

비컨테이너 항만물동량 예측모형 고도화 방안 연구(Ⅰ)

－ 자동차, 석탄, 고철, 시멘트를 중심으로

A Study on the Advancement of the Non-container Port
Throughput Forecasting Model (Ⅰ)
－ Focusing on Automobile, Coal, Scrap Metal and Cement

최석우 · 이화섭 · 김영훈 · 조성현 · 이기열



한국해양수산개발원
KOREA MARITIME INSTITUTE

저자	최석우, 이화섭, 김영훈, 조성현, 이기열			
내부연구진	연구책임자	최석우	한국해양수산개발원	항만연구본부 부연구위원
	공동연구원	이화섭	한국해양수산개발원	항만연구본부 전문연구원
	공동연구원	김영훈	한국해양수산개발원	항만연구본부 연구원
	공동연구원	조성현	한국해양수산개발원	항만연구본부 연구원
	공동연구원	이기열	한국해양수산개발원	항만연구본부 연구위원

연구기간 2022. 4. 11. ~ 2022. 10. 31.

보고서 집필내역

연구책임자	최석우 연구총괄, 제1장, 제2장 2절, 제3장 1절, 제4장 1절, 제5장 일부
내부연구진	이화섭 제2장 1절 일부, 제3장 4절, 제4장 4절
	김영훈 제2장 1절 일부, 제3장 2절, 제4장 2절
	조성현 제2장 1절 일부, 제3장 3절, 제4장 3절
	이기열 제5장 2절

산·학·연·정 연구자문위원	서영준 경북대학교 부교수
	이건우 한양대학교 부교수
	윤영수 해양수산부 사무관

※ 순서는 산·학·연·정 순임

발간사

2021년 우리나라의 항만물동량은 15억 8,284만 톤에 달하며, 컨테이너 물동량은 코로나19 이후 글로벌 경기회복에 힘입어 사상 처음으로 3,000만 TEU를 넘어섰다. 2010년~2021년 기간으로 확대하면 항만물동량은 연평균 2.5%, 컨테이너 물동량은 4.1% 성장하였으며, 특히 글로벌 금융위기의 여파로 경기가 침체한 2009년과 코로나19 확산으로 글로벌 물류 공급망에 충격이 가해진 2020년을 제외하면 물동량 증가 추세가 유지되고 있다. 이에 항만은 매년 증가하는 물동량에 대비하여 적절한 규모의 항만을 개발해야 하며, 여기에는 중장기 무역항 및 연안항 개발계획 수립, 부두 기능 재배치, 시설 증대 등 항만의 안정적인 개발 및 운영을 위해 정확한 항만물동량 예측이 필수적이다. 특히 개발 소요 기간 및 비용을 고려했을 때 계획수립의 근간이 되는 항만물동량 예측의 중요성은 아무리 강조해도 지나침이 없을 것이다.

항만물동량 예측의 정확도 향상을 위해서는 물동량의 변화 추이와 영향을 미칠 수 있는 대내외 환경변화 요인을 정확히 파악하고, 이를 적절히 반영해 줄 수 있는 예측 모형 설계가 무엇보다 중요하다. 특히 과거에 비해 기술 발전 속도가 빨라지고 산업의 구조적 변화를 불러오는 충격요인의 발생 빈도가 증가하는 상황은 지속적이고 빈번한 영향요인 발굴 및 예측모형 개선을 요구하고 있다. 해양수산부가 당원에 ‘항만수요예측센터’를 설치하여 물동량 모니터링과 수요예측을 실시하도록 하고, 물동량 연동 항만개발 시스템인 ‘트리거룰(Trigger Rule)’을 도입한 목적도 이러한 변화에 탄력적 대응을 기대한 바가 클 것이다.

본 연구는 자동차, 석탄, 고철, 시멘트 등 비컨테이너 주요 품목별 항만물동량 예측 모형의 발굴 및 실증분석을 실시하고, 기존 모형과 신규 모형의 예측력 비교검토를 통해 예측모형 고도화 방안을 마련하여 중장기 항만개발계획 수립 시 예측력이 높은 물동량 전망치를 제공할 수 있는 여건을 조성하고자 하였다. 이를 위해 광범위한 연구 방법론 및 선행연구 조사를 통해 항만을 중심으로 하는 물동량 예측기법을 발굴하고자 하였으며, 특히 최근에 각광받고 있는 다양한 머신러닝 기법들을 도입하고자 노력하였다. 또한 신규 모형 발굴 이후에는 교차검증법 등 모형 간 비교·검정 방법론을 활용하여 체계적인 예측모형 고도화 프로세스를 정비하고자 하였다. 최종적으로 본 연구를 통해 얻고자 하는 기대효과는 항만물동량 예측모형의 신뢰도 향상이다. 예측 프로세스

의 체계화를 통해 지속적으로 모형을 재검토하고 개선방안을 도출함으로써 모형 및 예측결과의 대내외적 신뢰도를 제고할 수 있을 것으로 기대한다.

끝으로 이 연구는 항만연구본부의 최석우 부연구위원이 책임을 맡았고, 동 본부의 이기열 연구위원과 이화섭 전문연구원, 김영훈 연구원, 조성현 연구원이 참여하였다. 아울러 이 연구 과정에서 자문과 조언을 아끼지 않았던 경북대학교 서영준 교수, 한양대학교 이건우 교수, 해양수산부 윤영수 사무관께 필자를 대신하여 감사드린다. 또한 본원 항만연구본부 김근섭 본부장께서 연구 진행 전반에 걸쳐 지적과 조언으로 보고서의 완성도를 높여 주셨다. 그 밖에 이 보고서가 출판되기까지 많은 도움을 주신 한국자동차산업협회, 한국철강협회 및 한국시멘트협회 관계자분들을 비롯한 모든 분들께도 이 지면을 통해서 심심한 감사의 말씀을 드린다.

2022년 12월
한국해양수산개발원
원장 김 종 덕

목차

요약_i

Executive Summary_ix

01	서론_1	
	제1절 연구의 배경 및 목적	1
	1. 연구 배경	1
	2. 연구 목적	4
	제2절 연구의 내용 및 방법	5
	1. 연구 내용	5
	2. 연구 방법	6
	제3절 선행연구 검토 및 차별성	9
	1. 선행연구 검토	9
	2. 선행연구와 본 연구의 차별성	13
02	항만수요예측 방법론 고찰_17	
	제1절 주요 수요예측 방법론	17
	1. 시계열 모형	17
	2. 회귀 모형	26
	3. 머신러닝 기반 모형	27
	제2절 예측모형의 적합도 비교·분석 방법론	35
	1. 주요 예측 정확도 측정 지표	35
	2. 주요 비교검증 방법론	36
	제3절 소결 및 시사점	38

03 기존 비컨테이너 항만물동량 예측모형 진단 _41

제1절 자동차 항만물동량 예측모형 진단	41
1. 국내외 자동차산업 동향 및 주요 이슈	41
2. 기존 자동차 항만물동량 예측모형 분석	49
3. 기존 예측모형의 특징 및 한계	61
제2절 석탄 항만물동량 예측모형 진단	62
1. 국내외 석탄산업 동향 및 주요 이슈	62
2. 기존 석탄 항만물동량 예측모형 분석	70
3. 기존 예측모형의 특징 및 한계	75
제3절 고철 항만물동량 예측모형 진단	77
1. 국내외 고철산업 동향 및 주요 이슈	77
2. 기존 고철 항만물동량 예측모형 분석	78
3. 기존 예측모형의 특징 및 한계	80
제4절 시멘트 항만물동량 예측모형 진단	82
1. 국내외 시멘트산업 동향 및 주요 이슈	82
2. 기존 시멘트 항만물동량 예측모형 분석	88
3. 기존 예측모형의 특징 및 한계	95
제5절 소결 및 시사점	97

04 비컨테이너 항만물동량 예측모형 고도화 방안 _99

제1절 자동차 항만물동량 예측모형 고도화	99
1. 국내 자동차 수요량 예측 정확도 개선	100
2. 수입자동차 수요 예측모형 개발	104
제2절 석탄 항만물동량 예측모형 고도화	108
1. 석탄 수요예측을 위한 신규 변수 발굴	109
2. 제철용 유연탄 수입물동량 예측모형 다각화	110
제3절 고철 항만물동량 예측모형 고도화	115

1. 국내 고철 수요 예측모형 개발	117
2. 국산 고철 공급량 예측모형 개발	119
3. 수급분석을 반영한 고철 수입물동량 예측모형 개발	124
제4절 시멘트 항만물동량 예측모형 고도화	126
1. 시멘트 수요예측을 위한 신규 변수 발굴	127
2. 주요 출하항의 단기 물동량 예측 정밀화	132
3. 지역별 지표 기반 입하 물동량 분배 규칙 개선	134
제5절 소결 및 시사점	136

05 결론 및 정책제언_139

제1절 요약 및 결론	139
제2절 정책제언	141
1. 상시적 항만물동량 예측모형 고도화 체계 구축	141
2. 예측모형 확장을 통한 데이터 기반 정책 개발 지원	142
3. 정부-연구기관 간 데이터 공유 체계 구축	143

참고문헌_145

부록_153

표 목차

〈표 1-1〉 본 연구 관련 주요 선행연구 현황	11
〈표 1-2〉 연구 차별성	15
〈표 2-1〉 주요 예측 정확도 측정 지표	35
〈표 2-2〉 주요 수요예측 방법론 및 특성	40
〈표 3-1〉 주요국 자동차산업 정책 동향	45
〈표 3-2〉 해외브랜드 수입 회귀 모형 분석결과	51
〈표 3-3〉 국내 자동차 수요 회귀 모형 분석결과	55
〈표 3-4〉 중고차 관련 주요 HS코드	57
〈표 3-5〉 우리나라 석탄소비 비중 추이	65
〈표 3-6〉 부문별 유연탄 소비 추이	66
〈표 3-7〉 제9차 전력수급기본계획의 주요 내용	67
〈표 3-8〉 국내 에너지전환 정책 발표 현황	69
〈표 3-9〉 무연탄과 유연탄의 HS코드	70
〈표 3-10〉 항만별 석탄 주요 수요처	73
〈표 3-11〉 주요 석탄 수입항 별 원단위	76
〈표 3-12〉 국내 고철 수급 현황	78
〈표 3-13〉 국내 고철 항만물동량 현황	78
〈표 3-14〉 고철 수입 물동량과 인근 항만 전기로 생산능력 관계	81
〈표 3-15〉 2021년 국가별 시멘트 생산량 및 비중	83
〈표 3-16〉 2021년 국가별 시멘트 소비량 및 비중	84
〈표 3-17〉 국내 시멘트 수출입 및 내수량	86
〈표 3-18〉 국내 주요 시멘트 제조사의 생산능력 및 가동률	87
〈표 3-19〉 건설수주액 추정 회귀분석 모형 요약	93
〈표 3-20〉 시멘트 수요 추정 회귀분석 모형 요약	93
〈표 3-21〉 품목별 예측모형 진단 결과 및 개선 방향	98
〈표 4-1〉 국내 자동차 수요 예측모형 변수 및 기초통계량	102
〈표 4-2〉 ANFIS 모형과 SLR 모형의 예측값 비교	103
〈표 4-3〉 ANFIS 모형과 SLR 모형의 예측 정확도 비교	104
〈표 4-4〉 수입 자동차 수요 예측모형 변수 및 기초통계량	105
〈표 4-5〉 VAR 모형과 SLR 모형의 예측값 비교	107

〈표 4-6〉 VAR 모형과 SLR 모형의 예측 정확도 비교	107
〈표 4-7〉 석탄 수입물동량 예측용 변수 상관분석 결과 요약	110
〈표 4-8〉 유연탄 수입물동량 SLR 분석결과 요약	111
〈표 4-9〉 유연탄 수입물동량 MLR 분석결과 요약	111
〈표 4-10〉 유연탄 수입물동량 모형별 예측 정확도 비교	112
〈표 4-11〉 광양항 유연탄 수입물동량 MLR 분석결과 요약	113
〈표 4-12〉 포항항 유연탄 수입물동량 MLR 분석결과 요약	113
〈표 4-13〉 항만별 유연탄 수입물동량 모형별 예측 정확도 비교	114
〈표 4-14〉 철강 수요 관련 변수 추이 및 전망	117
〈표 4-15〉 고철 수요예측 회귀 모형 요약	118
〈표 4-16〉 고철 수요예측 로그 추세모형 요약	118
〈표 4-17〉 자가발생 고철 발생량 추정모형 결과 요약	120
〈표 4-18〉 1인당 국내 철재 소비량 추정모형 결과 요약	121
〈표 4-19〉 국내 철강 축적량 추정모형 결과 요약	122
〈표 4-20〉 국내 철강 축적량	122
〈표 4-21〉 국내 고철 구입량 추정모형 결과 요약	123
〈표 4-22〉 고철수급 반영모형과 기존 모형의 예측값 비교	125
〈표 4-23〉 고철수급 반영모형과 기존 모형의 예측 정확도 비교	126
〈표 4-24〉 건설수주액 추정 변수 상관분석 결과 요약	128
〈표 4-25〉 총건설수주액 추정 회귀분석 모형 요약	129
〈표 4-26〉 총건설수주액 예측 결과 요약	130
〈표 4-27〉 건설수주액 추정 변수 상관분석 결과 요약	131
〈표 4-28〉 시멘트 수요 예측 회귀분석 모형 요약	131
〈표 4-29〉 시멘트 수요 예측 결과 요약	132
〈표 4-30〉 모형별 연안 출하 물동량 예측 결과	134
〈표 4-31〉 국내 주요 시멘트 제조사별 입하항	134
〈표 4-32〉 연안 입항 물동량 예측모형 결과 비교	135
〈표 4-33〉 품목별 예측모형 고도화 방안 및 개선 효과	137

