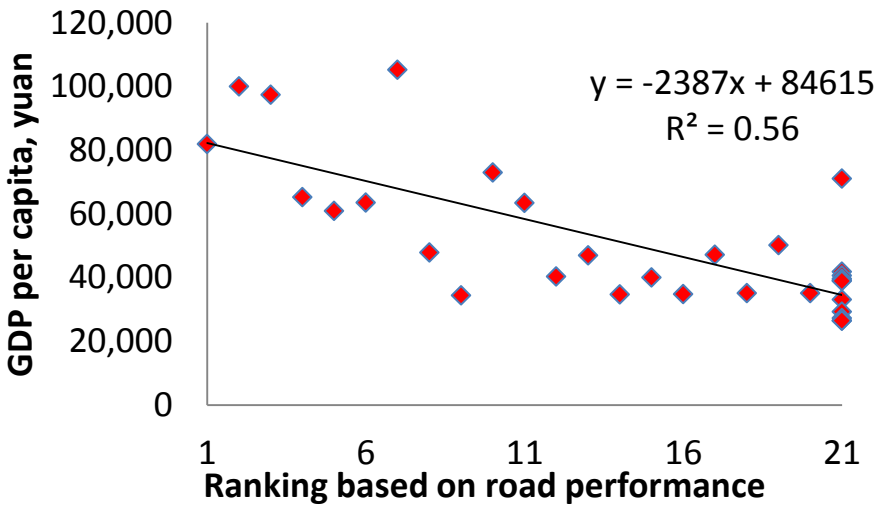
No. of employee in highway industry



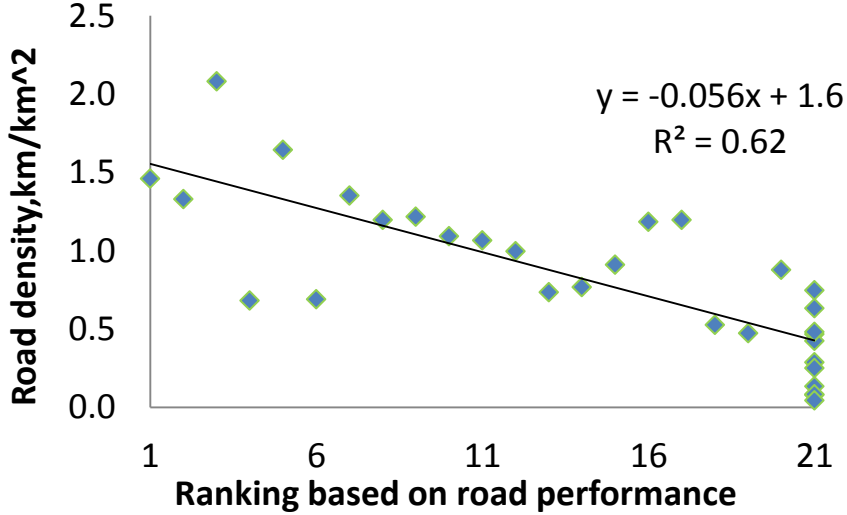
(Source: CEIC)

