

일반철도의 수송수요 변화 추정모형 개발

Development of Travel Demand Forecasting Model for Conventional Passenger Railway Service

김장욱*†, 금수희*, 오영택*, 신미화**

Janguok Kim*†, Suhui Keum, Youngteak Oh, Mi-hwa Shin

Abstract As short term demand forecasting of railway passenger service is based on transportation data of the past, it has limits to explain all the various changes on railway services including opening of new lines and changes on train operation frequencies. Therefore, this study establishes a model which is able to consider socio-economic indicators and the impacts of other travel modes. And it analyses the trip distribution characteristics by deriving usage probability depending on a line distance and threshold of the moving distance of the railway lines. Demand estimation model for the conventional passenger service is built by applying distance variables additionally obtained from passage characteristics analysis. Furthermore, improvement in model's suitability is obtained by building up separated model by lines rather than building up an integrated railway model as the leverage of the independent variables were different by lines in the statistics. Therefore, this study finalizes the forecasting model through selecting significant variables for the Saemaoul and Mugunghwa service in the 5 trunk lines(Gyeongbu line, Honam line, Jeolla line, Janghang line, Jungang line), and the sensitivity analysis shows that fare has greater impact on rail travel demand than travel time changes in the study.

Keywords : Rapid transit railway, Demand forecasting, Reliability theory, Fare, Travel times

초 록 일반철도의 단기수송수요 예측은 과거 수송실적에 근거하기 때문에 신설노선이나 열차운행횟수 변화들과 같은 다양한 철도공급서비스 변화를 설명하는 데 한계가 있다. 이에 본 연구에서는 사회경제지표와 더불어 타 교통수단의 영향 등을 함께 고려할 수 있는 모형을 구축하였다. 그리고 신뢰성이론을 이용하여 노선별 거리에 따른 이용확률 및 이동한계점을 도출하여 통행특성을 분석하였다. 통행특성분석을 통해 산출된 거리에 따른 변수를 추가로 반영하여 일반철도에 대한 수송수요 변화 추정모형을 구축하였다. 또한 노선별로 독립변수의 영향력이 통계적으로 상이하여 철도를 통합모형으로 구축하는 것 보다는 노선별로 분리하여 개별모형을 추정하는 것이 모형의 적합성을 향상시킬 수 있었다. 본 연구에서는 새마을 및 무궁화 열차의 5대 간선(경부선/호남선/전라선/장항선/중앙선)을 대상으로 모형을 구축한 결과 운임이 통행시간보다 영향력이 있는 것으로 분석되었다.

주요어 : 일반철도, 수요예측, 신뢰성이론, 철도운임, 통행시간

† 교신저자: 한국철도공사 연구원 경영연구처(kjw@korail.com)

* 한국철도공사 연구원 경영연구처, ** 한국철도공사 여객본부 역운영처

1. 서 론

일반여객 단기수송수요 예측에 주로 사용하던 시계열의 경우 과거 수송실적에 근거하기 때문에 다양한 철도공급 서비스 변화에 대한 수요량의 변화를 설명하는 데는 제한이 있다. 특히 급변하는 교통시장 및 고객니즈(needs)의 다양화 및 고급화 추세에 따라 철도서비스 품질에 대한 꾸준한 관리가 요구되는 현실에 비추어 볼 때 선택행위를 보다 구체적으로

표시할 수 있는 모형의 필요성이 강조되고 있는 실정이다. 이에 본 연구에서는 새마을과 무궁화 열차의 5대 간선(경부선/호남선/전라선/장항선/중앙선)을 대상으로 운임, 통행시간 등의 철도 서비스정책 변화에 따른 수요변화를 추정하고자 한다.