그림 목차

〈그림 1-1〉 우리나라 항만물동량 추이(2010~2021년)	2
〈그림 1-2〉 연간 친환경차 내수 판매 및 수출량	3
〈그림 1-3〉 단계별 연구 내용 및 방법	8
〈그림 2-1〉 ANN의 구조	29
〈그림 2-2〉 SVM의 개념도	31
〈그림 2-3〉 LSTM 셀의 구조	32
〈그림 2-4〉 XG부스트와 랜덤포레스트	33
〈그림 2-5〉 ANFIS 구조	34
〈그림 2-6〉 k -Fold 교차검증 프로세스	37
〈그림 2-7〉 몬테카를로 교차검증 프로세스	38
〈그림 3-1〉 2010년과 2020년의 세계 지역별 자동차 생산	42
〈그림 3-2〉 2010년과 2020년의 세계 지역별 자동차 판매	43
〈그림 3-3〉 2010년과 2020년의 세계 지역별 자동차 수출	43
〈그림 3-4〉 우리나라 자동차 생산 · 내수 · 수출 추이	47
〈그림 3-5〉 자동차 물동량 예측 프로세스	49
〈그림 3-6〉 해외브랜드 수입 로그 추세 모형 분석결과	51
〈그림 3-7〉 해외브랜드 수입차 판매량 예측결과	52
〈그림 3-8〉 2022~2029년 기간 글로벌 자동차 생산 전망	54
〈그림 3-9〉 2009~2029년 기간 글로벌 자동차 생산 로그 추세 모형 분석결과	54
〈그림 3-10〉 2011~2029년 기간 국내 자동차 생산 로그 추세 모형 분석결과	55
〈그림 3-11〉 국내 자동차 수요 로그 추세 모형 분석결과	56
〈그림 3-12〉 국내 수요 예측결과	56
〈그림 3-13〉 석탄 물동량 예측 프로세스	71
〈그림 3-14〉 고철 물동량 예측 프로세스	79
〈그림 3-15〉 세계 시멘트 생산량	82
〈그림 3-16〉 국내 시멘트 생산량	86
〈그림 3-17〉 국내 시멘트 가격 인상 추이	88
〈그림 3-18〉 연간 시멘트 총 물동량	89
〈그림 3-19〉 2021년 기준 시멘트 물동량 발생 구조	89
〈그림 3-20〉 시멘트 물동량 예측 프로세스	90

〈그림 3-21〉 수출 물동량 예측 방법	92
〈그림 4-1〉 고철 수입 물동량 고도화 방안 도식	116
〈그림 4-2〉 Port-MIS와 한국철강협회의 고철 수입 물동량 차이	116
〈그림 4-3〉 고철 수요예측 모형 비교	119
〈그림 4-4〉 Logistic growth curve를 통한 국내 철재 소비량 예측	121
〈그림 4-5〉 국내 철강 축적량	122
〈그림 4-6〉 고철 공급예측 모형	123

약어 목록

- ANFIS(적응형 뉴로 퍼지 추론시스템) : Adaptive Neuro Fuzzy Inference System
- ANN(인공신경망) : Artificial Neural Network
- ARMA(자기회귀이동평균) : Autoregressive Moving Average
- ARIMA(자기회귀누적이동평균) : Autoregressive Integrated Moving Average
- CAGR(연평균증감률) : Compounded Annual Growth Rate
- CV(교차검증) : Cross-Validation
- IEA(국제에너지기구) : International Energy Agency
- LSE(최소자승추정) : Least Squares Estimation
- LSTM(장단기 메모리) : Long Short-Term Memory
- MAE(평균절대오차) : Mean Absolute Error
- MAPE(평균절대비오차) : Mean Absolute Percentage Error
- MLE(최우도추정) : Maximum Likelihood Estimation
- MSE(평균제곱오차) : Mean Squared Error
- Port-MIS(항만운영정보시스템) : Port Management Information System
- RMSE(평균제곱근오차) : Root Mean Squared Error
- SARIMA(계절형 자기회귀누적이동평균) : Seasonal Autoregressive Integrated Moving Average model
- SMA(단순이동평균) : Simple Moving Average
- SVM(서포트벡터머신) : Support Vector Machine
- VAR(벡터자기회귀) : Vector Autoregressive
- XGBoost(익스트림 경사부스팅) : Extreme Gradient Boosting

요 약

비컨테이너 항만물동량 예측모형 고도화 방안 연구(Ⅰ)

- 자동차, 석탄, 고철, 시멘트를 중심으로

최석우 · 이화섭 · 김영훈 · 조성현 · 이기열

1. 연구의 목적

- 항만물동량에 영향을 미칠 수 있는 대내외적 요인을 발굴하고 이를 지속적으로 반영하기 위해 항만물동량 예측모형 고도화가 필요함
- 품목별 연관산업의 구조적 변화, 물류환경 변화 등 대내외 환경변화는 지속적인 모니터링과 예측모형의 정확도 제고가 요구됨
- 이에 물동량 예측모형에 대한 재검토와 신규 모형 개발 및 모형의 성능을 비교·평가할 수 있는 체계 마련이 시급함
- 이에 본 연구는 품목별 항만물동량 예측방법론을 검토하고, 실증분석 및 예측력 비교검토를 통해 모형의 고도화 방안 마련을 목적으로 함
- 광범위한 연구방법론 및 선행연구 조사를 통해 항만물동량 예측기법을 발굴하고 품목별(자동차/석탄/고철/시멘트), 유형별(수입/수출/환적/연안) 차별화된 방법론을 적용한 실증분석을 실시함

-
- 교차검증법(Cross-Validation)을 적용하여 기존 모형과 신규 모형의 예측력을 비교·검토하고, 이를 토대로 비컨테이너 품목별 항만물동량 예측 모형 고도화 방안을 마련함

2. 연구의 방법 및 특징

1) 연구의 방법

- 광범위한 문헌연구를 통해 주요 수요예측 방법론 및 예측모형의 적합도 비교·분석 방법을 검토함
- 주요 수요예측 방법론은 시계열 모형, 회귀 모형, 머신러닝 모형으로 구분하고 기존 모형 및 선행연구에서 활용한 모형의 특징 및 한계를 검토함
- 모형의 적합도 비교·분석을 위해 예측 정확도 측정 지표를 검토하고, k -Fold 교차검증법 등 표본 외 검정법(out-of-sample validation)을 도입함
- 기존 비컨테이너 품목별 항만물동량 예측모형의 특징 및 한계점을 분석하고 개선방향을 도출함
- 기존 품목별 항만물동량 예측모형 진단은 ① 연관 산업의 동향 및 주요 이슈 분석, ② 기존 예측모형의 활용변수 및 방법론 검토, ③ 기존 예측모형의 특징 및 한계점 분석 순으로 진행함
- 기존 예측모형 진단 결과를 토대로 한계점을 보완할 수 있는 고도화 방안을 제시함
- 제시된 고도화 방안은 실증분석을 통해 구체화 되고, k -Fold 교차검증법으로 기존 모형 대비 개선 효과를 정량적으로 측정함

2) 연구의 특징

- 본 연구는 기존에 활용 중인 항만물동량 예측모형에 대해 정밀한 진단을 실시하였으며, 이를 토대로 예측모형 고도화 방안을 제시함
- 주요 선행연구가 일부 품목 및 항만을 대상으로 특정 예측모형의 활용 가능성 제시에 그친 반면, 본 연구는 우리나라 전체 항만의 주요 품목을 대상으로 기존 모형에 대한 정밀 진단 및 신규 모형 발굴을 실시함
- 또한 품목별 차별화된 연구방법론을 적용하고, 교차검증법을 활용하여 다양한 예측방법론에 대해 비교·검정을 실시한 것이 특징임
- 특히 표본 내 검정(in-sample validation)의 과적합 문제를 해결하고, 다양한 모형 간 적합한 비교·검토를 통해 주어진 여건하에서 최적의 모형을 제시하였음

3. 연구 결과

1) 연구 결과 요약

- 주요 수요예측 방법론 및 모형의 적합도 비교·분석 방법론 검토
- 수요예측 연구는 시계열 자료를 활용한다는 점에서 SARIMA, VAR 등 시계열 모형이 주로 활용되며, 최근에는 예측 성능에서 좋은 평가를 받는 머신러닝 모형이 많이 제안됨
- 반면 회귀 모형은 사용이 쉽고 분석 결과에 대해 직관적인 해석이 용이함
- 신규 모형 개발에는 기존 모형과 신규 모형을 비교·평가할 수 있는 기준 마련이 필요하며, 이를 통해 모형의 우수성 및 신뢰도를 확보할 수 있음

〈주요 수요예측 방법론 및 특성〉

구분	주요 특성	비고
시계열 모형	• 장기 시계열 데이터 확보가 중요 • 일반적으로 단기 예측에 우수한 성능을 보이며, 모형 결합을 통해 예측 성능 향상 가능 • 다양한 투입변수를 활용한 논리적 설명이 가능한 모형 개발에는 한계	AR, MA, ARIMA, SARIMA, VAR, Bayesian VAR, QAR 등
회귀 모형	• 방대한 변수를 활용하여 해석이 용이하고 활용성이 높음 • 다양한 변수 활용에 있어서 다중공선성 등 문제 해결 필요 • 상대적으로 낮은 예측 성능	SLR, MLR, Logistic, Probit, Tobit, Quantile 등
머신러닝 기반 모형	• 폭넓은 변수 활용성 • 상대적으로 높은 예측 성능 • 복잡한 알고리즘으로 투입변수들의 직관적인 설명력은 낮음	SVM, Decision Tree, Random Forest, ANN, XG 부스트, ANFIS 등

자료: 저자 작성

- 주요 예측 정확도 측정 지표로는 MSE, RMSE, MAE, MAPE 등이 활용됨
- 모형의 성능 측정은 표본 내 검정(in-sample test) 보다 표본 외 검정(out-of-sample test)이 적합하며, 교차검증법(cross-validation)을 활용한 정밀 검증이 필요함
- 본 연구에서는 k-Fold 교차검증법과 몬테카를로 교차검증법을 검토함
- 기존 비컨테이너 항만물동량 예측모형 진단
 - 기존 품목별 항만물동량 예측모형 진단은 ① 연관 산업의 동향 및 주요 이슈 분석, ② 기존 예측모형의 활용 변수 및 방법론 검토, ③ 기존 예측모형의 특징 및 한계점 분석 순으로 진행함
 - (자동차) 연이은 공급망 위기 등 악재가 지속되면서 국내 자동차 공급 및 수요에 변화가 발생하고 있으며, 방법론적으로 대내외 환경변화를 반영할 수 있는 모형 개발이 필요함
 - (석탄) 타 품목에 비해 수요처가 명확하고 시설의 증설 및 감축이 발생하지 않는 한 매년 발생하는 수요가 비교적 일정한 것이 특징이나, 향후 국내외 에너지전환 정책이 가속화될 경우를 대비한 모형 개발이 필요함

- (고철) 철강산업에 대한 규제가 강화됨에 따라 고철 수요가 급격히 증가하고 있어, 국내 고철 자급률을 반영한 면밀한 고철 수급분석을 통해 고철 수입물동량 예측 정확도 제고가 필요함
- (시멘트) 시멘트 수요 기반 전국 물동량 추정을 위해 추가 설명변수 발굴 및 계절성과 추세를 반영한 연안 출하 비중 산정이 필요하며, 입하 물동량 분배 시 지역 수요 기반 비중 산정 방식 개선이 필요함

〈품목별 예측모형 진단 결과 및 개선 방향〉

품목	진단 결과	개선 방향
자동차	• 자동차 수요예측 시 단순선형회귀 모형 사용	• 대내외 환경변화를 반영하는 모형 개발을 통해 예측 정확도 제고
	• 수입물동량 예측 시 최근의 추세 반영 미흡	• 시계열 모형과 비교검토를 통해 최근 추세를 반영한 예측모형 개발
석탄	• 투입변수의 다양성 부족	• 상관분석을 통한 투입변수 발굴
	• 원단위법 및 단순 추세분석에 의존	• 다양한 분석방법론 도입으로 비교·검증 및 예측 정확도 개선
고철	• 기존 고철 수요예측 모형의 예측 정확도 저하	• 추가 설명변수 발굴 및 예측모형 개발
	• 국내 고철 자급률 반영 필요	• 국내 고철 수급분석을 통한 수입 물동량 예측 정확도 개선
시멘트	• 경험에 의한 변수 선정	• 상관분석을 통한 변수 선정
	• 단기 예측 결과에 따른 장기 예측치 변화가 큼	• 머신러닝을 적용한 단기 물동량 예측
	• 입하 물동량 분배 시 전년도 비중 사용	• 지역 수요 기반 입하 물동량 분배

자료: 저자 작성

■ 비컨테이너 항만물동량 예측모형 고도화 방안

- 기존 예측모형 진단 결과를 토대로 한계점을 보완할 수 있는 고도화 방안을 제시하고, 실증분석 및 교차검증으로 기존 모형 대비 개선 효과를 정량적으로 측정함
- (자동차) 머신러닝 및 시계열 분석기법을 도입하여 국내 자동차 및 수입자동차 수요 예측모형을 개선함

- (석탄) 석탄 물동량의 변동성이 커질 것을 대비하여 상관분석을 통해 신규 설명변수를 발굴하고, MLR 모형을 이용하여 제철용 유연탄 수입물동량 예측모형의 정확도를 개선함
- (고철) 증가하고 있는 국내 자급율을 반영한 수급분석 모형을 개발하고, 이를 통해 고철 수입물동량 예측모형의 정확도를 개선함
- (시멘트) 시멘트 수요에 영향을 미치는 신규 설명변수를 발굴하고, 이를 이용하여 주요 출하항의 단기 물동량 예측 및 입하 물동량 분배 규칙을 개선함

〈품목별 예측모형 고도화 방안 및 개선 효과〉

품목	고도화 방안	개선 효과
자동차	• 국내 자동차 수요량 예측 정확도 개선	• ANFIS 모형 도입 시 기존 모형 대비 예측 정확도 2.7%p 개선
	• 수입자동차 수요 예측모형 개발	• VAR 모형 도입 시 기존 모형 대비 예측 정확도 5.5%p 개선
석탄	• 석탄 수요예측을 위한 신규 변수 발굴	• 상관분석을 통해 중앙은행 정책금리, 실질 거래 환율 등 신규 변수 발굴
	• 제철용 유연탄 수입물동량 예측모형 다각화	• MLR 모형 도입으로 기존 모형 대비 예측 정확도 1.3%p 개선
고철	• 국내 고철 수요 예측모형 개발	• 로그 추세모형으로 국내 고철 수요 예측 모형 개발
	• 국산 고철 공급량 예측모형 개발	• 고철 발생 방식에 따라 특화된 국산 고철 공급량 예측모형 개발
	• 수급분석을 반영한 고철 수입물동량 예측 모형 개발	• 수급분석 반영 모형 개발을 통해 예측 정확도 4.9%p 개선
시멘트	• 시멘트 수요예측을 위한 신규 변수 발굴	• 신규변수 발굴을 통해 기존 모형 대비 예측 정확도 5.4%p 개선
	• 주요 출하항의 단기 물동량 예측 정밀화	• XG부스트 모형 도입 시 기존 모형 대비 예측 정확도 1.1%p 개선
	• 지역별 지표 기반 입하 물동량 분배 규칙 개선	• 지역별 인구수 반영 및 SMA 보정 시 기존 모형 대비 예측 정확도 14.5%p 개선