Correlation between Provincial Road Network Ranking and other Factors

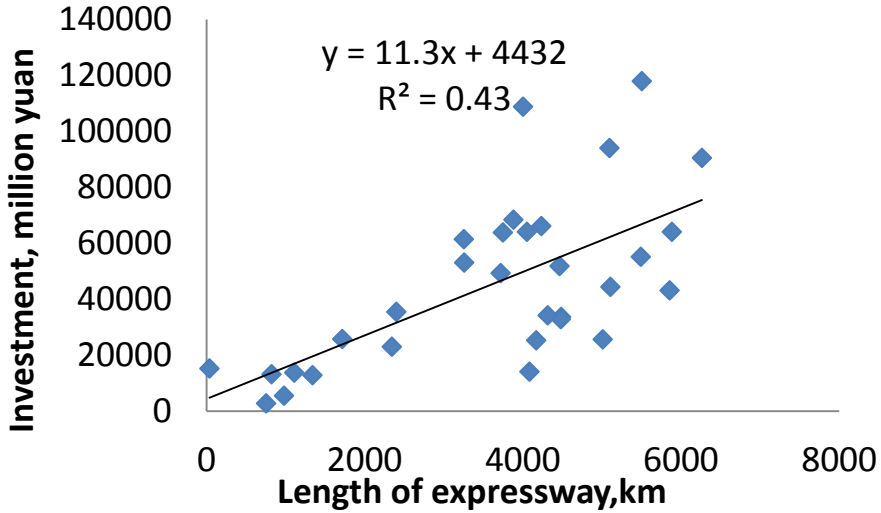
GDP/capita vs road performance



Road density vs road performance



Expressway vs investment

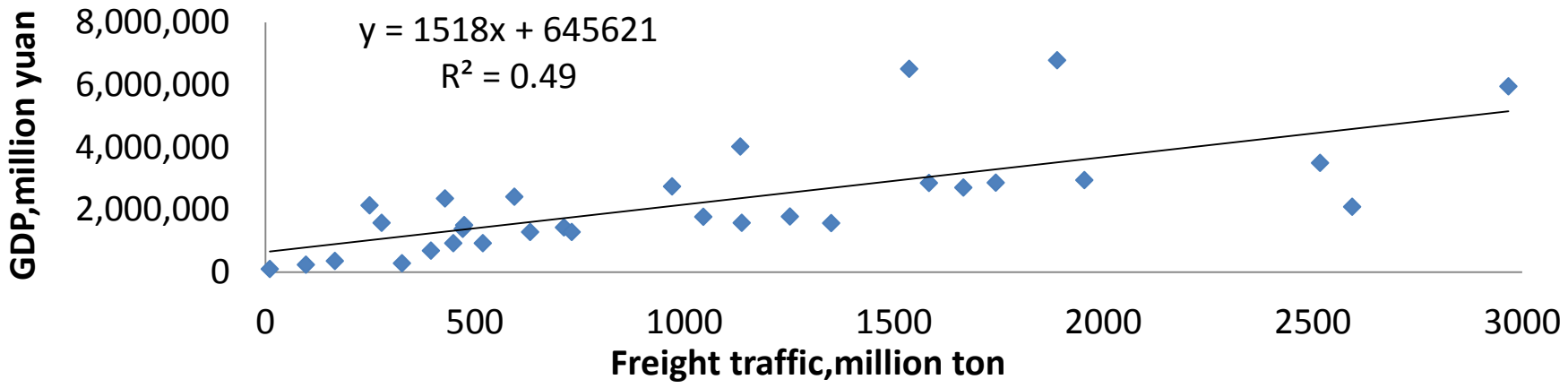


Findings

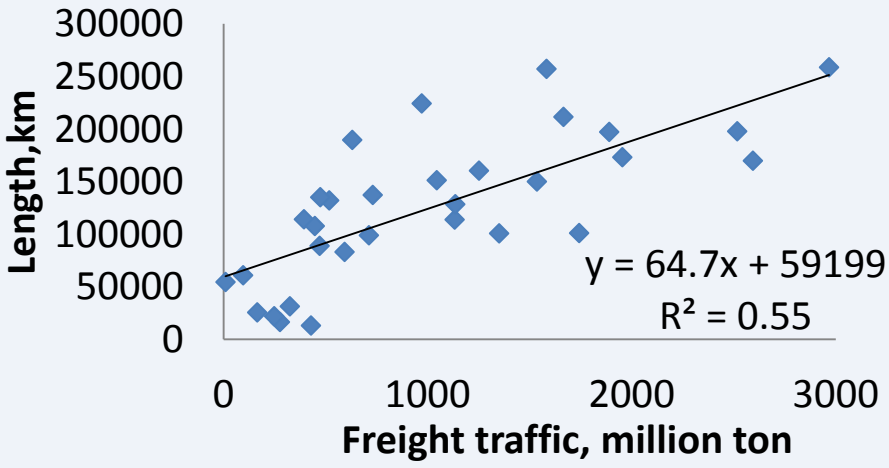
- Provincial economy condition has a prominent relation to road network performance.
- Provinces with high road density care more about road network performance.
- No obvious relation between highway investment and road network performance.

Salient Effect of Freight Traffic on GDP, Population, and Road Length in Different Provinces

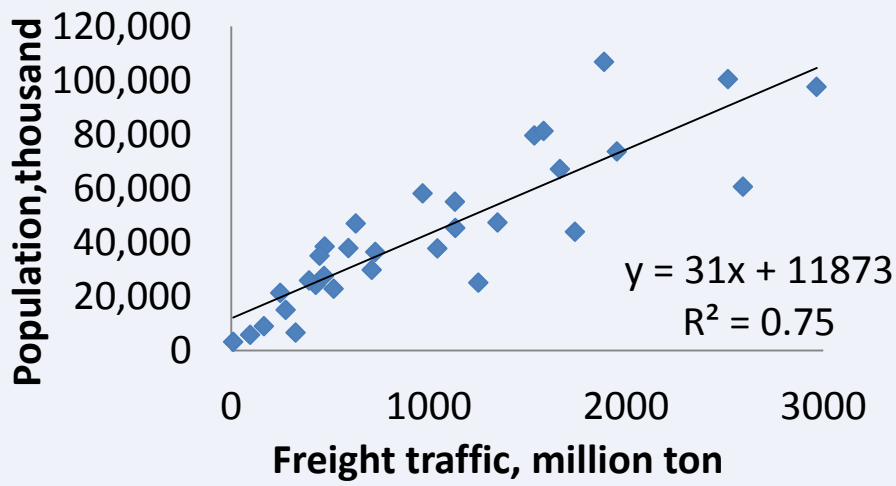
GDP vs freight



Length vs freight



Population vs freight



Stepwise Regression Model

$$\textbf{Ranking} = 28.9 - 5.5\ \textbf{Density} - 0.00016\ \textbf{GDP1} - 0.00183\ \textbf{Freight}$$

$R^2 = 77\%$

Statistics	Density	GDP/capita	Freight
P-value	0.006	0.000	0.064
Variance inflation factor (VIF)	2.1	1.8	1.4

Where,
Ranking= Division ranking based on road performance;
Density=Road density is calculated as the ratio of total road length and land area , km/km^2;
GDP1=GDP per capita, yuan;
Freight= Commercial freight traffic, million ton.

Interpretation

- High road density, GDP per capita, and freight traffic are potential signs to indicate good road performance.
- Estimate possible ranking of other 10 divisions that has no ranking during the 2015 road inspection.

Effect of Road Network on Intercity Transport

Efficient road network improves provincial economy performance by increasing passenger and freight traffic.

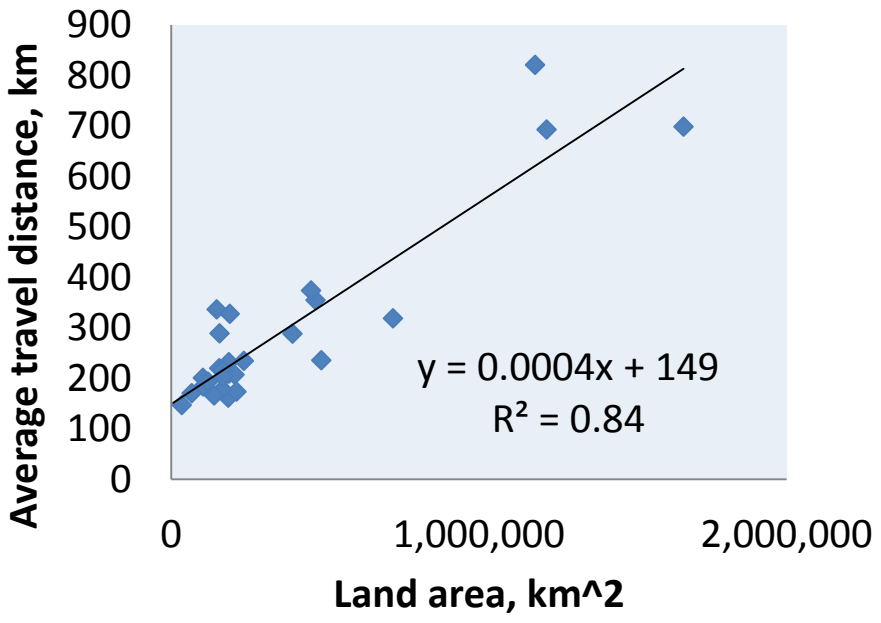
- Largest cities benefit from transportation infrastructure through greater concentration of resources in the PRC.
- Investigate the road network among the large cities is necessary for the understanding of provincial road network.
- Ten cities with the highest GDP were selected for each province and then the travel information from the top city to the rest 9 cities was summarized.