2. 일반철도의 수요변화 추정모형 개발 방법

2.1 수요변화 추정모형 개발 방향 및 개선사항

철도요금과 열차통행시간의 수요탄력성 분석방법에는 크게 교통수단선택 모형에서 얻어진 시간과 비용 파라미터(Parameter)를 해석하는 방법과, 수요량을 종속변수로 간주하고 요금이나 시간변화를 독립변수로 간주하여 회귀모형식을 추정하고 변수들의 영향을 분석하는 방법이 있는데 본 연구에서는 직접수요모형을 이용하여 정책변수에 대한 영향력을 분석하고자 한다. 대부분의 직접수요모형 구축에 관한 연구는 고속철도(KTX) 및 일반철도 경부선과 호남선 위주로 분석되었고, 다른 노선에 대한 연구는 미흡한 실정이다. 노선별 독립변수의 영향력이 통계적으로 다르므로 개별모형을 추정하는 것이 모형의 적합성을 향상시킬 수 있으나, 기존 연구에서는 통합된 모형으로 구축하거나 공통된 변수를 가지고 분석하여 모형의 신뢰도가 낮았다. 이에 본 연구에서는 새마을 및 무궁화 열차의 5대 간선을 대상으로 열차 및 노선별로 유의한 변수를 선정하여 각각의 개별모형을 구축하였다.

2.2 자료 구축 및 추정방법

직접수요모형 구축을 위해 횡단면 자료(Cross-sectional data)를 이용하여 지역간 기종점 통행량(O/D)의 연총량을 종속변수로 하고, 독립변수는 교통수요를 유발 추정되는 다양한 사회경제지표와 열차 정책요소인 서비스 공급지표, 그리고 개인교통 수단의 영향, 거리별 더미(Dummy)로 구분하여 연구가설을 설정하였다. 또한 변수 선정시 일반열차(새마을 및 무궁화 열차)의 경우 정차역이 많고 같은 도시내에서 여러번 정차하는 경우가 많아 비슷한 특징을 가진 권역별로 구분할 필요가 있다. 이에 노선별로 시군 단위의 존(Zone)을 설정하고 중력모형을 이용하여 두 지역간 선정된 변수의 곱에 자연로그를 취한 이중로그 회귀모형을 이용하여 분석하였다. 선정된 변수간 설명력 높은 변수를 도출하기 위해 상관관계 분석을 실시한 결과, 사회경제지표로는 인구, GRDP, 종사자수, 도시지역면적이 철도서비스 공급지표로는 열차운임, 열차통행시간, 열차운행거리가 공선성이 존재하는 것으로 분석되었다.

Table 1 Development of data for direct demand model

		Contents	Methods of data development
Dependent variable		Annual regional(O/D) passengers in 2013 by train, line	Transport record in KORAIL
Independent variable	Socio-economic index	Population GRDP Number of employees Number of car registered/population Length of highway/length of road	Statistics Korea (Korean Statistical Information Service)
	Service supply index (Policy factors)	Train fares Train operation hours Train operation frequency Train operation distance	KORAIL Train Service Information
	Effects on car usage	Car driving distance Car driving hours Toll	Korea Expressway Corporation Ministry of Government Administration and Home Affairs (Fare Estimation Standard)
	Others	Dummy distance	KORAIL Distance between stations(O/D)

Remarks) In case of population variables, economically active population means figures including number of employees and unemployed people over 15 years old. Therefore, it decided to use number of employees and number of population which indicates number of working people and scale respectively for final variables in this study.

이러한 공선성 존재로 인해 열차운임과 열차통행시간을 기준으로 단계별 입력방식을 통해 각각 4개의 모형을 구축하고 편상관계수 및 t값, R²을 활용하여 기여도가 높은 변수 및 적합한 모형을 도출하였다.