자료: 저자 작성

2) 정책대안 제시내용 및 정책화 활동

- 상시적 항만물동량 예측모형 고도화 체계 구축
 - 항만 수요예측 고도화 사업을 상시적·전문적으로 추진할 수 있는 체계 구축을 제안함
 - 고도화 사업의 내용적 범위는 기존 예측모형의 검증 및 신규 모형 개발을 포함하며, 이를 위해 방대한 양의 데이터를 효율적으로 수집·분석할 수 있는 시스템과 통계분석 툴 및 전문인력 확보가 필요함
- 예측모형 확장을 통한 데이터 기반 정책 개발 지원
 - 기존 증장기 물동량 예측모형 고도화뿐만 아니라 단기적 관점에서의 물동량 예측 및 외부 충격에 따른 영향 추정모형 개발이 필요함
 - 월별 혹은 분기별 자료를 이용한 단기 물동량 예측에는 머신러닝 기반 예측모형의 활용도가 높을 것으로 예상됨
- 정부-연구기관 간 데이터 공유 체계 구축
 - Port-MIS와 항만수요예측센터 간 데이터 공유 체계 구축을 통해 고빈도 대용량 데이터의 활용도 제고가 필요함
 - 공유된 데이터를 항만수요예측센터에서 지속적으로 모니터링하고 시의적절하게 분석결과를 정부에 제공함으로써 공공데이터의 활용도와 정책 연구의 품질을 향상시킬 수 있음

3) 정책적 기여 등 기대효과

- 항만물동량 예측 정확도 제고 및 상시예측 시스템 구축에 활용
 - 다양한 예측모형 및 영향요인 발굴을 통해 항만물동량 예측모형의 정확도를 개선하고 상시적인 예측시스템 구축에 활용함

-
- 단기/중장기 항만물동량 전망자료 주기적인 제공
 - 정책담당자, 항만당국, 부두 운영사 등 항만 개발 및 운영 효율성 강화
주체들에게 신뢰성 높은 항만물동량 전망자료의 주기적 제공을 기대함

EXECUTIVE SUMMARY

A Study on the Advancement of the Non-container Port Throughput Forecasting Model (I)

**– Focusing on Automobile, Coal, Scrap Metal
and Cement**

Seokwoo Choi · Hwaseop Lee · Younghoon Kim · SeongHyun Cho
· Kiyoul Lee

1. Purpose

- It is necessary to advance port throughput forecasting models in order to discover and continuously reflect internal and external factors affecting the port throughput.
- Continuous monitoring and improvement in the accuracy of forecasting models are required due to internal and external changes in circumstances, including structural and logistical changes in related industries by item.
- Accordingly, it is urgent to revisit port throughput forecasting models, develop a new model and prepare a system to compare /evaluate the performance of the models.

-
- Therefore, this study aims to review methodologies for forecasting port throughput by item and develop a plan for advancing port throughput forecasting models through empirical analysis and the comparative review of prediction power.
 - Through the investigation of extensive research methodologies and preceding studies, this study discovered port throughput forecasting methods, while conducting an empirical analysis by applying methodologies differentiated to items (automobile/ coal /scrap metal/cement) and types (import/export/transshipment /coast).
 - With the application of cross-validation, the predictability of existing and new models were compared and reviewed. Based on this result, this study prepared measures for advancing port throughput forecasting models by non-container item.

2. Methodology and Features

1) Methodology

- A broad array of literature was studied to review major demand forecasting methodologies as well as methods for comparing and analyzing the goodness of fit of forecasting models.
- With major demand forecasting methodologies divided into time-series, regression and machine learning models, features and limitations of existing models and the ones used in preceding studies were reviewed.

- In order to compare and analyze the goodness of the fit of the model, this study reviewed indicators for measuring the prediction accuracy, while employing out-of-sample validation such as k-Fold cross-validation etc.
- The features and limitations of existing port throughput forecasting models were analyzed by non-container item, and improvement measures were developed.
- Diagnosis on existing item-specific port throughput forecasting models was conducted in the following order: ① analyzing the trend and major issues of related industries; ② reviewing variables and methodologies utilized in existing forecasting models; and ③ analyzing the features and limitations of existing forecasting models.
- Based on the diagnosis result of existing forecasting models, advancement measures were presented to make up for the limitations.
- The advancement measures proposed by this study were materialized through an empirical analysis, and the k-Fold cross-validation method quantitatively measured the improvement effect compared to existing models.

2) Features

- This study carried out precise diagnosis on existing port throughput forecasting models and accordingly, presented measures to advance and upgrade the forecasting models.

-
- While major preceding studies were limited to presenting the possibility of utilizing specific forecasting models targeting certain items and ports, this study conducted precise diagnosis on existing models for major items of all ports across the country and developed a new model.
 - In addition, this study carried out comparison as well as validation for a variety of forecasting models by applying methodologies differentiated by items and employing cross-validation.
 - Specifically, it solved the problems of overfitting in in-sample validation while presenting an optimized model under given condition by carrying out comparison and review befitting a variety of models.

3. Result

1) Summary

- Review on major demand forecasting methodologies as well as comparison and analysis methodologies for goodness of fit of a model
- The demand forecasting study mainly utilized time-series models such as SARIMA and VAR in that time-series data were employed. In recent days, machine learning models which are well evaluated in terms of prediction performance are highly suggested.

- On the other hand, regression models are easy to use and readily allow intuitive interpretation of analysis results.
- In developing new models, it is necessary to prepare the criteria which allow comparison and evaluation between existing models and new models, which, in turn, will help secure the excellence and reliability of the models.

〈Major demand forecasting methodologies and features〉

Classification	Major features	Remarks
Time-series model	• Important to secure long-term time-series data • Showing great performance in short-term forecasting in general and able to improve the predictability through the combination of models • Posing limitations in developing a model available for logical explanation with a variety of input variables	AR, MA, ARIMA, SARIMA, VAR, Bayesian VAR, QAR etc.
Regression model	• Easy to interpret and highly applicable by utilizing a wide array of variables • Need to solve the problem of multicollinearity in utilizing a variety of variables • A relatively low predictability	SLR, MLR, Logistic, Probit, Tobit, Quantile etc.
Machine learning-based model	• Able to apply a broad array of variables • A relatively high predictability • Input variables have low intuitive explanation power due to complex algorithms	SVM, Decision Tree, Random Forest, ANN, XGBoost, ANFIS etc.

Source: Prepared by the authors

- MSE, RMSE, MAE, MAPE, etc. were used as indicators for measuring prediction accuracy.
- For measuring the performance of models, out-of-sample testing, rather than in-sample testing, is suitable while precise validation through cross-validation is required.

-
- In this study, the k-Fold and Monte Carlo cross-validation methods were reviewed.
 - **Diagnosis on existing non-container port throughput forecasting models**
 - Existing item-specific port throughput forecasting models were diagnosed in the following order: ① analyzing the trend and major issues of related industries; ② reviewing variables and methodologies utilized in existing forecasting models; and ③ analyzing the features and limitations of existing forecasting models.
 - (Automobile) As unfavorable conditions such as a series of supply chain disruptions have continued, the supply and demand of domestic automobiles are changing and hence, it is necessary to develop a model which can reflect internal and external changes from a methodological perspective.
 - (Coal) The source of demand is clearer than other items and the annual demand is relatively constant unless facilities are expanded and reduced. However, it is required to develop a model in case domestic and overseas policies for energy conversion would be accelerated in the future.
 - (Scrap metal) Since the demand for scrap metal has been rapidly increasing with strengthened regulations on the steel industry, it is necessary to improve the accuracy of forecasting scrap metal import throughput through a thorough analysis on the supply and demand of scrap metal reflecting its self-sufficiency rate.
 - (Cement) For forecasting the nationwide throughput based on cement demand, it is required to discover additional explanatory

variables, and calculate the share of coastal shipment by reflecting seasonal changes and trends. In addition, a regional demand-based calculation method should be improved in distributing incoming throughput.

〈Diagnosis results and improvement directions of forecasting models per item〉

Item	Diagnosis result	Improvement direction
Automobile	Utilizing simple linear regression models in forecasting the demand of automobiles	Improving the forecasting accuracy by developing a model that reflects internal and external environmental changes
	Insufficient to reflect recent trends in forecasting import throughput	Developing a forecasting model that reflects recent trends through cross-check with time-series models
Coal	Lack of diversity in input variables	Discovering input variables through correlation analysis
	Relying on unit load method and trend analysis	Improving the accuracy of forecasting, comparison and verification by adopting a variety of analysis methodologies
Scrap metal	Declining prediction accuracy of existing scrap metal demand forecasting models	Discovering additional explanatory variables and developing a forecasting model
	Need to reflect domestic self-sufficiency rate of scrap metal	Improving the accuracy of forecasting import throughput through analysis on the supply and demand of scrap metal
Cement	Selecting experience-based variables	Selecting variables based on correlation analysis
	Long-term forecasts largely varies depending on short-term forecasting results	Forecasting the short-term throughput based on machine learning
	Using the figure of previous year when distributing incoming throughput	Distributing regional demand-based incoming throughput

Source: Prepared by the authors

-
- Measures for advancing non-container port throughput forecasting models
 - This study presented advancement measures to make up for the limitations of existing forecasting models based on the diagnosis results, and quantitatively measured improvement effects compared to existing models through empirical analysis and cross-validation.
 - (Automobile) Demand forecasting models for domestic and imported automobile were improved with machine learning and time-series analysis methodologies.
 - (Coal) In preparation for an increase in the volatility of coal throughput, new explanatory variables were discovered through correlation analysis. The accuracy of the forecasting model for bituminous coal import throughput for steelmaking was improved with MLR model.
 - (Scrap metal) The supply and demand analysis model was developed by reflecting the increasing domestic self-sufficiency rate of scrap metal, which in turn improved the accuracy of import throughput forecasting model for scrap metal.
 - (Cement) After discovering new explanatory variables affecting the cement demand, this study used them to forecast short-term throughput by major port of shipment and improve incoming throughput distribution rules.

〈Measures for advancing forecasting models per item and improvement effects〉

Item	Advancement measures	Improvement effect
Automobile	Improving the accuracy of domestic automobile demand	The forecasting accuracy can be improved by 2.7%p compare to existing models when introducing an ANFIS model
	Developing a demand forecasting model of imported automobile	The forecasting accuracy can be improved by 5.5%p compare to existing models when introducing a VAR model
Coal	Discovering new variables for forecasting coal demand	Discovering new variables such as the central bank policy rate and the real exchange rate
	Diversifying the forecasting model of bituminous coal import throughput	The forecasting accuracy can be improved by 1.3%p compare to existing models with the adoption of a MLR model
Scrap metal	Developing a forecasting model for the demand of domestic scrap metal	Developing the forecasting model for the demand of domestic scrap metal with log-linear trend model
	Developing a forecasting model for the supply of domestic scrap metal	Developing the forecasting model for the supply of domestic scrap metal specialized in accordance with scrap metal development
	Developing a forecasting model for the import of scrap metal that reflects the demand and supply analysis	The forecasting accuracy can be improved by 4.9%p by developing a model reflecting the demand and supply analysis
Cement	Discovering new variables for forecasting the demand of cement	The forecasting accuracy can be improved by 5.4%p compared to existing models by discovering new variables
	Refining the accuracy of short-term throughput of major ports of shipment	The forecasting accuracy can be improved by 1.1%p compare to existing models when introducing an XGBoost model
	Improving the distribution rules of regional indicator-based incoming throughput	The forecasting accuracy can be improved by 14.5%p compare to existing models when reflecting population by region and refining SMA

Source: Prepared by the authors

2) Policy suggestions and policy-making activities

- Establishing a regular advancement system for port throughput forecasting models
 - This study suggested the establishment of a system which can carry out a project to advance demand forecasting in a regular as well as a professional manner.
 - The content scope of the advancement project includes the validation of the existing forecasting models and development of a new model. To this end, it is necessary to secure a system that can efficiently collect and analyze a vast amount of data, statistical analysis tools and professionals.
- Supporting the development of data-based policies through the expansion of forecasting models
 - In addition to the advancement of existing mid-and long-term throughput forecasting models, it is required to develop a model for forecasting throughput from a short-term perspective and estimating the effect of external shocks.
 - Machine learning-based forecasting models are expected to be highly utilized for short-term throughput forecasting which uses monthly or quarterly data.
- Formulating a data-sharing system between the government and research institutions
 - It is necessary to improve the utilization of high-frequency, massive data through the establishment of a data-sharing system between Port-MIS and the Port Demand Analysis Center.

- If the Port Demand Analysis Center carries out continuous monitoring on shared data and provide the government with their analysis results in a timely manner, it would improve the utilization of public data and the quality of policy studies.

3) Expected benefits including policy contribution

- This study can be used to improve the accuracy of port throughput forecasting and establish a regular forecasting system.
- This study can be used to improve the accuracy of port throughput forecasting models by discovering a variety of forecasting models and influence factors while establishing a regular forecasting system.
- The short-, mid-, to long-term port throughput forecasting data can be provided on a regular basis.
- Highly reliable port throughput forecasting data can be provided to those in charge of developing ports and strengthening operating efficiency including policymakers, port authorities and port operators.

01

서론

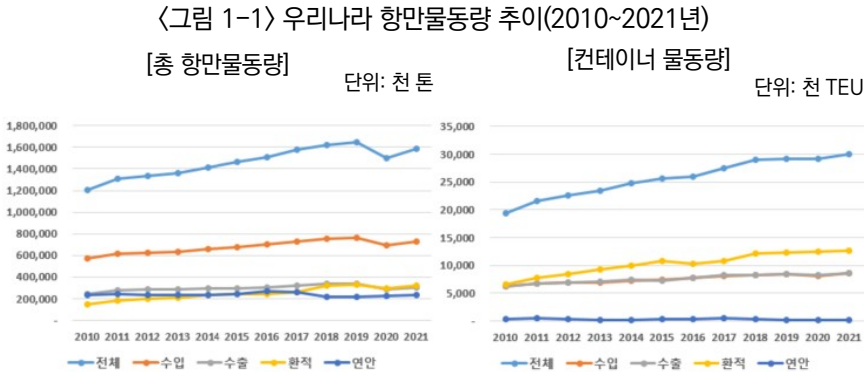
제1절 연구의 배경 및 목적

1. 연구 배경

2021년 기준 우리나라의 전체 항만물동량은 15억 8,284만 톤에 이르고 있으며, 컨테이너 물동량은 코로나19 이후 글로벌 경기회복에 힘입어 사상 처음으로 3,000만 TEU를 넘어섰다. 또한 2010~2021년 기간 전체 항만 물동량은 연평균 2.5%, 컨테이너 물동량은 4.1% 성장하여 동 기간 우리나라 경제성장률 4.2%와 근접한 성장세를 보였다. 특히 글로벌 금융위기의 여파로 경기가 침체한 2009년과 코로나19 확산으로 글로벌 물류 공급망에 충격이 가해진 2020년을 제외하면 매년 물동량이 증가하고 있다.