Travel information among top ten cities in Jiangsu

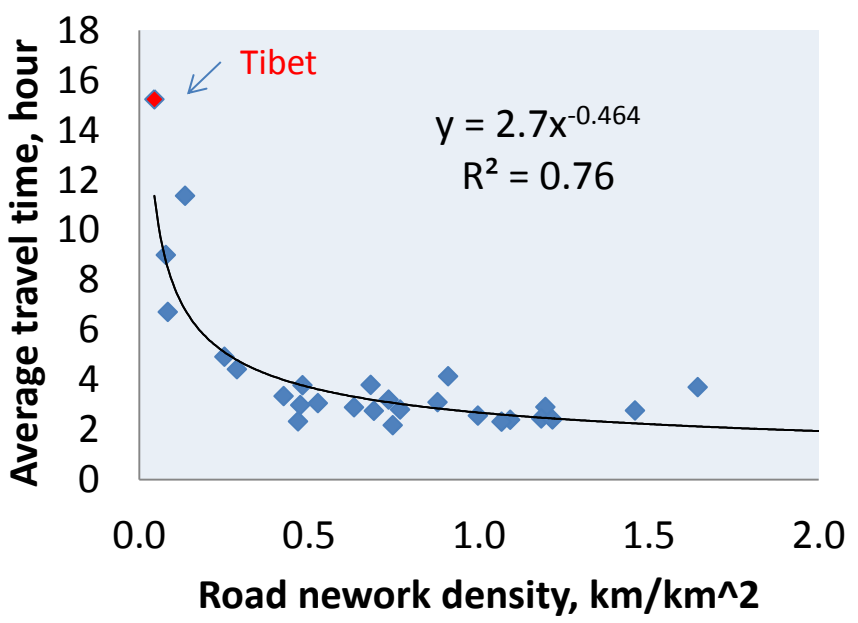
Ranking	City	GDP, hundred million yuan	Population, ten thousand	Travel time, hour	Travel distance, km	Traffic light	Average speed, km per hour
1	Suzhou	14504	1060	0	0		
2	Nanjing	9721	822	3.1	218	29	71
3	Wuxi	8518	650	1.1	54	20	49
4	Nantong	6148	730	1.9	112	16	58
5	Xuzhou	5320	863	6.3	525	33	83
6	Changzhou	5273	470	1.8	102	34	58
7	Yancheng	4213	722	3.4	258	26	76
8	Yangzhou	4017	448	2.6	201	20	78
9	Taizhou	3656	464	2.5	166	31	66
10	Zhenjiang	3502	317	2.3	168	19	73
Average		6487	655	2.8	200	25	68

Average Travel Information in Different Provinces

Travel distance vs area



Travel time vs density



Interpretation

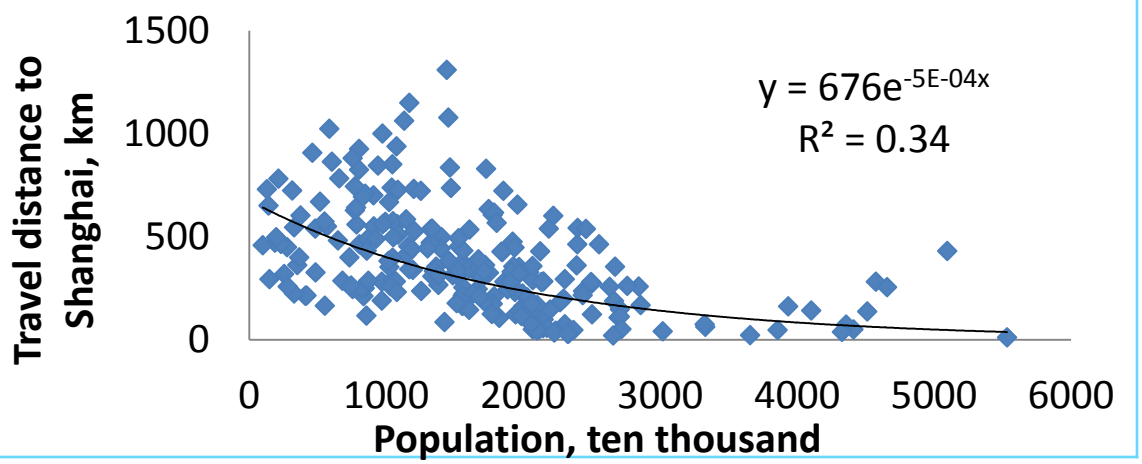
- Travel distance among large cities is long in those provinces with large land area.
- Road density effect is prominent when the road density is at low levels.

Effect of Major Cities on Cities in Different Provinces

Travel distance of top ten cities in Guangdong to the three major cities

Ranking	City	Beijing (km)	Shanghai (km)	Guangzhou (km)
1	Guangzhou	2138	1442	0
2	Shenzhen	2154	1455	138
3	Foshan	2142	1472	34
4	Dongguan	2168	1465	71
5	Huizhou	2072	1369	144
6	Zhongshan	2227	1529	89
7	Zhanjiang	2531	1855	418
8	Maoming	2455	1779	342
9	Jiangmen	2207	1531	92.9
10	Zhuhai	2256	1557	129