3. 일반철도의 수요변화 추정모형 개발

3.1 새마을 경부선 직접수요모형

열차운임과 통행시간 모두 유효한 변수로 유의확률 95% 이내에서 채택되었으나, 운임의 설명력이 더 높은 것으로 분석되었다. 운임을 베이스(base)로 한 직접수요모형 구축 결과 열차운행빈도, 도시지역면적, 열차운임, 자동차등록대수/인구가 최종변수로 채택되었고, 통행량에 가장 큰 영향을 미치는 변수로는 열차운행빈도로 나타났다.

$$\ln O / D \text{간수송량} = 19.218 + 1.211 \ln \chi_1 + 1.563 \ln \chi_2 - 0.746 \ln \chi_3 - 2.385 \ln \chi_4 \quad (1)$$

χ_1 : 두지역간열차운행빈도

χ_2 : 두지역간도시지역면적 곱

여기서,

χ_3 : 두지역간열차운임

χ_4 : 두지역간자동차등록대수/인구

Table 2 Analysis result of direct demand modeling for Saemaedul service on Gyeongbu line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	19.218	7.638		0.013		
Train service frequency	1.211	0.135	0.446	0.000	0.744	1.344
Area of urban region	0.563	0.090	0.404	0.000	0.446	2.244
Train fares	-0.746	0.144	-0.238	0.000	0.878	1.139
Number of cars registered/population	-2.385	0.811	-0.184	0.004	0.473	2.116
R-Square			0.725			
Adjusted R-Square			0.717			

3.2 새마을 호남선 직접수요모형

열차운임과 열차통행시간은 수송량과의 상관관계가 낮고 유의하지 않아 모형 구축시 채택되지 않았고, 사회경제지표 중 도시지역면적, 열차운행빈도, 고속도로연장/총 도로길이, 자동차등록대수/인구가 최종 변수로 채택되었다. 통행량에 가장 큰 영향을 미치는 변수로는 열차운행빈도로 나타났다.

$$\ln O / D \text{간수송량} = 7.756 + 0.545 \ln \chi_1 + 2.066 \ln \chi_2 - 0.127 \ln \chi_3 - 1.764 \ln \chi_4 \quad (2)$$

χ_1 : 두지역간도시지역면적 곱

χ_2 : 두지역간열차운행빈도

여기서,

χ_3 : 두지역간고속도로연장/총도로길이

χ_4 : 두지역간자동차등록대수/인구

3.3 새마을 전라선 직접수요모형

호남선과 마찬가지로 열차운임과 열차통행시간은 수송량과의 상관관계가 낮고 유의하지 않아 모형구축시 채택되지 않았으며, 종사자수, 운행빈도, 고속도로연장/총 도로길이, 단거리더미가 최종변수로 채택되었다.

$$\ln O / D \text{간수송량} = -4.926 + 0.487 \ln \chi_1 + 2.318 \ln \chi_2 - 0.072 \ln \chi_3 + 0.566 \ln \chi_4 \quad (3)$$

여기서, χ_1 : 두지역간종사자수급
 χ_2 : 두지역간열차운행빈도
 χ_3 : 두지역간고속도로연장 / 총도로길이
 χ_4 : 두지역간거리100km미만이면1,그외0

Table 3 Analysis result of direct demand modeling for Saemaeul service on Honam line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	7.756	4.111		0.060		
Area of urban region	0.545	0.042	0.421	0.000	0.503	1.986
Train service frequency	2.066	0.092	0.569	0.000	0.824	1.213
Length of highway/total length of road	-0.127	0.027	-0.135	0.000	0.644	1.552
Number of cars registered/population	-1.764	0.491	-0.132	0.000	0.396	2.524
R-Square				0.862		
Adjusted R-Square				0.860		

Table 4 Analysis result of direct demand modeling for Saemaeul service on Jeolla line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	-4.926	0.787		0.000		
Number of employees	0.487	0.032	0.660	0.000	0.875	1.143
Train service frequency	2.318	0.231	0.440	0.000	0.851	1.176
Length of highway/total length of road	-0.072	0.033	-0.093	0.034	0.879	1.138
Short distance dummy	0.566	0.144	0.161	0.000	0.974	1.027
R-Square				0.774		
Adjusted R-Square				0.767		

3.4 새마을 중앙선 직접수요모형

도시지역면적, 고속도로연장/총도로길이, 단거리더미가 통계적으로 유의확률 95%내에서 최종 변수로 채택되었으며, 열차운임은 수송량과의 상관관계가 낮아 모형 구축시 채택되지 않았다.