이에 항만은 물동량을 처리하기 위한 항만시설 확보가 필수적이며, 매년 증가하는 물동량에 대비하여 적절한 규모의 항만개발을 위해 정부 재정과 민간자본이 투입된다. 중장기 무역항 및 연안항 개발계획 수립, 부두기능 재배치, 시설 증대 등 항만의 안정적인 개발 및 운영을 위해 정확한 항만물동량 예측이 수반되어야 함은 물론이다. 특히 개발 소요 기간 및 비용을

고려했을 때 계획수립의 근간이 되는 항만물동량 수요예측의 중요성은 아무리 강조해도 지나침이 없을 것이다.

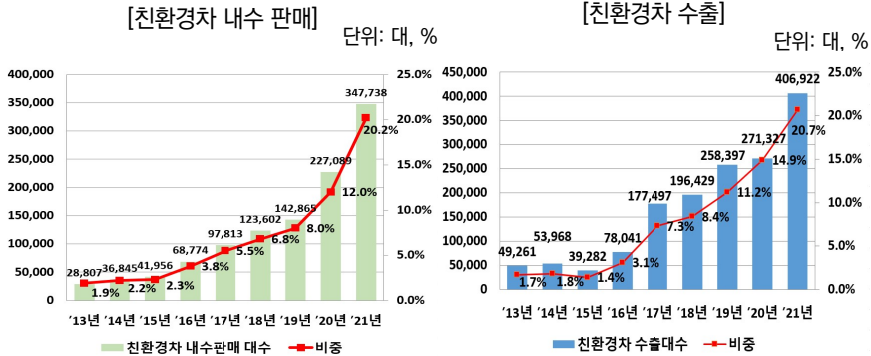


자료: 해양수산부 Port-MIS(검색일: 2022.9.12.)

한편 항만물동량 수요예측 정확도 향상을 위해서는 물동량의 변화 추이와 영향을 미칠 수 있는 대내외 환경변화 요인을 정확히 파악하고 이를 적절히 반영해 줄 수 있는 예측모형 설계가 무엇보다 중요하다. 특히 과거에 비해 기술 발전 속도가 빨라지고 산업의 구조적 변화를 불러오는 충격요인의 발생 빈도가 증가하는 상황은 지속적이고 빈번한 영향요인 발굴 및 예측모형 개선을 요구하고 있다.

주요 비컨테이너 품목 중 하나인 자동차의 경우 대내적으로 수입차 및 친환경차 구매가 증가하고 있으며, 대외적으로 동유럽 분쟁, 반도체 공급부족 등 공급충격 요인이 발생함과 동시에 수소·전기차, 자율주행차 등 최첨단 기술 도입에 따라 산업의 패러다임이 변화 중이다. 이에 따라 국내 수입차 판매량은 2021년 기준 27만 6천 대로 2003년 이후 연평균 15.9% 증가하고 있으며, 친환경차의 내수 판매량은 2021년 기준 34만 8천 대를 기록하며 전년대비 54.5% 증가하고 수출 판매량은 40만 7천 대로 전년대비 50.0% 증가하였다.

〈그림 1-2〉 연간 친환경차 내수 판매 및 수출량



자료: 산업통상자원부 보도자료(2022.01.18.)

또 다른 주요 품목들인 석탄과 고철의 경우 석탄이 수입에 의존하는 반면 고철은 국내 자급률이 80% 이상으로 서로 상반된 산업 구조를 가지고 있다. 하지만 정부의 2050 탄소중립 시나리오에 석탄발전 중단, 철강공정에 수소환원제철 방식 도입 등이 포함됨에 따라 석탄과 고철의 수요와 공급이 크게 변화할 것으로 예상된다. 이밖에 시멘트도 정부의 민간택지 분양가 상한제 시행, 역대 최대 SOC 투자, 원자재 및 유연탄 가격 상승 등 건설업 여건 변화에 따라 생산량 및 재고 출하량 변동으로 물동량 예측에 불확실성이 가중되고 있다.

이처럼 품목마다 관련 산업의 구조적 변화, 대내외 환경 변화 등이 나타나고 있으며, 이는 항만물동량에 직·간접적 영향을 미치고 있다. 이에 품목별 항만물동량에 영향을 미치는 대내외적 요인을 지속적으로 발굴 및 반영하고 중장기적으로 항만개발계획까지 효율적으로 운용하기 위해서 항만물동량 예측모형의 고도화가 필요하다. 특히 글로벌 물류에서 핵심적인 연결점 역할을 하는 항만의 물동량 변화를 정확히 인지하는 과정은 단순히 물동량의 증감을 확인하는데 그치지 않고 우리나라 산업의 구조적 변화와 중장기적 정책 환경의 변화 모습을 추론할 수 있게 한다는 점에서 매우 중요하다.

2. 연구 목적

본 연구는 자동차, 석탄, 고철, 시멘트 등 비컨테이너 품목별 항만물동량의 예측모형 발굴 및 실증분석을 목적으로 한다. 궁극적으로는 기존 모형과 신규 모형의 예측력 비교검토를 통해 예측모형 고도화 방안을 마련하여 중장기 항만개발계획 수립 시 예측력이 높은 물동량 전망치를 제공할 수 있는 여건을 조성하고자 한다.

이를 위해 우선 광범위한 연구방법론 및 선행연구 조사를 통해 항만을 중심으로 하는 물동량 예측 기법을 발굴하고자 한다. 이는 대부분의 기존 연구들이 특정 산업 또는 품목의 수요와 공급을 전망하거나 기종점 간 해상물동량을 예측하는 데 중점을 두었다면, 본 연구는 항만의 안정적인 개발 및 운영을 최우선 가치로 두고 항만을 의사결정주체로 하는 물동량 예측 방법론 개발을 원칙으로 삼았다. 이어서 품목별(자동차/석탄/고철/시멘트 등), 유형별(수입/수출/환적/연안) 차별화된 방법론을 적용하고 실증분석을 실시하고자 한다.

또한 신규모형을 적용한 예측치 산출 이후에는 교차검증(Cross-Validation, CV) 등 검증된 모형 간 비교·검정 방법론을 활용하여 기존 모형과 신규 모형의 예측력 비교검토를 통해 예측모형 고도화를 위한 방안을 제시하고자 한다. 다만 중장기 항만물동량 예측에 있어서는 기 수립된 항만 및 연관 산업시설의 개발계획을 반영하는 등 과거 데이터에 기반한 모형의 예측력 비교만으로 모형의 성능을 비교하기에는 한계가 있다. 이에 데이터 기반 모형의 예측력뿐만 아니라 모형의 범용성과 효율성 등을 다각적으로 검토하여 예측모형의 고도화 방안을 마련하고자 한다.

최종적으로 본 연구를 통해 얻고자 하는 기대효과는 항만물동량 예측모형의 신뢰도 향상이다. 예측 프로세스의 체계화를 통해 지속적으로 모형을

재검토하고 개선방안을 도출함으로써 모형 및 예측결과의 대내외적 신뢰도를 제고할 수 있을 것으로 기대한다.

제2절 연구의 내용 및 방법

1. 연구 내용

본 연구는 총 다섯 개의 장으로 구성되어 있다. 제1장은 본 연구의 배경과 목적, 연구의 내용 및 방법, 선행연구 검토 및 연구의 차별성 등 연구의 개요를 기술하였다.

제2장에서는 다양한 분야의 수요예측에 활용되는 방법론을 모형의 특성별로 구분하여 연구방법론을 검토하였다. 일반적으로 수요예측에는 데이터의 종속성을 반영한 시계열 모형이 많이 활용되며, 수요에 영향을 미치는 다양한 외부 요인들의 영향을 반영하기 위해 회귀 모형을 활용하기도 한다. 최근에는 과거에 비해 가용한 데이터의 종류와 양이 급격히 증가하면서 다양한 머신러닝 기반의 모형도 제안되고 있다. 이에 본 장에서는 각 모형의 특징 및 활용분야를 분석하여 항만물동량 수요예측에 적용이 가능한 최적의 방법론을 도출하고자 하였다.

제3장에서는 신규 모형 개발에 앞서 기존 품목별 항만물동량 예측모형의 특징 및 한계점을 진단하였다. 분석대상 품목은 한국해양수산개발원 항만수요예측센터에서 매년 수요예측을 진행 중인 14개 주요 품목¹⁾ 중 연구

1) 14개 주요 품목은 컨테이너와 비컨테이너 13개 품목(양곡, 시멘트, 모래, 석탄, 철광석, 기타광석, 화학공업생산물, 목재, 고철, 철재, 자동차, 잡화)로 구성됨

기간 및 예산 등을 고려하여 자동차, 석탄, 고철, 시멘트 등 4개 품목을 선정하였다. 각 품목별 항만물동량 예측모형을 진단하는 과정은 연관 산업의 동향 및 주요 영향요인을 분석하고 기존 예측모형과 활용 변수 및 관련 시설 개발계획이 포함된 법정계획 등을 검토하는 것이 포함되었다. 이를 통해 기존 예측모형의 특징과 한계점을 도출하고자 하였다.

제4장에서는 앞서 3장에서 진단한 품목별 예측모형의 한계점을 보완하여 모형의 예측력 및 활용도를 향상시킬 수 있는 고도화 방안을 제시하였다. 각 품목의 특성 및 가용한 데이터의 양과 질을 고려하여 적절한 예측모형을 선별하였으며, 실증분석을 통해 예측치를 전망하였다. 또한 교차검증법 등을 활용하여 기존 모형과 신규 모형의 예측력을 계량적으로 비교·검토하였으며, 최종적으로 모형의 적합도, 활용도, 효율성 등을 기반으로 모형의 고도화 방안을 제시하였다.

마지막 제5장은 본 연구의 결과를 요약하고 한계점 및 향후 계획 등을 제시하였다. 특히 항만물동량 예측에 있어서 다양한 연구방법론을 비교·검토하여 점진적인 예측모형 고도화를 위한 체계를 구축하고 예측 결과의 정책적 활용성을 제고하기 위한 방안을 제시하였다. 이를 위해 해양수산부와 항만공사 등 정책수요자와 각 연관 산업별 전문가 및 이해 관계자의 의견을 최대한 수렴하였다.

2. 연구 방법

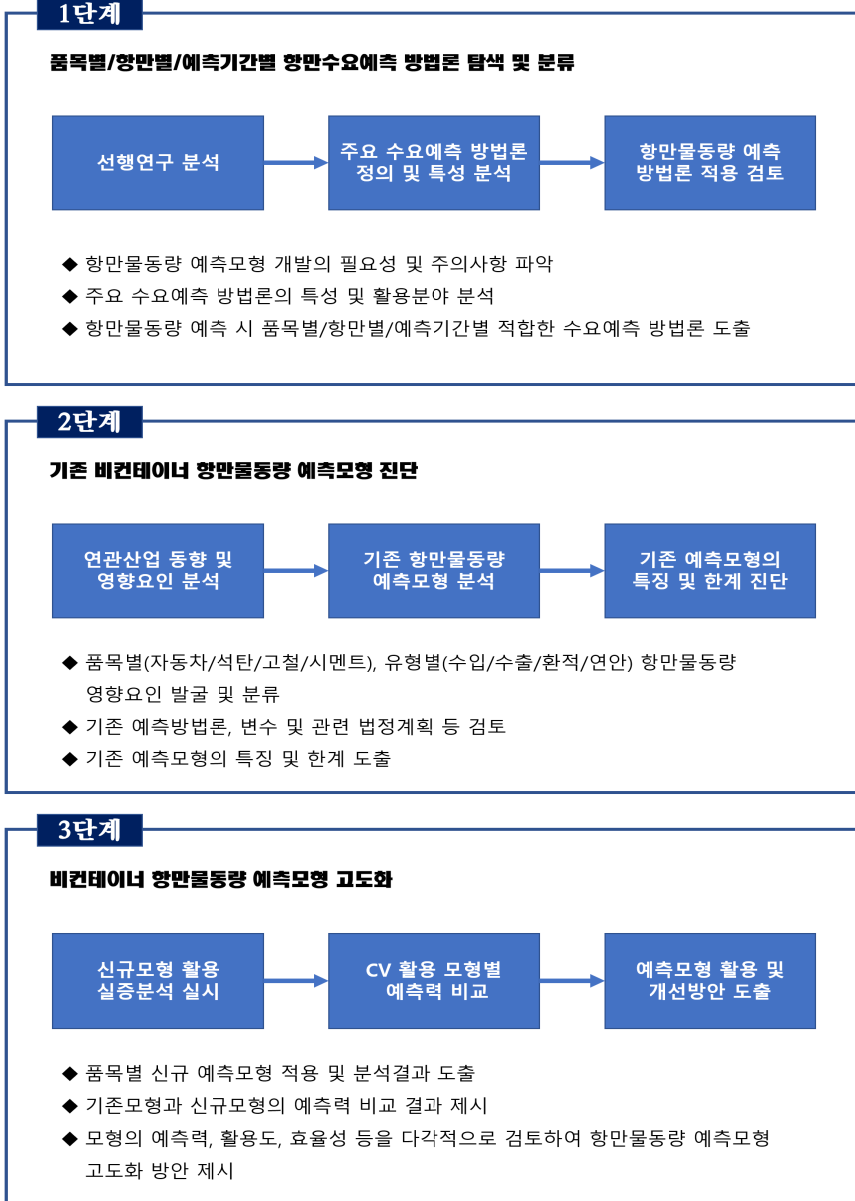
본 연구의 주요 연구 방법은 <그림 1-3>과 같이 크게 3단계로 구분된다. 먼저 1단계에서는 국내외 선행연구 분석을 통해 주요 수요예측 방법론을 탐색하고 특성을 분석한다. 이를 토대로 품목별, 항만별, 예측기간별 항만물동량 예측 방법론을 탐색하고 분류한다. 문헌자료는 다양한 보고서 및

학술논문을 망라하여 수집하였으며, 대상 품목과 예측모형을 기준으로 분류하여 검토하였다. 또한 품목의 특성과 입수 가능한 자료의 범위 등을 고려해 적합한 방법론을 도출하기 위해 전문가 자문 및 연구진 브레인스토밍을 활용하였다.