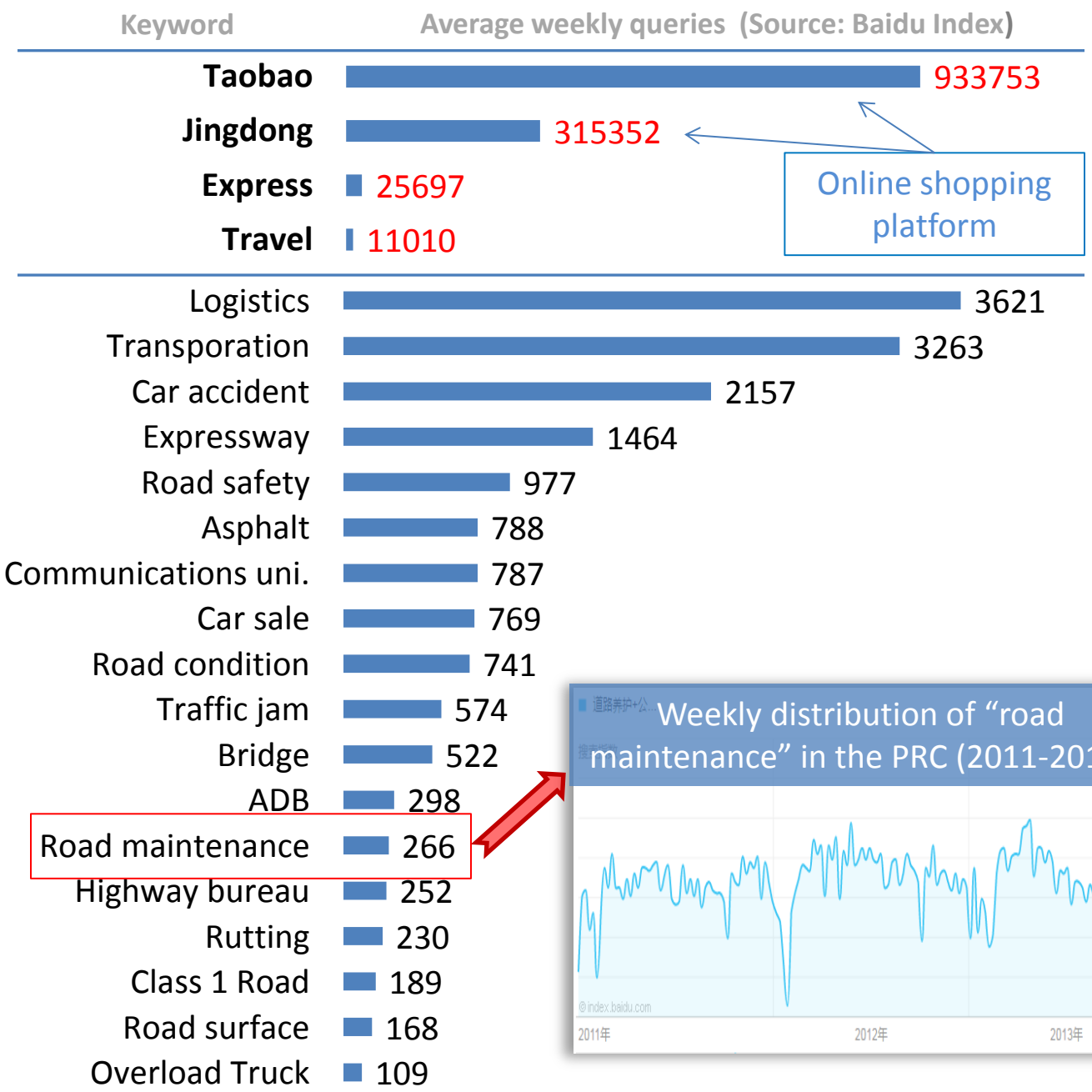
Travel distance to Shanghai vs population (274 cities)



Interpretation

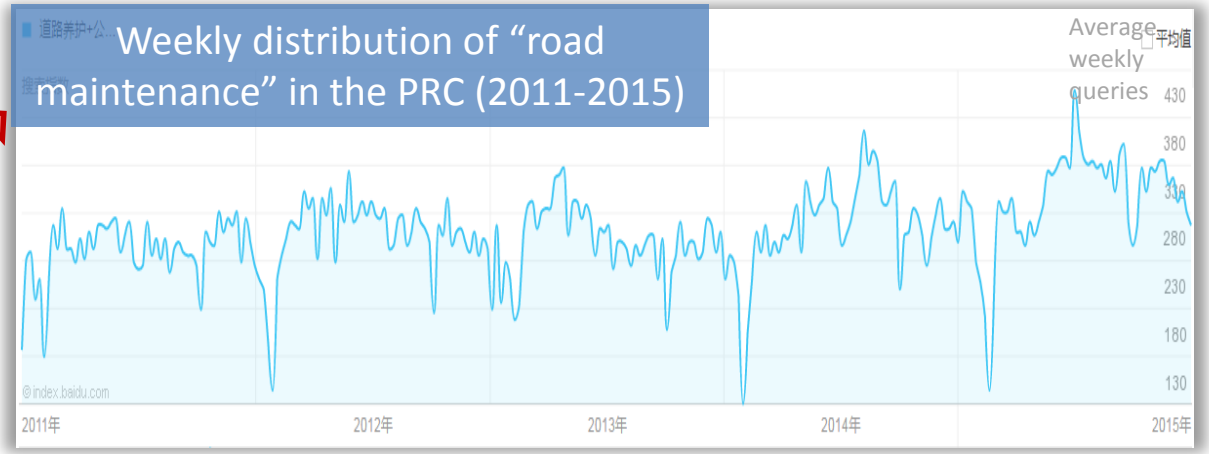
- Economic radiation effect of Shanghai to the development of other cities.
- Road network between the large cities plays an important role in the growth of the cities.
- Improvement in the road network and the reduction of travel distance and time among large cities can facilitate the development of regional economy.