$$\ln O / D \text{간수송량} = -2.846 + 0.607 \ln \chi_1 - 0.242 \ln \chi_2 + 0.799 \ln \chi_3 \quad (4)$$

χ_1 : 두지역간도시지역면적 곱

여기서, χ_2 : 두지역간고속도로연장 / 총도로길이

χ_3 : 두지역간거리100km미만이면1,그외0

3.5 새마을 장항선 직접수요모형

열차운임과 통행시간 중 운임이 유효한 변수로 유의확률 95%이내에서 채택되었으나, 결정계수(R-Square)값이 0.207로 O/D간 수송량 예측의 약 20% 설명력을 가지는 것으로 다른 노선에 비해 낮은 설명력을 가진 것으로 분석되었다. 이는 새마을 장항선의 경우 출퇴근 및

관광통행 등과 같은 타 노선과는 다른 수송특성으로 인해 현재 채택된 변수만으로는 장항선의 수송분포 특성을 반영하지 못해 모형의 신뢰도가 낮은 것으로 판단됨.

$$\ln O/D \text{간수송량} = 10.119 - 1.325 \ln \chi_1 + 0.347 \ln \chi_2 \quad (5)$$

여기서, χ_1 : 두지역간열차운임
 χ_2 : 두지역간GRDP급

Table 5 Analysis result of direct demand modeling for Saemaedul service on Jungang line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	-2.846	1.605		0.084		
Area of urban region	0.607	0.061	0.800	0.000	0.599	1.669
Length of highway/total length of road	-0.242	0.087	-0.226	0.008	0.577	1.733
Short distance dummy	0.799	0.183	0.285	0.000	0.905	1.106
R-Square			0.854			
Adjusted R-Square			0.843			

Table 6 Analysis result of direct demand modeling for Saemaedul service on Janghang line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	10.119	4.086		0.016		
Train fares	-1.325	0.397	-0.440	0.002	0.859	1.164
GRDP	0.347	0.124	0.368	0.007	0.859	1.164
R-Square			0.207			
Adjusted R-Square			0.178			

3.6 무궁화 경부선 직접수요모형

열차운임과 통행시간 모두 유효한 변수로 유의확률 95%이내에서 채택되었으나, 새마을과 달리 통행시간의 설명력이 더 높은 것으로 분석되었다. 행시간을 베이스(base)로 한 직접수요 모형 분석 결과 열차운행빈도, GRDP, 열차통행시간, 고속도로 연장/총 도로연장, 장거리 더미, 단거리 더미가 최종변수로 채택되었다.

$$\ln O/D \text{간수송량} = -7.824 + 1.337 \ln \chi_1 + 0.541 \ln \chi_2 - 0.380 \ln \chi_3 - 0.250 \ln \chi_4 - 0.685 \chi_5 + 0.733 \chi_6 \quad (6)$$

χ_1 : 두지역간열차운행빈도
 χ_2 : 두지역간GRDP급
 여기서, χ_3 : 두지역간열차운행시간
 χ_4 : 두지역간고속도로연장 / 총도로길이
 χ_5 : 두지역간거리300km이상이면1, 그외0
 χ_6 : 두지역간거리100km미만이면1, 그외0

Table 7 Analysis result of direct demand modeling for Mugunghwa service on Gyeongbu line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	-7.824	1.362		0.000		
Train service frequency	1.337	0.076	0.471	0.000	0.726	1.377
GRDP	0.541	0.031	0.520	0.000	0.578	1.730
Train operation hours	-0.380	0.126	-0.146	0.003	0.218	4.594
Length of highway/total length of road	-0.250	0.059	-0.129	0.000	0.560	1.787
Long distance dummy	-0.685	0.188	-0.107	0.000	0.601	1.664
Short distance dummy	0.733	0.203	0.160	0.000	0.264	3.795
R-Square			0.833			
Adjusted R-Square			0.830			