2단계에서는 품목별 연관 산업 관련 각종 보고서, 언론 기사, 법정 계획 등을 검토하여 산업 동향을 파악하였으며, 이를 토대로 각 품목의 항만물동량에 영향을 미칠 수 있는 요인을 발굴하였다. 특히 수·출입, 환적, 연안 등 물동량의 유형별로 구분하여 영향요인을 분류하였다. 이어서 기존 항만물동량 예측모형을 분석하여 품목별, 유형별 예측모형과 활용된 변수를 검토하였으며, 각 모형의 특징과 한계점을 도출하였다.

마지막으로 3단계에서는 2단계에서 도출한 기존 예측모형의 한계점을 보완하여 예측력을 향상시킬 수 있는 모형을 1단계의 예비모형 중에서 선정하였다. 일차적인 모형의 선정은 연구진의 해당 품목과 모형에 대한 학습을 통해 선별하였으며, 선정된 모형은 계량분석 전문가와 품목의 연관 산업 전문가 자문을 통해 구체화되었다. 선정된 모형을 적용한 실증분석에는 최대한 국내외 기관의 공신력 있는 데이터를 수집하여 분석에 활용하였으며, R, SAS, MATLAB, PYTHON 등 통계패키지를 활용하여 모형 구동 및 모형 간 예측력 검정을 실시하였다. 최종적으로 수립된 항만물동량 예측모형 고도화 방안은 품목별 수요예측 자문위원회를 통해 수정·보완되었다.

〈그림 1-3〉 단계별 연구 내용 및 방법



자료: 저자 작성

제3절 선행연구 검토 및 차별성

1. 선행연구 검토

중장기 비컨테이너 항만물동량 예측모형 개발을 주제로 연구한 국내외 사례는 많지 않다. 이에 본 절에서는 검토하고자 하는 선행연구의 범위에 컨테이너 물동량 예측을 포함시켰으며, 연도별 중장기 예측뿐만 아니라 단기 또는 이벤트 발생 시 영향 정도 추정 연구도 검토하였다. 방법론 측면에서도 일반적으로 예측모형으로 많이 활용되는 시계열 또는 회귀 모형 외에 최근 많이 활용되는 머신러닝 기반의 다양한 기법들을 활용한 연구를 최대한 검토하고자 하였다. 본 연구에서 검토한 주요 선행연구 현황은 다음의 <표 1-1>과 같으며, 전체 선행연구 현황은 <부록 1>에서 확인할 수 있다.

먼저 국내 연구를 살펴보면, 전찬영·송주미(2006)가 인공신경망(Artificial Neural Network, ANN) 모형을 활용하여 주요 수출입 품목에 대해 항만 물동량을 예측한 바 있으며, 하준수 외(2021), 민경창·하현구(2014), 신창훈(2011), 김두환·이강배(2020) 등은 다양한 시계열 모형 또는 데이터 마이닝 기법 등을 활용하여 항만별 컨테이너 물동량을 전망하였다. 주요 품목별 수요예측 연구로는 자동차의 경우 박상수·김경유·윤자영(2017)이 확률적 계수 로짓모형을 통해 우리나라 자동차산업의 수요 및 공급합수를 추정하였으며, 석탄 관련 연구로는 장원익(2010), 조상호 외(2020) 등이 회귀 모형을 활용하여 철광석 및 석탄 물동량을 예측한 바 있다. 또한 시멘트 물동량 관련 연구는 서홍용(2014)이 전국 30개 무역항의 시멘트 물동량을 대상으로 요인분석 및 네트워크 분석을 실시하였으며, 조찬혁(2008)은 지수평활모형 및 ARIMA 모형으로 동해항의 시멘트 물동량을 추정한 바 있다.

해외 연구의 경우에도 국내 연구에서와 마찬가지로 개별 연구자들이 각자의 연구 목적에 따라 제한적인 범위에서 연구를 수행한 사례가 대부분이다. Dragan et al.(2021), Farhan & Ong(2018) 등은 ARIMA 기반의 시계열 모형을 활용하여 특정 항만의 품목별 물동량 예측모형을 개발하였으며, Jiang et al.(2019)는 베이지안 회귀 모형을 통해 중국 연안항만의 품목별 물동량을 예측하였다. 이밖에도 VAR 모형을 활용한 Gao, Luo, & Zou(2016), 인공신경망 모형을 활용한 Lam et al.(2004), Grey 모형을 활용한 Zhang, Huang, & Zhao(2013) 등이 항만물동량 예측 관련 연구를 수행하였다. 품목별 수요예측 연구로는 Islam et al.(2016), Zhang et al.(2017), Vahahi, Hosseininia, & Alborzi(2016) 등이 시계열 모형을 활용하여 자동차 판매량 예측 연구를 수행하였으며, Chan & Lee(1997), Benalcazar, Krawczyk, & Kamiński(2017), Li & Li(2019) 등은 석탄 소비량을 예측하였다. 또한 Oda, Akimoto, & Tomoda(2013), Pauliuk et al.(2013), Dworak & Fellner(2021) 등은 고철 수요량 예측 연구를 수행하였으며, Fiskin & Cerit(2019), Fradinata, Suthummanon, & Suntiamorntut(2015), Liu et al.(2008) 등은 시멘트 수요 예측 연구를 수행한 바 있다.

국내외 선행연구 검토를 통해 도출한 특징은 다음과 같다. 첫째, 항만물동량 관련 연구는 주로 컨테이너를 대상으로 이뤄졌으며, 일부 품목 대상 연구는 해당 품목의 연관 산업 수요 및 공급 전망이 주로 수행되었다. 둘째, 대부분의 선행연구가 빈도수가 높은 월별 또는 분기별 시계열 자료를 활용하고 있는 반면, 항만 건설계획 수립의 기초가 되는 중장기 예측보다는 단기 예측에 목적을 두고 있다. 셋째, 방법론적 측면에서 시계열 모형, 회귀 모형, 인공신경망 모형 등 단일 모형을 주로 적용하여 항만물동량 예측에 있어서 개별 모형의 활용 가능성을 제시하고 있으나, 모형별 비교·검증을 통한 예측시스템 개선방안 제시는 미흡하다.

〈표 1-1〉 본 연구 관련 주요 선행연구 현황

NO	연구 목적	연구 방법	주요 연구 내용
항만물동량 예측 관련 연구			
1	• 과제명: 데이터 마이닝 기법을 이용한 항만물동량 예측 활용 방안 연구 • 연구자(년도): 전찬영 · 송주미 (2006) • 연구목적: 인공신경망모형을 활용한 단기 품목별 항만물동량 예측	• 국내외 문헌자료 분석 • 인공신경망(Artificial Neural Network) 모형 • 몬테카를로 시뮬레이션	• 데이터마이닝 기법 고찰 • 주요 물동량 추세 변화 분석 및 설명변수 설정 • 품목별 예측의 적합도 비교분석 • 몬테카를로 시뮬레이션 분석
2	• 과제명: SARIMA 모형을 이용한 우리나라 컨테이너물동량 예측 • 연구자(년도): 민경창 · 하현구 (2014) • 연구목적: 분기별 물동량 자료를 활용하여 단기 수요를 예측	• SARIMA(Seasonal Autoregressive Integrated Moving Average)모형	• 분기별 항만 컨테이너 물동량 예측 • ARIMA모형과 SARIMA 모형간 예측결과 비교
3	• 과제명: Throughput forecasting of different types of cargo in the adriatic seaport Koper • 연구자(년도): Dragan et al. (2021) • 연구목적: 품목별 항만물동량 예측모형 개발	• 선행연구 분석 • Dynamic Factor Analysis-ARIMAX • Principal component analysis • 다중회귀모형 • Monte Carlo simulation	• 54개의 거시경제변수를 DFA를 활용해 동적 요인을 추출하고, ARIMAX 모형을 통해 품목별 물동량 예측 • 2001~2017년 분기 데이터를 이용해 각 모형을 통계적 비교 검증
자동차 수요예측 관련 연구			
4	• 과제명: An analysis of factors that affecting the number of car sales in Malaysia • 연구자(년도): Islam et al. (2016) • 연구목적: 거시경제 변수 활용 회귀모형으로 자동차 수요추정	• 선행연구 분석 • 거시경제 변수 선정 • 회귀분석 • ANOVA를 활용한 가설 검정	• 말레이시아 자동차 판매량에 영향을 주는 거시경제 변수를 활용하여 회귀분석을 통해 모형 설계 • 자동차 판매량과 개별 거시변수간 상관관계를 ANOVA로 검정

NO	연구 목적	연구 방법	주요 연구 내용
5	• 과제명: A sales forecasting model in automotive industry using adaptive neuro-fuzzy inference system and genetic algorithm • 연구자(년도): Vahabi et al. (2016) • 연구목적: ANFIS 모형을 활용한 연간 자동차 판매량 예측	• 선행연구 분석 • ANFIS 모형 분석 • 모형 적합도 검정	• 인구, 실업률, 환율, 생산 지수, 제조업 판매지수, 전체 소비, 소비자 물가, 가스 가격, GDP, 물가상승률, 대출 기준금리, 주가 지수, 1인당 자본, 자동차 수출 세금 등 방대한 거시경제 데이터를 활용하여 ANFIS 모형으로 판매량 예측

석탄 수요예측 관련 연구

6	• 과제명: 광물자원이 해상물동량 전망에 관한 연구: 철광석 및 석탄을 중심으로 • 연구자(년도): 장원익(2010) • 연구목적: 조선경기 예측을 위해 건화물(철광석, 석탄)의 물동량을 전망	• 국내외 선행연구 검토 • 선형 회귀분석	• 철광석은 조강생산과 세계 GDP를 독립변수로 설정 • 석탄은 조강생산과 세계 발전량을 독립변수로 설정 • 선형 회귀분석을 통해 석탄과 철광석의 단기/중기 해상물동량 전망
7	• 과제명: Forecasting global coal consumption: An artificial neural network approach • 연구자(년도): Pablo Benalcazar et al.(2017) • 연구목적: 인공신경망 분석방법론을 이용 글로벌 석탄 소비 예측	• 선행연구 분석 • 인공신경망 분석 • 다층 퍼셉트론 신경망 (MLP) 모형	• 1970~ 2016년까지 IMF 세계 경제전망 데이터를 이용, 신경망 모델을 적용하여 세계 석탄 소비량을 예측 • 예측의 성과 개선을 위해 하이브리드 ANN 기반 기법 및 사회학적 변수 추가 적용을 제안

고철 수요예측 관련 연구

8	• 과제명: Long-term global scrap availability of steel scrap • 연구자(년도): Oda, Akimoto & Tomoda(2013) • 연구목적: 2050년 고철 가용성 파악	• 선행연구 분석 • Material flow analysis (MFA) • Life-time model	• 고철 발생원 중 폐기 철에 따른 재활용률을 분석하고, 여러 시나리오에 따른 장래 철 생산량을 가정하여 장래 고철 가용성을 파악
---	--	--	--

NO	연구 목적	연구 방법	주요 연구 내용
9	- 과제명: Global Scrap Trading Outlook Analysis for Steel Sustainability - 연구자(년도): Lee & Sohn (2015) - 연구목적: 국가 간 연도별 고철 무역구조 파악 및 고철 수요 예측	- 선행연구 분석 - 네트워크 분석 - 시계열 분석(ARIMA)	- 국가 간 고철 수출입 물동량 네트워크 구조 분석 - 시계열 기법을 통해 철 생산량 및 고철 수요량 예측
시멘트 수요예측 관련 연구			
10	- 과제명: 동해항 시멘트 물동량의 추정에 관한 연구 - 연구자: 조찬혁(2008) - 연구목적: 동해항의 시멘트 반출량 예측을 위해 Holt-Winter 지수평활방식 및 ARIMA 예측력 비교	- Holt-Winter 지수평활 모형 적용 - ARIMA(Box-Jenkins 방식)모형 적용	- 동해항의 12년간 월별 물동량 데이터를 IMF 이전, IMF 기간, IMF 이후로 구분하여 물동량 예측 - Holt-Winter 지수평활모형과 ARIMA 방식의 예측능력 비교
11	- 과제명: Forecasting Domestic Shipping Demand of Cement: Comparison of SARIMAX, ANN and Hybrid SARIMAX-ANN - 연구자: Fiskin & Cerit (2019) - 연구목적: 시멘트 항만수요 예측을 위한 SARIMA, ANN 및 하이브리드 SARIMA&ANN의 예측력 비교	- Shapiro-Wilk 및 Jarque-Bera 데이터 정규성 검정 - Augmented Dickey-Fuller 및 Philips Perron 일관성 분석 - SARIMAX 예측모형 적용 - ANN 예측모형 적용 - Hybrid SARIMAX & ANN 예측모형 적용	- 터키의 시멘트 연안 물동량 예측을 위한 SARIMAX, ANN 및 하이브리드 SARIMAX&ANN 비교 13년간 월별 물동량을 활용하여 단기(19개월) 물동량 예측 - 하이브리드 SARIMA&ANN 모형이 가장 좋은 예측력을 보임

자료: 저자 작성

2. 선행연구와 본 연구의 차별성

본 연구는 우리나라 30개 무역항을 대상으로 품목별 항만물동량을 수·출입, 환적, 연안 물동량으로 구분하여 중장기 예측모형을 개발하고 실증분석을 실시하였다. 본 연구가 항만물동량을 예측한다는 측면에서 일부 선행

연구와 유사성이 있으나, 기본적으로 연구 목적과 범위, 방법 등에서 기존의 연구와 차별성을 가진다.

연구의 목적 측면에서 기존에 활용 중인 항만물동량 예측모형에 대해 정밀한 진단을 실시하고, 주요 비컨테이너 품목을 대상으로 예측모형의 고도화 방안을 제시한다. 이는 선행연구들이 주로 항만물동량 예측에 있어서 개별 예측모형의 활용 가능성을 제시하는 데 목적을 두는 것과 명확한 차별성이 존재한다. 특히 다양한 유형의 방법론을 종합적으로 고려하여 최적의 예측모형을 제시한다는 측면에서 연구의 방향성이 차별화된다.

연구의 범위에서도 본 연구와 유사한 선행연구를 찾을 수 없다. 본 연구의 시간적 범위는 우리나라 항만물동량이 현재의 체계로 구축되어 공표되기 시작한 1994년부터 2021년까지의 연도별 자료를 기반으로 2040년까지의 중장기 항만물동량을 예측하고 있으며, 공간적 범위는 우리나라 30개 무역항을 대상으로 한다. 또한 내용적 범위는 비컨테이너 품목 중 주요 모니터링 대상인 자동차, 석탄, 고철, 시멘트를 대상으로 물동량의 유형을 수입, 수출, 환적, 연안 물동량으로 구분하여 예측모형을 개발한다.