Provincial Road Network and Public Interest on Transportation



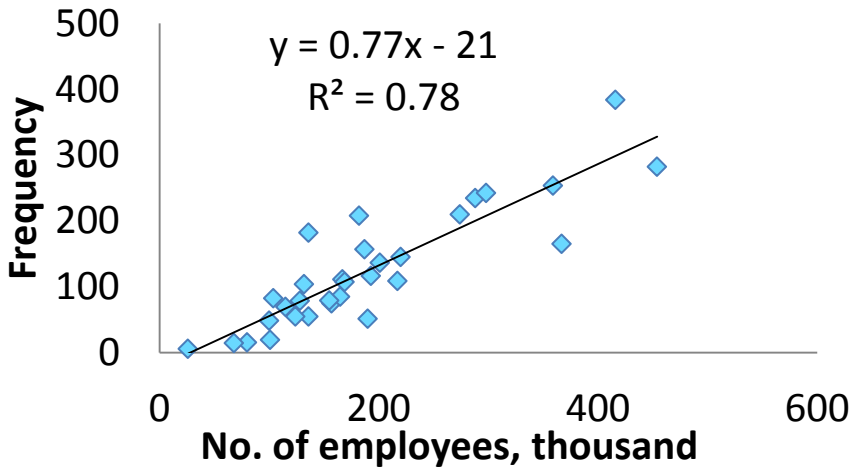
Explanation

- Public awareness on transportation may have certain correlation to provincial road network.
- Use keyword tool to investigate the search patterns of keyword in different provinces.
- Online search can result in travel plan, purchase, report, or an investment decision, which are related to transportation.

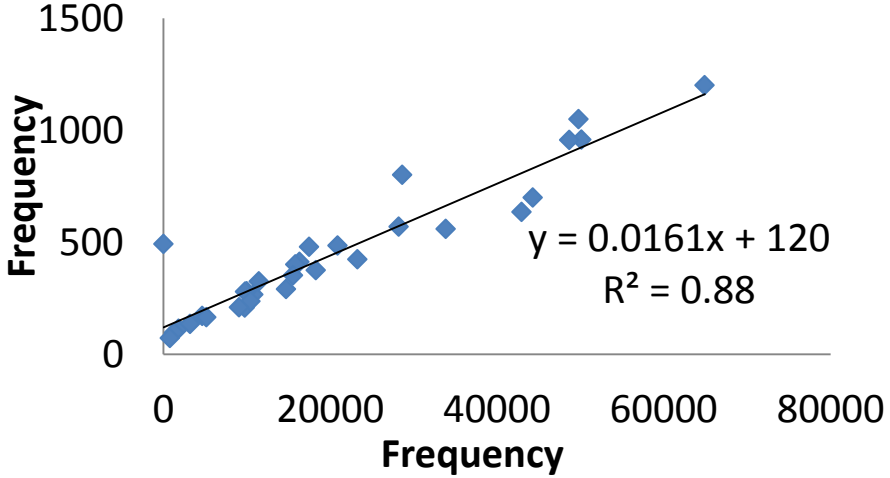


Relation among Keywords in Different Provinces

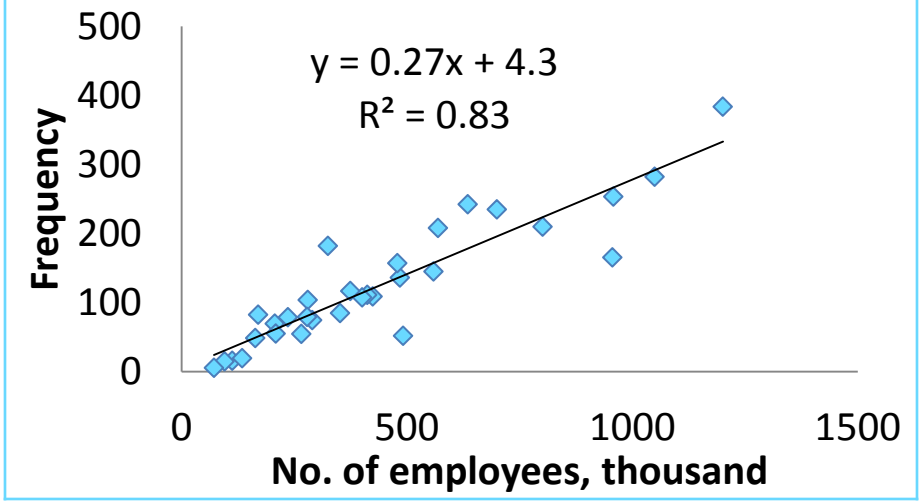
No. of employees vs “Transportation”



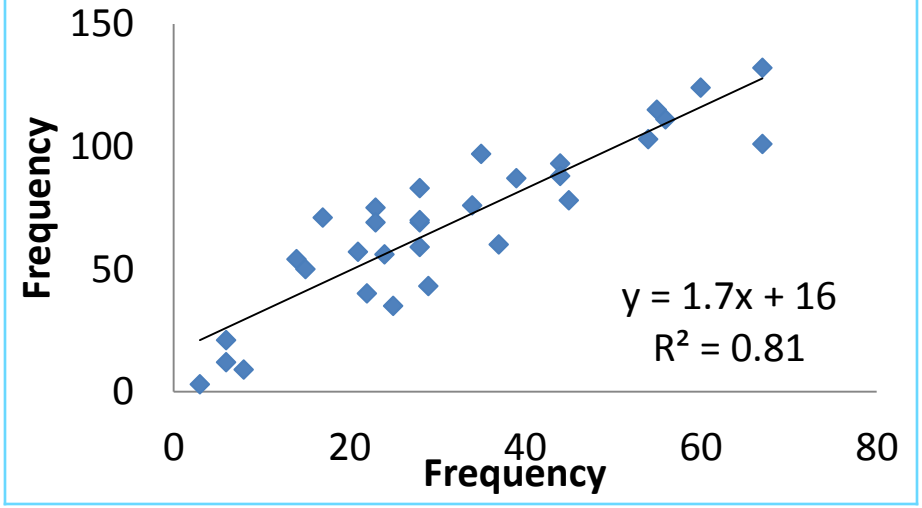
“Taobao" vs “Travel"



No. of employees vs “Travel"