3.7 무궁화 호남선 직접수요모형

열차운임과 통행시간 모두 유효한 변수로 유의확률 95% 이내에서 채택되었으나, 통행시간의 설명력이 더 높은 것으로 분석되었다. 통행시간을 베이스(base)로 한 직접수요모형 분석 결과 열차운행빈도, GRDP, 자동차등록대수/인구, 장거리더미, 고속도로 연장/총 도로길이, 열차통행시간이 최종변수로 채택되었고, 열차운행빈도가 통행량에 가장 큰 영향을 미치는 것으로 분석되었다.

$$\ln O/D \text{간수송량} = 11.119 + 1.54 \ln \chi_1 + 0.386 \ln \chi_2 - 2.146 \ln \chi_3 - 0.088 \ln \chi_4 - 0.177 \chi_5 - 0.520 \chi_6 \quad (7)$$

χ_1 : 두지역간열차운행빈도

χ_2 : 두지역간 GRDP

여기서, χ_3 : 두지역간자동차등록대수/인구

χ_4 : 두지역간고속도로연장 / 총도로길이

χ_5 : 두지역간열차운행시간

χ_6 : 두지역간거리300km이상이면1,그외0

Table 8 Analysis result of direct demand modeling for Mugunghwa service on Honam line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	11.119	4.725		0.019		
Train service frequency	1.541	0.074	0.657	0.000	0.650	1.540
GRDP	0.386	0.034	0.471	0.000	0.385	2.600
Number of cars registered/population	-2.146	0.554	-0.191	0.000	0.267	3.749
Long distance dummy	-0.520	0.161	-0.096	0.001	0.739	1.354
Length of highway/total length of road	-0.088	0.028	-0.100	0.002	0.655	1.527
Train operation hours	-0.177	0.057	-0.089	0.002	0.784	1.276
R-Square			0.800			
Adjusted R-Square			0.796			

3.8 무궁화 전라선 직접수요모형

열차운임과 통행시간 모두 유효한 변수로 유의 확률 95%이내에서 채택되었으나, 열차운임의 설명력이 더 높은 것으로 분석되었다. 운임을 베이스(base)로 한 직접수요 모형 분석 결과 인구, 열차운행빈도, 열차운임, 고속도로연장/총 도로길이가 최종변수로 채택되었고, 인구가 통행량에 가장 큰 영향을 미치는 것으로 분석되었다.

$$\ln O / D \text{간수송량} = -4.385 + 0.573 \ln \chi_1 + 1.632 \ln \chi_2 - 0.432 \ln \chi_3 - 0.056 \ln \chi_4 \quad (8)$$

여기서, χ_1 : 두지역간인구급
 χ_2 : 두지역간열차운행빈도
 χ_3 : 두지역간열차운임
 χ_4 : 두지역간고속도로연장 / 총도로길이

Table 9 Analysis result of direct demand modeling for Mugunghwa service on Jeolla line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	-4.385	0.916		0.000		
Population	0.573	0.025	0.683	0.000	0.909	1.100
Train operation frequency	1.632	0.075	0.619	0.000	0.965	1.036
Train fares	-0.432	0.070	-0.177	0.000	0.936	1.068
Length of highway/total length of road	-0.056	0.025	-0.065	0.028	0.916	1.092
R-Square			0.791			
Adjusted R-Square			0.788			

3.9 무궁화 중앙선 직접수요모형

열차운임과 통행시간 모두 수송량과의 상관관계가 낮아 모형구축 시 채택되지 않았고, 열차운행빈도, GRDP가 유의확률 95%에서 통계적으로 유의한 최종변수로 채택되었으며, 열차운행빈도가 통행량에 가장 큰 영향을 미치는 것으로 분석되었다.

$$\ln O / D \text{간수송량} = -2.669 + 1.997 \ln \chi_1 + 0.266 \ln \chi_2 \quad (9)$$

여기서, χ_1 : 두지역간열차운행빈도
 χ_2 : 두지역간GRDP급

Table 10 Analysis result of direct demand modeling for Mugunghwa service on Jungang line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	-2.669	1.362		0.052		
Train operation frequency	1.997	0.120	0.764	0.000	0.920	1.087
GRDP	0.266	0.047	0.261	0.000	0.920	1.087
R-Square			0.764			
Adjusted R-Square			0.760			