연구의 방법에서는 본 연구가 다양한 예측모형 간 적합한 비교·검토를 통해 주어진 여건하에서 최적의 모형을 제시한다는 측면에서 선행연구와 차별성이 있다. 본 연구는 수많은 예비모형 중에서 최적의 모형을 선별하는 기준으로 기존의 연구에서 많이 활용된 검정통계량과 비교·검정 방법론을 차용하는 한편 최종적으로 모형의 활용도와 효율성 등을 다각적으로 활용한다. 이는 기존 선행연구들이 예측력을 기반으로 신규 모형의 제안에 그친 반면 본 연구는 제안된 모형을 정책 개발에 직접적으로 활용하고자 함에 따른 특수성에 기반하고 있다.

〈표 1-2〉 연구 차별성

구분	선행연구와의 차별성		본 연구
	요약	참고문헌	
연구 목표	항만물동량 예측 시 개별 모형의 활용 가능성 제시에 그침 상시적 예측시스템 구축을 위한 방안 제시는 미흡함	전찬영 · 송주미(2007), 이민규 외(2014), 민경창 · 하현구(2014), 신창훈(2011) 등	기존 항만물동량 예측 모형에 대해 정밀한 진단을 실시함 비컨테이너 항만물동량 예측 모형 고도화 방안을 제시함
연구 대상	일부 품목별 물동량 또는 제한된 항만을 대상으로 실증분석을 실시함	하준수(2021), 박상수 외(2017), Fiskin & Cerit(2019) 등	우리나라 전체 항만의 주요 품목을 연구대상으로 설정함
연구 범위	상대적으로 짧은 시계열 자료를 활용함 중장기 예측보다는 단기 예측에 집중함	하준수 외(2021), 민경창 · 하현구(2014), 김두환 · 이강배(2020) 등	모형의 예측력, 활용도, 효율성을 다각적으로 검토하고 단기/중장기 예측에 적합한 모형을 제시함
연구 방법	시계열 모형, 회귀 모형, 인공신경망 모형 등 단일 모형 적용함	KMI(2016), 하준수(2021), 장원익(2010), Lee & Sohn(2015) 등	다양한 모형을 적용하고 비교검정을 통해 최적의 모형을 제시함

자료: 저자 작성

02

항만수요예측 방법론 고찰

제2장에서는 다양한 분야의 수요예측에 활용되는 방법론을 모형의 특성별로 구분하여 연구방법론을 검토한다. 주요 수요예측 방법론으로 시계열 모형, 회귀 모형, 머신러닝 기반 모형을 살펴보았으며, 예측모형 간 비교·검정을 위해 일반적으로 활용되는 검정통계량과 비교·검정 방법론의 특징과 활용방안을 기술한다.

제1절 주요 수요예측 방법론

1. 시계열 모형

시계열 모형은 시간 순서에 따라 기록된 시계열 자료를 분석하여 분석 대상의 추세를 파악하고 예측하기 위한 모형이다. 시계열 자료는 관측치 간 확률적으로 독립이 아니기에, 관측치 간 종속 관계를 잘 활용하면 과거 자료를 통해 현재와 미래 자료를 어느 정도 예측할 수 있다. 시계열 자료를 분석하는 접근법은 크게 탐색적 분석기법과 확률적 분석기법으로 분류할 수 있다.²⁾ 탐색적 분석기법의 경우 대표적으로 이동평균 평활법, 지수 평활법, 회귀모형

평활법과 같은 평활법(Smoothing)과 요소분해법(Time series decomposition) 등이 자주 활용된다. 시계열 자료에 내포되어 있는 패턴을 단순하지만 효율적인 방법으로 모형화하는 탐색적 분석기법과 달리, 확률적 분석기법은 추정값과 관측값과의 차이인 잔차를 확률변수로 가정하여 분석한다. 이러한 확률적 분석기법에는 ARMA, ARIMA, SARIMA, VAR, VECM, GARCH 등이 존재한다. 본 장에서는 다양한 모형 중 항만물동량 예측모형 개발과 관련하여 기존 연구에서 사용한 모형에 대해 소개한다. 구체적으로 탐색적 분석기법 중에서는 단순이동평균 평활법과 가중이동평균 평활법에 대해 설명하고, 확률적 분석기법 중에서는 ARIMA, SARIMA, VAR 등에 대해 기술한다.

1) 탐색적 분석기법

이동평균 평활법을 비롯하여 탐색적 분석기법의 장점 중 하나는 상대적으로 시계열 자료의 양이 적을 때에도 분석할 수 있다는 점이다. 특히 변동폭이 작을 것으로 기대되는 분석 대상의 경우, 간단한 탐색적 분석기법의 활용으로도 충분히 유의미한 예측값을 얻을 수 있다.

(1) 단순 이동평균 평활법(Simple Moving Average Smoothing)

단순 이동평균 평활법은 시계열 분석 대상의 최근 N 시점 자료를 산술평균하여 예측치를 산정하는 기법이다. 예를 들어, 분석하고자 하는 t 시점의 시계열 자료를 y_t 라고 정의했을 때, $t+1$ 시점의 예측치를 산정하는 수식은 다음과 같이 표현된다.

$$y_{t+1} = (y_t + y_{t-1} + \cdots + y_{t-N+1})/N \quad (\text{II-1})$$

2) 이재길(2017), pp. 89-95.

일반적으로, N 의 값이 클수록 시계열 자료의 예측치들은 더 고르게 분포하며, N 이 작을수록 예측치의 최근 데이터에 의해 변동되는 폭이 늘어나게 된다. 이때 데이터의 특성에 따라 가장 예측력이 높은 N 의 값을 선택해야 하고, 이를 위해 MSE, MAPE 등의 예측 정확도 지표를 활용할 수 있다. 예측 정확도 지표에 대한 설명은 제2절에서 후술한다.

(2) 가중 이동평균 평활법(Weighted Moving Average Smoothing)

단순 이동평균 평활법은 t 시점의 예측값에 대해 과거 N 시점의 관측치가 모두 동일한 수준의 영향력을 가진다고 가정한다. 그러나 일반적으로 시계열 예측에 있어서 최근의 관측값이 먼 과거의 관측값보다 더 큰 영향력을 가질 것으로 기대된다. 이에 가중 이동평균 평활법은 이러한 부분을 과거 N 시점의 관측치에 각각의 가중치를 부여하는 방법으로 해소한다. 가중 이동평균 평활법을 통해 $t+1$ 시점의 예측치를 산정하는 수식은 다음과 같이 표현할 수 있다.

$$y_{t+1} = (w_t y_t + w_{t-1} y_{t-1} + \dots + w_{t-N+1} y_{t-N+1}) / N \quad (\text{II-2})$$

여기서 가중치 w_t 는 $w_t + w_{t-1} + \dots + w_{t-N+1} = 1$ 을 만족해야 한다.

2) 확률적 분석기법

시계열 자료의 관측치가 생성될 때 그 관측치가 특정한 확률 분포에서 추출된 표본이라고 가정하면, 이러한 관측치들의 집합을 확률과정(stochastic process)이라고 표현한다. 그리고 t 시점의 시계열 자료 y_t 와 확률과정 $\{y_t\}$ 에 대해서 임의의 자연수 s 에 대해 $\{y_t\}$ 와 $\{y_{t+s}\}$ 가 같은 결합확률분포를 가질 경우 이 확률과정은 정상성(stationarity)을 가진다고 한다. 확률과정 $\{y_t\}$ 가 정상성을 가지면 다음과 같은 조건을 만족한다.

$$E(y_t) = E(y_{t+s}) \quad (\text{II-3})$$

$$\text{Var}(y_t) = \text{Var}(y_{t+s})$$

$$\text{Cov}(y_t, y_{t+s}) = \gamma(s)$$

즉, 정상 시계열 자료는 기댓값과 분산이 일정하며, 두 개의 시계열 자료 간 공분산이 t 와 무관하게 두 시점 간의 차이인 s 에만 의존한다. 또한 이 세 조건을 만족하는 시계열 자료를 약한 정상성(weakly stationary) 시계열이라고 한다. 그리고 정상성이 없는 시계열 자료를 비정상 시계열 자료라고 한다. 일반적으로 시계열 자료 간 결합확률분포를 파악하는 것은 매우 어렵기에 시계열 분석에서의 정상성은 대부분 약한 정상성을 의미한다.

(1) ARMA 모형

ARMA 모형(Autoregressive Moving-Average model, 자기회귀이동평균모형)은 정상성을 갖는 단일 변수 시계열 자료를 분석하는 보편적인 방법으로 AR 모형(Autoregressive model, 자기회귀모형)과 MA 모형(Moving Average model; 이동평균모형)이 합쳐진 형태이다. 먼저 AR 모형은 특정 시점의 시계열 자료를 과거 시점 관측치들의 선형결합으로 설명할 수 있다고 가정한다. 이에 따라 AR 모형은 종속변수 y_t 를 설명변수 $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ 로 설명하는 회귀식의 모형을 가진다. 이때, AR(1) 모형은 종속변수 y_t 를 한 시점 전 과거 자료인 y_{t-1} 와 잔차항 e_t 만을 사용하여 설명하는 모형으로 다음과 같이 표현할 수 있다.

$$y_t = \phi y_{t-1} + e_t, \quad e_t \sim i.i.d.(0, \sigma_e^2) \quad (\text{II-4})$$

이를 일반화한 AR(p) 모형은 종속변수 y_t 를 p 개의 과거 데이터를 통해 설명하는 모형이며, 다음 식과 같이 표현할 수 있다.

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + e_t, \quad e_t \sim i.i.d.(0, \sigma_e^2) \quad (\text{II-5})$$

한편, MA 모형은 특정 시점의 시계열 자료를 과거 시점 잔차항들의 선형결합으로 설명한다. 즉, MA 모형은 종속변수 e_t 를 $e_{t-1}, e_{t-2}, \dots, e_{t-q}$ 로 설명하는 회귀식의 모형을 가진다. 한 시점 전의 잔차항을 설명변수로 사용하는 MA(1) 모형은 다음과 같이 표현할 수 있다.

$$y_t = e_t + \theta e_{t-1}, \quad e_t \sim i.i.d.(0, \sigma_e^2) \quad (\text{II-6})$$

이를 일반화한 MA(q) 모형은 종속변수 y_t 를 q 개의 과거 잔차항으로 설명하는 모형이며, 다음 식과 같이 표현할 수 있다.

$$y_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q}, \quad e_t \sim i.i.d.(0, \sigma_e^2) \quad (\text{II-7})$$

ARMA 모형은 이 두 모형을 결합한 모형으로 특정 시점의 시계열 자료를 과거 시점 관측치 및 잔차항을 모두 활용하여 설명한다. 종속변수 y_t 를 한 시점 전 과거 자료인 y_{t-1} 와 한 시점 전 잔차항 e_{t-1} 을 통해 설명하는 모형인 ARMA(1,1) 모형의 식은 다음과 같다.

$$y_t = \phi y_{t-1} + e_t + \theta e_{t-1}, \quad e_t \sim i.i.d.(0, \sigma_e^2) \quad (\text{II-8})$$

이를 일반화한 ARMA(p, q) 모형은 p 개의 과거 관측치와 q 개의 과거 잔차항을 사용하는 모형으로 다음 식과 같이 표현된다.

$$\begin{aligned} y_t &= \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} \\ &= \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j e_{t-j} + e_t, \quad e_t \sim i.i.d.(0, \sigma_e^2) \end{aligned} \quad (\text{II-9})$$

(2) ARIMA 모형

ARMA 모형의 경우 정상 시계열 자료를 가정하고 있으나 현실적으로 마주하는 시계열 자료는 시간 변화에 따라 평균, 분산, 자기공분산 등이 변화하는 비정상 시계열 자료가 대부분이다. 이러한 비정상 시계열 자료에도

ARMA 모형을 적용하기 위해서는 정상 시계열로의 변환 방법이 필요하다. 이 중 대표적인 기법인 Box-Jenkins 방법은 비정상 시계열 자료에 대해 차분(differencing)을 적용하여 ARMA 모형을 사용할 수 있는 정상 시계열 자료로 변환한다. 이때 1차 차분을 적용하였음에도 비정상성이 존재할 경우 이보다 높은 d 차 차분을 적용하며, 이를 수식으로 표현하면 다음과 같다.

$$(1차 차분) \Delta y_t = y_t - y_{t-1} \quad (\text{II-10})$$

$$(d차 차분) \Delta^d y_t = \Delta y_t - \Delta y_{t-d} \quad (\text{II-11})$$

비정상 시계열 자료에 d 차 차분을 적용한 자료가 정상성을 가지게 된 경우 ARMA 모형을 적용할 수 있으며, 이를 ARIMA 모형(Autoregressive Integrated Moving Average model)이라고 한다. d 차 차분을 적용한 정상 시계열 자료 $TRIANGLE^d y_t$ 에 $ARMA(p, q)$ 모형을 적용한 경우 $ARIMA(p, d, q)$ 라고 표현하며, 일반적으로 실제 시계열 자료를 다룰 때 차분의 차수는 3을 초과하지 않는다.

ARIMA 모형을 식으로 간단하게 표현하기 위해 먼저 차수가 m 인 후향 연산자(backward shift operator) B^m 을 정의한다. 또한 종속변수 y_t 의 m 시점 전 과거 자료인 y_{t-m} 은 $(B^m)y_t$ 로 표현된다. 이를 이용하여 $(B^3)y_t = y_{t-3}$, $(1-B)y_t = y_t - y_{t-1}$ 과 같은 관계가 성립된다고 가정하면, $ARIMA(p, d, q)$ 모형은 다음 식과 같이 표현할 수 있다.

$$\begin{aligned} & (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d y_t \\ & = (1 - \theta_1 B - \theta_1 B^2 - \dots - \theta_q B^q) e_t, \quad e_t \sim i.i.d.(0, \sigma_e^2) \end{aligned} \quad (\text{II-12})$$

위 식에서 $(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)y_t$ 는 모형 내 AR 부분을 나타내며, $(1 - \theta_1 B - \theta_1 B^2 - \dots - \theta_q B^q)e_t$ 는 MA 부분을 나타낸다. 그리고 $(1 - B)^d y_t$ 는 y_t 의 d 차 차분을 의미한다. 이때 AR과 MA의 계수 부분은 각각 다음과

같은 연산자로 표현할 수 있다.

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (\text{II-13})$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (\text{II-14})$$

위의 연산자를 앞선 식 (II-12)에 대입하면 ARIMA(p, d, q)는 다음과 같이 간단하게 나타낼 수 있다.

$$\phi_p(B)(1-B)^d y_t = \theta_q(B) e_t, \quad e_t \sim i.i.d.(0, \sigma_e^2) \quad (\text{II-15})$$

하준수 외(2021)는 부산항 9개 부두별 일간 물동량 데이터를 통해 ARIMA 모형을 적용하여 42일 치 단기 수요예측을 시행했고, 해당 모형이 비정상적인 물동량 흐름의 감지 및 대응에 활용될 수 있음을 보인 바 있다.