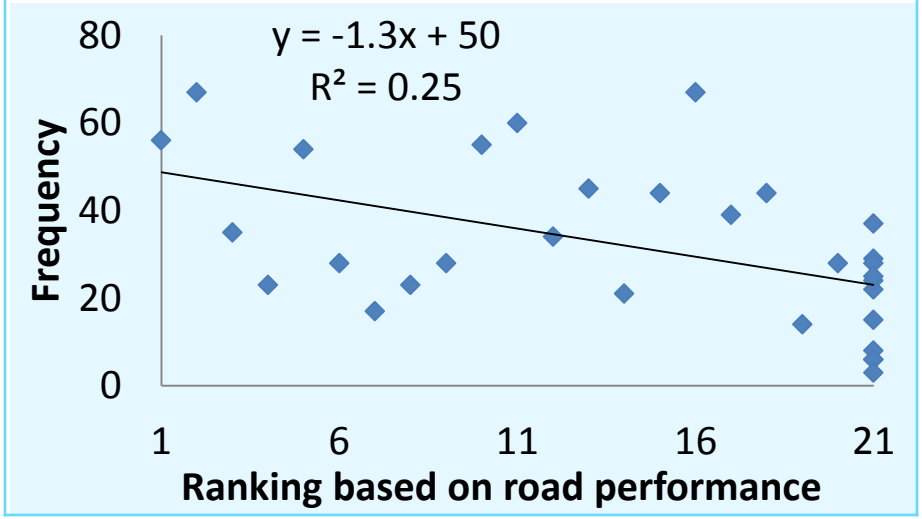


“Traffic jam" vs “Road maintenance"