3.10 무궁화 장항선 직접수요모형

열차운임과 통행시간 중 운임이 유효한 변수로 유의확률 95% 이내에서 채택되었으나, 결정계수(R-Square) 값이 0.294로 O/D간 수송량 예측의 약 29% 설명력을 가지는 것으로 다른 노선에 비해 낮은 설명력을 가진 것으로 분석되었다. 이는 무궁화 장항선의 경우 출퇴근 및 관광통행 등과 같은 타 노선과는 다른 수송특성으로 인해 현재 채택된 변수 만으로는 장항선의 수송분포 특성을 반영하지 못해 모형의 신뢰도가 낮은 것으로 판단되었다.

$$\ln O / D\text{간수송량} = -59.405 - 1.064 \ln \chi_1 + 0.191 \ln \chi_2 - 9.705 \chi_3 \quad (10)$$

χ_1 : 두지역간열차운임

여기서, χ_2 : 두지역간 GRDP

χ_3 : 두지역간자동차등록대수 / 인구

Table 11 Analysis result of direct demand modeling for Mugunghwa service on Janghang line

Model	Unstandardized coefficient		Standardized coefficient	Significance probability	Collinearity statistic	
	B	Standard error	Beta		Tolerance	VIF
(constant)	-59.405	28.787		0.044		
Train fares	-1.064	0.310	-0.430	0.001	0.863	1.159
GRDP	0.191	0.116	0.205	0.104	0.884	1.132
Number of cars registered	9.705	3.779	0.303	0.013	0.972	1.029
R-Square			0.294			
Adjusted R-Square			0.254			

4. 결론

본 연구에서는 새마을과 무궁화 열차의 5대간선(경부선, 호남선, 전라선, 장항선, 중앙선)을 대상으로 운임, 통행시간 등의 철도정책 서비스변화에 따른 수요변화 추정모형을 개발하였다. 특히 운임과 통행시간은 거리에 대한 운영체제로 서로 상관관계가 존재하여 단계별 입력 방식을 통해 운임과 통행시간 각각을 기준으로 4개의 개별 모형이 구축되었고, 열차종 및 노선별로 채택되는 변수가 유의하였다. 직접수요모형을 이용한 추정모형의 탄력성 분석을 통해 얻은 결과를 기반으로 노선별 열차운행계획을 위한 기초자료로 활용될 수 있을 것이다.

참고문헌

- [1] 김효종·정찬묵(2012) 철도수요예측을 위한 직접수요모형 개발에 관한 연구, 한국철도학회 학술발표대회논문집, pp. 2166-2178.
- [2] 대한교통학회(2013) 수도권 고속철도 운영관련 수송수요 예측 연구.
- [3] 도명식(2010) 신뢰성 개념을 이용한 포장의 평균수명 및 신뢰도 예측, 대한토목학회, 대한토목학회논문집, 제30권 제5호, pp. 497-504.
- [4] 이용상·권용장(2003) 수송가격 및 수요와 열차운행, 한국철도기술연구원.
- [5] 엄진기 (2009) 수송실적자료를 이용한 철도교통 수요변화 추정 연구, 한국철도학회논문집 제12권 제3호, pp 420~426.
- [6] 정철, 김시곤, 김찬성(2007) 지역간 철도수요분석에서 비용과 시간탄력성의 비교연구: KTX 수요에 대한 탄력성을 중심으로, 대한토목학회논문집, 제27권 제5호, pp. 547-553.
- [7] 정찬묵(2012), 주성분해석을 통한 철도이용객 수요에 미치는 사회경제지표 분석, 한국콘텐츠학회논문지 제12권 제7호, pp 437~444.
- [8] 한국철도공사 연구원(2011) 간선여객 단기수송수요 예측방법 연구.
- [9] 한국철도공사 연구원(2012) 2013년도 KTX 및 일반여객 수송수요예측.
- [10] 한국철도공사 연구원(2013) 2014년도 일반여객 수송수요예측.