(3) SARIMA 모형

시계열 자료에 계절성 등의 주기가 관찰되는 경우에는 계절차분을 적용하여 계절적인 변동성을 제거한 시계열 자료로 변환하여 분석한다. 기존 ARMA 모형에 계절차분까지 고려된 모형을 SARMA 모형(Seasonal Autoregressive Moving Average model)이라고 한다. 예를 들어 분석 대상 시계열 자료가 월별로 주기가 나타난다면 계절주기를 12로 설정하고, SARMA 모형은 월별 ARMA 모형을 각각 고려하여 미래 예측치를 산정한다. 이때 각 계절 ARMA 모형의 AR과 MA 차수를 각각 P 와 Q 라고 한다면, 시계열 자료의 주기적인 특성은 다음과 같이 나타낼 수 있다.

$$\Phi_P(B^{12})y_t = \Theta_Q(B^{12})e_t \quad (\text{II-16})$$

여기서 $\Phi_P(B^{12})$ 와 $\Theta_Q(B^{12})$ 는 각각 다음과 같이 정의된다.

$$\Phi_P(B^{12}) = 1 - \phi_1 B^{12} - \phi_2 B^{24} - \dots - \phi_P B^{12P} \quad (\text{II-17})$$

$$\Theta_Q(B^{12}) = 1 - \Theta_1 B^{12} - \Theta_2 B^{24} - \dots - \Theta_Q B^{12Q} \quad (\text{II-18})$$

기존 ARMA(p, q) 모형과 위 주기적 특성을 곱해 주기가 s 인 시계열 자료의 계절성 및 비계절성을 함께 고려하여 일반화한 SARMA(p, q)(P, Q) s 모형은 다음과 같이 표현된다.

$$\phi_p(B)\Phi_P(B^s)y_t = \theta_q(B)\Theta_Q(B^s)e_t \quad (\text{II-19})$$

또한 비정상 계절 시계열 자료를 분석하기 위해서는 추세차분과 계절차분을 함께 고려한 SARIMA 모형(Seasonal Autoregressive Integrated Moving Average model)을 활용한다. 시계열 자료 $\{y_t\}$ 의 비계절차수가 p, d, q 이고 계절차수가 P, D, Q 이며 계절주기가 s 라고 할 때, SARIMA(p, d, q)(P, D, Q) s 모형은 다음과 같이 나타낼 수 있다.

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)e_t \quad (\text{II-20})$$

SARIMA 모형을 활용하여 항만물동량을 예측한 선행연구에는 민경창·하현구(2014), Farhan & Ong(2018) 등이 있다. 민경창·하현구(2014)는 분기별 국내 전체 항만 컨테이너 물동량을 SARIMA 모형을 통해 예측했고, 항만 물동량의 단기 수요 예측에 있어서 기존 ARIMA 모형보다 적합하다고 주장했다. Farhan & Ong(2018)은 세계 20위권 컨테이너 항만의 월간 물동량데이터를 대상으로 SARIMA 모형을 적용하였고, 대부분의 항만에서 높은 예측 정확도가 나타났음을 보였다.

(4) VAR 모형

앞서 설명한 시계열 모형들은 모두 단변량 시계열 모형으로써 분석 대상인 시계열 자료를 설명하기 위한 변수로 과거의 해당 시계열 자료만을 사용한다. 그러나 시계열 자료를 분석하기 위해 추가적인 설명변수를 사용하여 할 경우 다변량 시계열 모형을 활용할 수 있다. 그중 VAR 모형(Vector Autoregressive model)은 두 개 이상의 시계열 자료가 서로의 값을 설명

하는 동시성(simultaneity)을 가질 때 활용할 수 있다. 예를 들어 GDP와 실업률, 온도와 습도 등의 시계열 자료는 두 변수를 같이 분석할 때 더 많은 시사점을 제공할 수 있다. 이처럼 VAR 모형은 각각의 변수에 대해서 개별 모형을 구성하고 추정하는 것보다 변수 간의 관계를 고려할 수 있다는 점에서 논리적인 강점을 가지며, 이를 통해 더 다양한 정보를 도출할 수 있다는 장점이 있다.

VAR 모형은 서로 연관이 있는 k 개의 시계열 자료 $y_{i,t}$ ($i = 1, \dots, k$)에 대해 각 자료가 자기 자신과 상대방의 과거 자료에 의해 설명될 수 있다고 가정한다. $y_{i,t}$ 가 서로의 과거 데이터 p 개를 통해 설명되는 VAR(p) 모형을 수식으로 표현하면 다음과 같다.

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_{t-p} + e_t \quad (\text{II-21})$$

여기서 y_t 는 $y_{1,t}, \dots, y_{k,t}$ 로 이루어진 k -벡터이며, c 는 모형의 절편인 k -벡터, A_p 는 시불변(time-invariant) $k \times k$ -행렬, e_t 는 k -벡터 오차항이다.

VAR 모형을 활용한 수요예측 관련 선행연구에는 Gao, Luo, & Zou(2016), Zhang et al.(2017) 등이 있다. Gao, Luo, & Zou(2016)은 경쟁 관계에 있는 선전항과 홍콩항의 월별 컨테이너 물동량을 예측하기 위해 VAR 모형을 활용하였다. 이를 통해, 선전항 물동량 증가율이 홍콩항에 미치는 영향보다 홍콩항 물동량 증가율이 선전항에 미치는 영향이 더 크다는 것을 확인하였다. Zhang et al.(2017)은 중국 전기 자동차 판매량을 예측하기 위해 VAR을 사용하였다. Battery 방식과 Plug-in 방식의 두 유형으로 구분한 전기 자동차의 판매량뿐 아니라, 소비자 물가지수, 소비자 신뢰지수, 공산품 가격 지수, 소매 연료 가격, 자동차 가격, Baidu 검색 지수 등 6개의 외생변수를 VAR 모형에 함께 적용하여 단기(1년, 월간) 및 중장기(5년, 연간) 예측에 활용하였다.

2. 회귀 모형

회귀 모형은 다양한 통계분석 모형 중에서 추정 방식의 편의성과 합리성, 활용 방식의 범용성 등으로 가장 널리 활용되는 모형이다. 기본적으로 종속변수와 설명변수 간의 관계(함수)를 가정하고 설명변수가 종속변수에 미치는 영향력을 함수의 계수로 추정한다. 추정된 계수를 이용하여 특정 설명변수 값에 대응되는 종속변수의 값을 예측할 수 있으며, 또한 두 변수의 관계에 대한 가설을 설정하고 실측 데이터를 통하여 이를 검정할 수 있다. 일반적으로 종속변수와 독립변수의 관계를 선형(linear)으로 가정하는 선형 회귀분석과 비선형(non-linear)으로 가정하는 비선형 회귀분석으로 나뉘며, 모형에 포함된 독립변수의 수에 따라 단순회귀분석(Simple Regression)과 다중회귀분석(Multiple Regression)으로 구분된다.

수요예측 관련 선행연구에서 가장 보편적으로 활용된 다중선형회귀(Multiple Linear Regression, MLR) 모형의 일반식은 다음과 같다.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \epsilon_i, \epsilon_i \sim i.i.d.N(0, \sigma^2) \quad (\text{II-22})$$

위 식은 n 의 관측치와 한 개의 종속변수, p 개의 독립변수를 가정하고 있다. 또한 y_i 는 종속변수의 i 번째 관측치, x_{ij} 는 j 번째 독립변수의 i 번째 관측치이며, β_j 는 추정해야 할 계수, ϵ_i 는 i 번째 오차이다.

회귀 모형의 계수 추정에는 일반적으로 최소자승법(Least Squares method, LS)이 사용되며, 그 외에도 최우도추정법(Maximum Likelihood Estimation, MLE), 최소절대편차(Least Absolute Deviation, LAD) 등 다양한 추정법이 사용된다. 회귀 모형이 다양한 분야에서 예측모형으로 널리 활용되는 것에는 추정법이 상대적으로 간편하며, 추정량의 분포를 가정한 통계적 추론이 정형화되어 있어 모형 선택 및 가설검정이 용이하다는 점이 크게 작용한다. 또한 종속변수에 영향을 미칠 수 있는 다양한 독립변수를 활용하여

예측력을 향상시킬 수 있으며, 예측 결과에 대한 해석이 상대적으로 용이하여 활용도가 높다는 장점이 있다.

다만 항만물동량 예측 관련 학술 연구에서 회귀 모형을 단독으로 활용한 연구는 많지 않다. 이는 회귀 모형이 타 모형들과 비교하여 상대적으로 오랜 기간 연구되어 방법론적으로 연구의 차별성 및 참신성을 확보하기 어렵기 때문으로 풀이된다. 그럼에도 여전히 일부 연구에서는 회귀 모형을 부분적으로 활용하고 있다. 특히 실증분석에 있어서 모형의 특수성보다는 연구 결과의 해석이 중요하거나, 또는 신규 모형을 제시하는 데 있어서 비교군의 역할로 회귀 모형을 활용하는 경우이다.

장원익(2010)은 조선업 경기 예측을 위해 선형적으로 해상물동량에서 큰 비중을 차지하는 철광석 및 석탄의 물동량을 전망하였으며, 이를 위해 조강생산능력, 세계 GDP, 세계 발전량 등을 독립변수로 하는 선형회귀분석을 실시하였다. 한편 Dragan et al.(2021)은 아드리아해 코퍼(Koper)항의 품목별 항만물동량을 ARIMAX 모형으로 예측하였으며, 모형의 예측력을 선형회귀분석을 통한 결과와 비교 분석하였다. 또한 Lam et al.(2004)은 홍콩항의 품목별 항만물동량 예측을 위해 신경망 모형을 활용하였으며, 회귀 모형을 통한 결과와 비교·검증한 바 있다.

3. 머신러닝 기반 모형

Mitchell & Mitchell(1997)은 머신러닝을 ‘컴퓨터 프로그램이 작업에서 경험을 얻어 성능을 향상하는 것’으로 정의했다. 항만 수요예측 문제에 머신러닝을 적용하는 목적은 과거 수요 데이터에서 특성(feature)을 추출하여 미래 수요에 대한 예측력을 향상하는 것이다. 머신러닝은 지도학습(Supervised learning), 비지도학습(Unsupervised learning) 및 강화학습

(Reinforcement learning)으로 구분할 수 있다. 지도학습은 모형이 예측하고자 하는 결과값(label)이 알려져 있으며, 비지도학습은 결과값이 알려져 있지 않은 것이 특징이다. 강화학습은 상태(state)에 따른 행동(action)에서 얻을 수 있는 보상(reward)을 통해 최적의 행동을 선택하는 연구 분야이다.

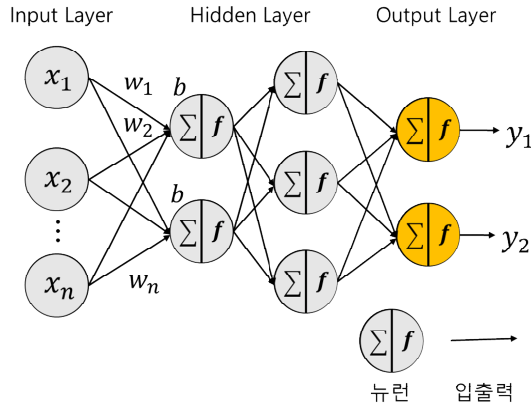
최근 머신러닝 기반 항만 수요예측에 대한 연구가 활발하게 진행되고 있다. 항만 수요예측은 사회경제적 변수와 같은 이산형 데이터와 실제 항만물동량과 같은 연속형(시계열) 데이터를 사용한다. 문제 해결 방식 및 모형 설계에 따라 지도학습과 비지도학습 모두 적용이 가능하나 일반적으로 비지도학습 문제로 접근한다. 본 절에서는 항만 수요예측 문제를 해결하기 위해 적용된 머신러닝 방법론의 개념을 설명하고 주요 사례연구를 통해 머신러닝 방법론의 활용방안에 대해 기술한다.

1) 인공신경망(Artificial Neural Network, ANN)

ANN의 개념은 헵의 학습(Hebb, 1949)과 임계값 논리(McCulloch & Pitts, 1943)에 기반하고 있으며, 이후 퍼셉트론(Rosenblatt, 1958)과 역전파 알고리즘(Rumelhart, Hinton, & Williams, 1986) 개발로 인해 현재의 ANN 모형으로 진화되었으며 다양한 분야에서 널리 사용되고 있는 머신러닝 방법론이다.

〈그림 2-1〉은 인공신경망의 기본 구조이다. ANN은 입력층(input layer), 은닉층(hidden layer)과 출력층(output layer)로 구성되어 있다. 노드는 뉴런, 아크는 입출력을 의미하며 완전연결층(fully connected layer)으로 구성되어 있다. 뉴런은 입력값 별 가중치(w_n)와 편향(b)의 합을 구하고 활성화함수(f)를 통해 최종 출력값을 도출한다. 모형을 구성하는 뉴런의 개수, 은닉층의 개수, 활성화함수 등은 하이퍼 파라미터(hyper parameter)라고 하며 문제의 특성과 계산 효율성에 맞게 모형 설계자가 직접 설정한다.

〈그림 2-1〉 ANN의 구조



자료: Rumelhart, Hinton & Williams(1986), p. 535. 저자 재작성

Lapedes & Farber(1987)는 ANN이 비선형 시계열 분석에도 적용될 수 있음을 보여준다. 전찬영·송주미(2007)는 10개 품목의 물동량 예측을 위해 회귀분석과 ANN을 적용하고 두 모형을 적합성, 신뢰성, 강건성 측면에서 비교하였다. 본 연구는 항만물동량 예측 문제에 ANN의 활용 가능성을 제시했다. 신창훈 외(2008)는 ARIMA와 ANN을 결합한 하이브리드 모형을 제시하였으며 부산, 인천, 광양, 울산 및 마산항의 컨테이너 수출입 물동량을 ARIMA, ANN 및 하이브리드 모형으로 예측하고 비교·검증하였다. 조상호 외(2020)는 광양항의 석탄 및 철광석 물동량 예측을 위해 단계적 회귀분석과 ANN 모형을 적용하였으며 그 결과를 비교 분석했다. 김동윤 외(2021)는 다층신경망(Multi Layer Perceptron, MLP)을 활용한 부산항의 공컨테이너 수요예측 모형을 제안하였으며 다중선형회귀분석과의 단기 수요 예측력을 비교했다.