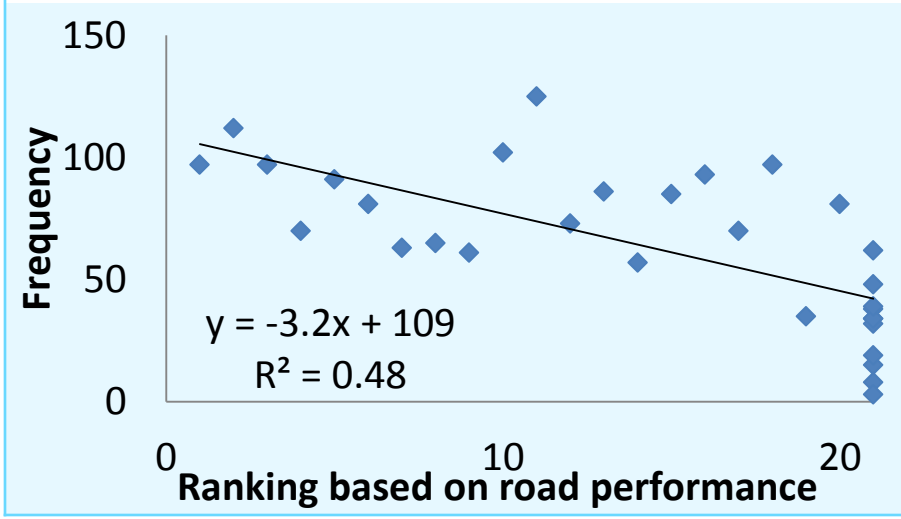


Relationship between Ranking and Keyword Query

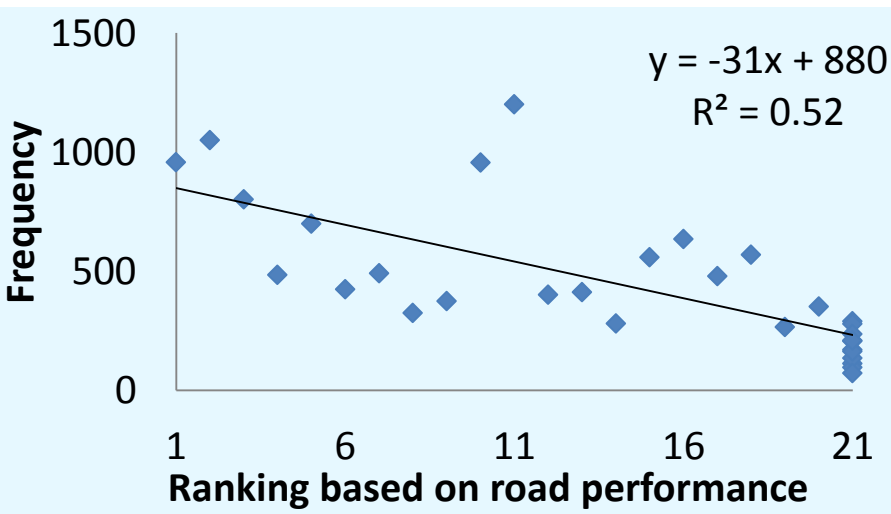
“Road maintenance" vs ranking



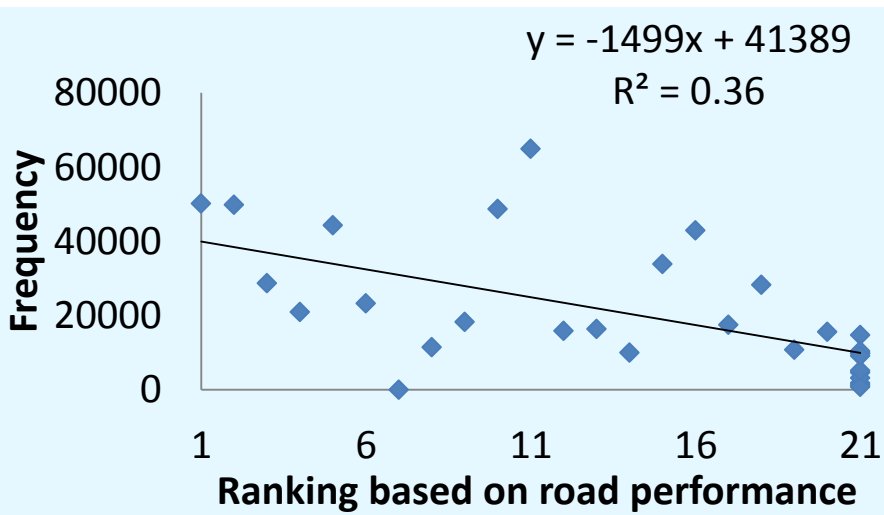
“Freight" vs ranking



“Travel" vs ranking

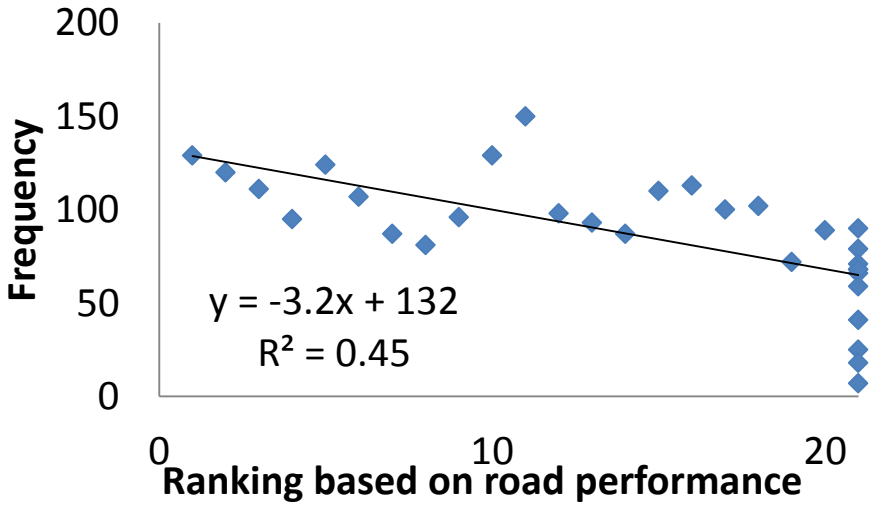


“Taobao" vs ranking

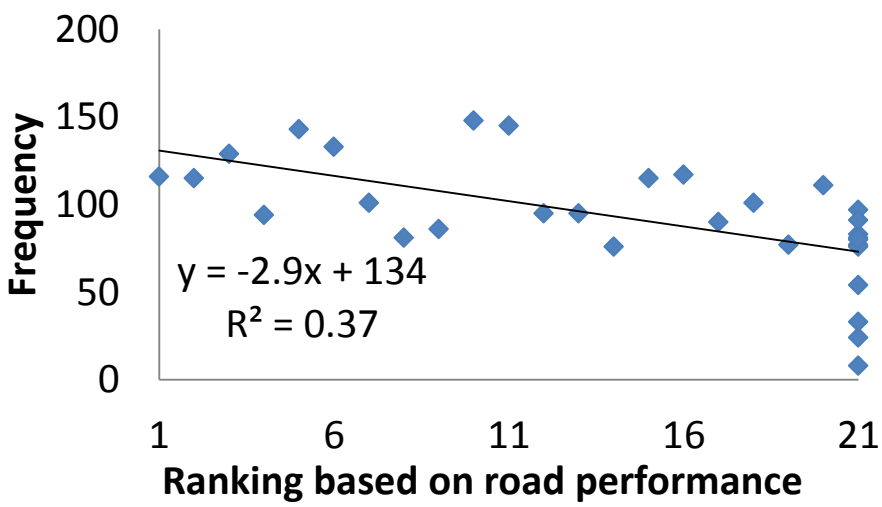


Keyword: a Good Indicator for Road Network Condition

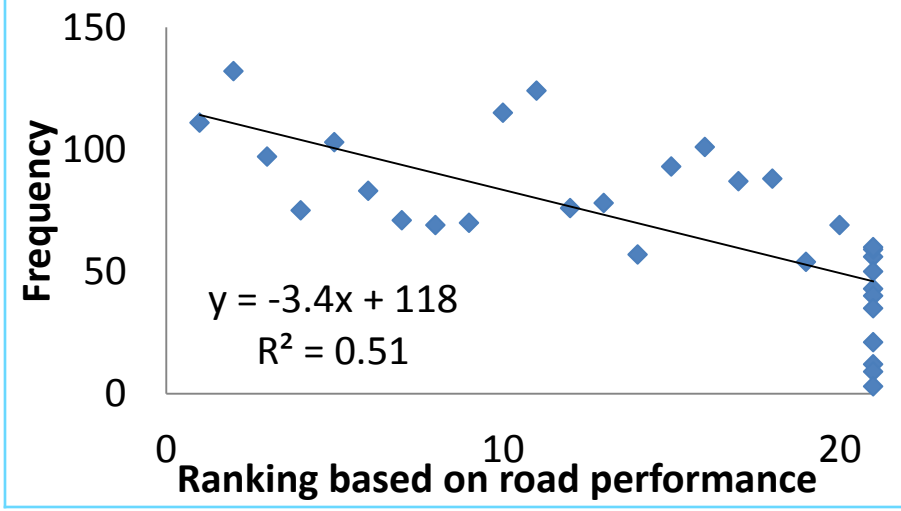
"Asphalt" vs ranking



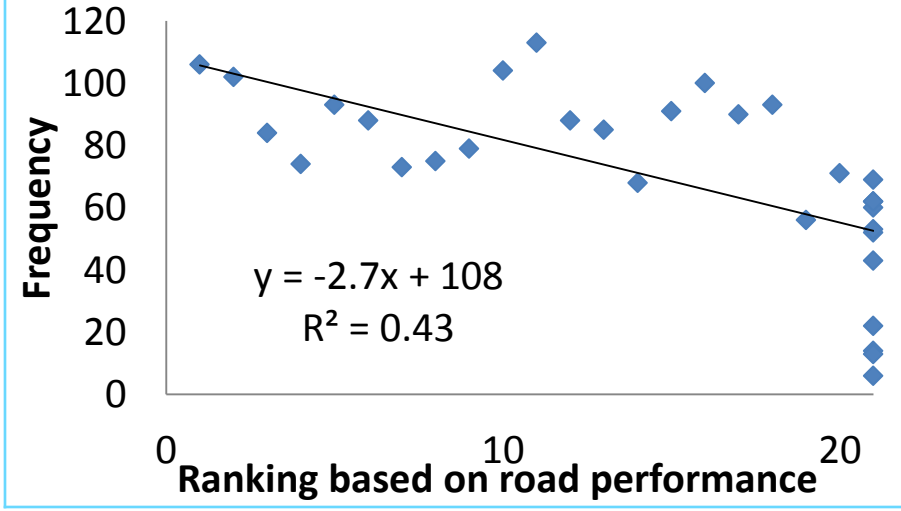
"Road safety" vs ranking



"Travel jam" vs ranking



"Bridge" vs ranking



Influence of Research on Transportation

Researches in transportation are key stimulus for the development of transportation sector in the PRC.

Influence

- Large-scale transportation projects require research skill and professional experience to resolve technical obstacles.
- Advancement in provincial road network is then achieved through the fulfillment of the transportation projects.
- Inspire other practitioners in the transportation sector.
- Amount of scientific papers has certain correlation to the performance of transportation infrastructure in different provinces.

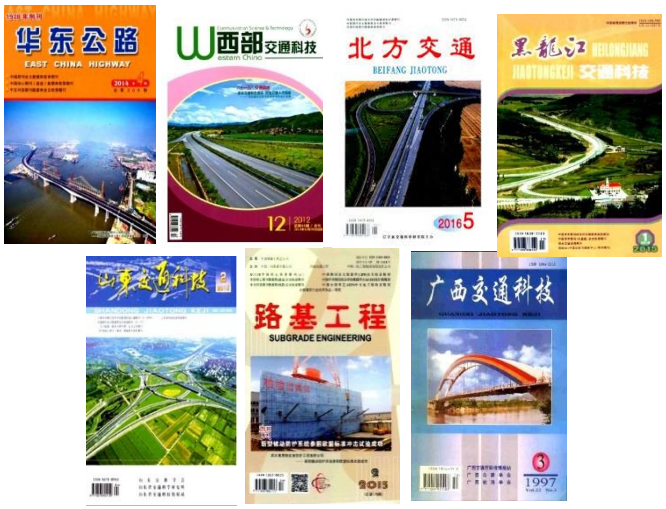
Bibliometric analysis focuses on highway related paper		
✓		✗
• Highway material • Highway construction • Highway foundation • Highway geometry	• Bridge • Tunnel • Traffic • Logistic	• Airport • Railway • Waterway • Building • Car design

Transportation Journals

Approximately 1000 papers published in 2011 from 20 journals were analyzed.

National transportation journal

- Journal of Southeast University
- Journal of Tongji University
- Journal of Harbin Institute of Technology
- Journal of Wuhan University of Technology
- China Journal of Highway and Transport
- China Civil Engineering Journal
- Journal of Highway and Transportation R&D
- Journal of Transportation Systems Eng.& IT
- Journal of Building Materials
- Journal of Transport Information and Safety



Provincial transportation journal

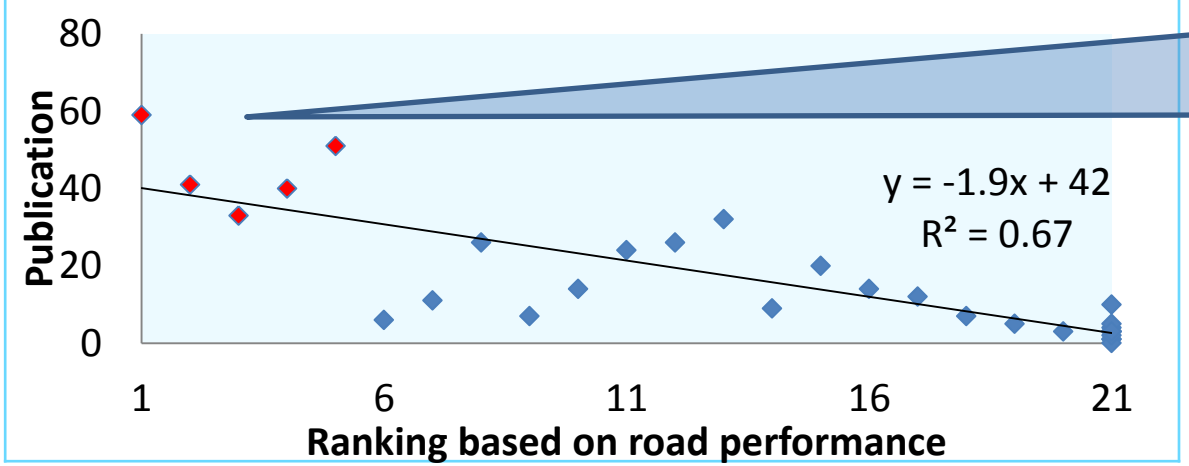
- Foundation Engineering
- Northern Transportation
- Eastern Highway
- Shandong Transportation Technology
- Guangxi Transportation Technology
- Heilongjiang Transportation Technology
- Western Transportation Technology

Effect of Transportation Research on Provincial Road

Type	Name of institution
University	Shandong Architecture University
Research institute	Expressway Maintenance Technology Laboratory Shandong Transportation Institute
Local highway bureau	Jiyang Highway Management Bureau Binzhou Highway Management Bureau Guangrao Transport Agency
State-owned enterprise	Shandong Highway Engineering Construction Group Co., Ltd Shandong Province Highway Design Consulting Co, Ltd, Qing-Lin Expressway Operations Management Center
Privately owned enterprise	Liaocheng Highway Engineering Corporation Jinan Jinyue Highway Engineering Corporation Shandong Taihe Highway Engineering Co., Ltd

Workplace of first authors from Shandong Province

Publication vs road performance



Authors from Jiangsu, Beijing, Shanghai, Liaoning, Shandong publish large amount of scientific papers annually.

Conclusion and Recommendation

Factor	Data source	Finding	Recommendation
Economy 	CEIC, China Statistics Bureau	• Province with good road performance usually has high GDP per capita, road density and freight traffic.	More money should be spent in road maintenance.
Geography 	Gaode Map	• High road density has positive effect on reducing travel time. • Economic radiation of Shanghai to the development of other cities is prominent.	Investment focuses on reducing travel time among large cities and improve regional economy.
Awareness 	Baidu Index	• Increasing public interest in transportation could lead to the elevated expenditure and investment in transportation sector.	Use transportation related keywords to predict road network condition.
Research 	Online publications	• Professional capability of transportation practitioners is important in improving provincial road network.	Fund more research projects and improve research skill.

Thank you for your attention

Any questions?