Gosassng et al.(2011, 2018)은 거시경제 지표와 항만 수요의 관계와 같은 복잡한 비정상성 데이터(nonstationary data)를 처리하는 것은 전통적 통계 방법론보다 ANN이 더 적합하다고 지적했다. Fradinate et al.(2014)는

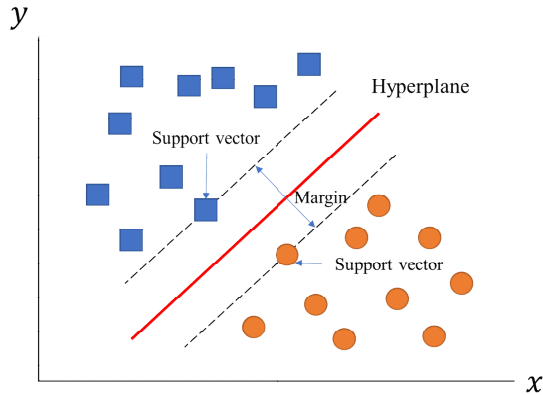
인도네시아 Aceh 지방의 시멘트 판매량과 판매 수요를 예측하기 위해 ANN과 ARIMA를 통한 예측모형을 제안했다. 22개월의 데이터로 단기수요(2개월)의 시멘트 수요를 예측하고 ANN, ARIMA와 MA의 예측력을 비교했다. Fradinate, Suthummanon, & Suntiamorntut(2015)는 추가 연구를 통해 인도네시아의 GDP 및 인구 등의 외생변수를 활용한 ANN 모형을 통해 벌크 물동량의 단기수요(2개월)를 예측했다.

2) SVM(Support Vector Machine)

SVM은 Boser, Guyon, & Vapnik(1992)에 의해 제안되었다. SVM은 의사결정 범위 내에서 서포트벡터(support vector) 간의 거리인 마진(margin)을 최대로 하는 초평면(hyperplane)을 찾는 학습 알고리즘으로 이진 분류(binary classification) 문제나 예측 문제에 주로 사용된다(〈그림 2-2〉 참고). 수요예측 문제를 해결하기 위해서는 SVR(Support Vector Regression)이나 ANN 및 군집화 모형과 결합한 SVM이 적용된다.

Ding et al.(2019)는 2002년부터 2014년까지의 Ningbo항과 원저우항의 컨테이너 수출입 물동량 예측을 위해 역전파 ANN-SVM의 결합 모형을 제안하고 결합 모형의 우수성을 검증하였다. Ruiz-Aguilar et al.(2020)은 자기조직화지도(Self-Organizing Map, SOM), SARIMA 및 SVM을 통한 군집화 기반 시계열 분석모형을 제안하고 항만 컨테이너 물동량을 예측하였다. 제안된 군집화 기반 시계열 분석모형은 SOM을 통해 데이터 차원을 축소 및 군집화하여 각기 다른 군집에서 SARIMA와 SVR을 통한 예측 결과를 도출하며 1일 혹은 7일의 컨테이너 물동량 예측을 위한 초단기 예측 모형이다. Li et al.(2017)은 가우스함수를 적용한 Gauss- ν SVM을 제안하고 중국의 주요 사회경제적 외생변수를 사용해 상하이항과 텐진항의 항만 수요를 예측했다. 또한 ARIMA 및 ANN 등 기존 예측 방법론과의 비교를 통해 모형의 우수성을 검증했다.

〈그림 2-2〉 SVM의 개념도



자료: Boser, Guyon, & Vapnik.(1992), p. 146. 저자 재작성

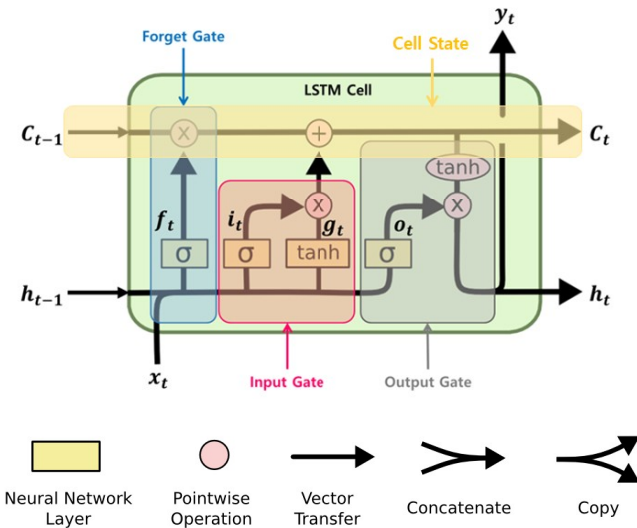
3) LSTM(Long Short-Term Memory)

LSTM은 순차적 데이터(sequential data)를 학습하고 처리할 수 있는 RNN(Recurrent Neural Network)에서 개선된 개념이기 때문에 자연어 처리(natural language processor), 시계열 분석 등에 주로 쓰인다.

LSTM의 기본 단위인 셀의 구조는 〈그림 2-3〉과 같다. LSTM 셀은 Cell state, Forget gate, Input gate와 Output gate로 구분할 수 있다. Cell state는 다음 셀이 정보를 전달하는 것으로 Forget gate, Input gate, 그리고 Output gate에서 결정된 데이터를 다음 셀로 전달한다. Cell state는 RNN에서는 셀의 개수가 늘어날수록 발생하기 쉬운 기울기 소실 및 폭발(gradient vanishing and exploding) 문제를 해결하기 위해 적용되었다. Forget gate는 과거 정보를 잊기 위한 것, Input gate는 새로운 정보를 입력하기 위한 것이며 Output gate는 어떤 정보를 출력할지 결정하는 것이다. LSTM의 구체적인 알고리즘에 대한 자세한 내용은 Hochreiter & Schmidhuber(1997)에 기술되어 있다.

김두환 · 이강배(2020)는 LSTM 모델을 활용하여 부산항의 컨테이너 물동량을 예측했다. 2004년~2016년 기간의 부산항 컨테이너 월별 물동량을 학습하여 2016년~2018년의 물동량을 예측했다. 모델을 구성하는 하이퍼파라미터(뉴런의 개수, 레이어 수, 배치 사이즈, 에포크, 드롭아웃, 활성화 함수)는 경험적 실험을 통해 결정했으며 성능 평가를 위해서 SARIMA 모형과 MAPE(Mean Absolute Percentage Error)와 RMSE(Root Mean Squared Error)를 통해 예측력을 비교했다. 본 실험에서는 LSTM의 예측 정확도가 SARIMA 모형보다 높은 것으로 나타났다.

〈그림 2-3〉 LSTM 셀의 구조



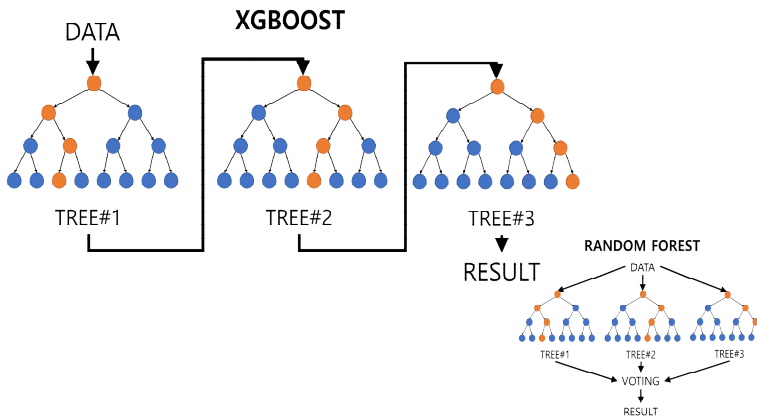
자료: 김두환 · 이광배(2022), p. 57, Olah(2015), p. 1.

4) XG부스트(Extreme Gradient Boosting)

XG부스트는 다수의 약한 의사결정나무를 결합하여 결과를 도출하는 기법이며 머신러닝 기법 중 앙상블 모형에 해당한다. XG부스트는 약한 의사결정나무가 도출한 결과를 순차적으로 취합하는 방식인 부스팅(Boosting)

기법을 사용하며 여러 개의 의사결정나무 중 일부를 선택해 결과를 취합하는 배깅(Bagging) 기법을 사용하는 랜덤포레스트와는 양상불 방식에 차이가 있다. XG부스트는 분류 및 예측 문제에 적용이 가능하며 최근 시계열 분석에도 탁월한 성능을 보여주고 있다(Chen & Guestrin, 2016). XG부스트는 주식, 연속적 센서 데이터 등 다양한 시계열 데이터 예측을 위한 연구에 활용되었으나, 아직 항만수요 및 물동량 예측을 위한 연구에 활용되지는 않았다.

〈그림 2-4〉 XG부스트와 랜덤포레스트



자료: Singh(2020), p. 1. 저자 재작성

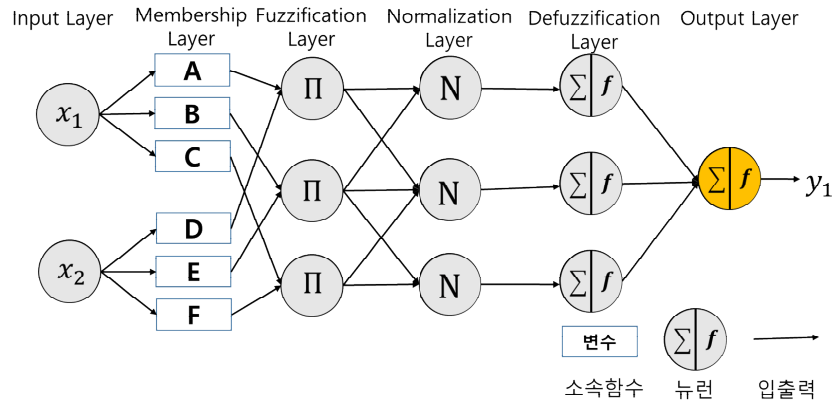
5) ANFIS(Adaptive Neuro Fuzzy Inference System)

ANFIS는 퍼지 이론과 인공신경망(ANN)이 결합한 형태의 머신러닝 모형이다(Jang et al, 1993). 인공신경망은 블랙박스과 같아서 모형이 학습하여 도출한 결과에 대한 논리적 설명이 어렵고, 퍼지 이론은 IF-THEN 규칙에 기반한 소속함수(membership function)로 결과도출에 대한 논리적 설명이 가능하지만 소속함수 최적화가 어렵다. ANFIS는 두 이론의 약점을 보완하기 위해 개발된 모형으로 인공신경망의 학습 방법을 통해 소속함수의

최적 파라미터를 구할 수 있고 별도의 수리적 모형을 정의하지 않아도 비선형 예측에 효과적인 모형이다. ANFIS는 날씨, 에너지소비 및 항공기 수요 등 다양한 시계열 예측 문제에 적용되었지만 항만수요 및 물동량 예측을 위한 연구는 아직 진행되지 않았다.

ANFIS는 ANN과 동일한 입출력층을 제외하고 네 가지 특수한 레이어로 구성되어 있다. 소속함수층(membership layer)은 소속함수와 전제 매개변수를 통해 입력데이터가 퍼지집합에 속하는 정도를 확인해 퍼지집합 내 소속등급(membership grade)을 라벨링하여 출력한다. 퍼지화층(fuzzification layer)은 소속함수층에서 출력된 모든 소속등급을 결합해 출력한다. 정규화층(normalization layer)은 퍼지화층으로부터 입력된 정보를 정규화해서 정규화된 신호를 출력한다. 마지막으로 역퍼지화층(defuzzification layer)은 정규화된 신호와 결과 매개변수를 통해 출력신호를 발생시킨다. ANFIS는 인공지능망의 학습 방식을 통해 모형을 구성하고 있는 매개변수를 최적화하여 예측력을 높인다.

〈그림 2-5〉 ANFIS 구조



자료: Pae et al. (2018), p. 6. 저자 재작성

제2절 예측모형의 적합도 비교 · 분석 방법론

1. 주요 예측 정확도 측정 지표

앞서 제1절에서 살펴본 시계열 모형, 회귀 모형, 머신러닝 기반 모형 등 다양한 예측모형의 예측 결과를 비교하기 위한 기준으로 활용되는 대표적인 지표로는 MSE(Mean Squared Error), RMSE(Root Mean Squared Error), MAE(Mean Absolute Error), MAPE(Mean Absolute Percentage Error) 등이 있다.

각 지표들은 기본적으로 예측 기간 전반의 오차(실제값과 예측값의 차이)를 합산하여 모형의 성능을 평가한다. 각 지표의 일반적인 수식과 특징 및 단점을 요약하면 다음의 <표 2-1>과 같다.

<표 2-1> 주요 예측 정확도 측정 지표

구분	수식	특징	단점
MSE	$\sum_{t=1}^N \frac{(Y_t - \hat{Y}_t)^2}{N}$	오차를 제곱함 함으로써 예측 기간 전체의 오차를 측정함	관측치와 단위가 다름 이상치에 민감하게 반응함
RMSE	$\sqrt{\sum_{t=1}^N \frac{(Y_t - \hat{Y}_t)^2}{N}}$	MSE의 단위 차이에 따른 해석의 어려움을 보완함	모수 추정 시 편향성을 가짐
MAE	$\sum_{t=1}^N \frac{ Y_t - \hat{Y}_t }{N}$	이상치에 둔감하며, 모든 오차에 동일한 가중치를 부여함	모수 추정 및 추론 시 계산이 어려움
MAPE	$\frac{100}{N} \times \sum_{t=1}^N \left \frac{Y_t - \hat{Y}_t}{Y_t} \right $	평가 지표를 비율로 확인 함으로써 직관적인 이해가 쉬움	실제값이 0인 경우 사용 불가함 비율 해석이 의미 있는 값에만 적용이 가능함

자료: 저자 작성

2. 주요 비교검증 방법론

다양한 예측모형 중 최적의 모형을 선택하는 방법에는 부트스트랩(Bootstrap), 래퍼(Wrappers), 교차검증법 등 몇 가지 방법론이 존재한다. 이 중 교차검증법은 모형의 예측력에 대한 기댓값의 통계적 추정치를 제공한다는 점에서 다른 방법론과 비교하여 활용도가 높다.³⁾ 일반적인 교차검증법은 표본을 훈련용 세트(training set)와 검증용 세트(test set)로 구분한다. 이때 시계열 자료의 경우 표본의 관측 순서가 중요함으로 세트가 연속적으로 구성되어야 한다.

교차검증법의 일반적인 단계는 다음과 같다. 첫째, 훈련용 세트를 이용하여 모형을 추정한다. 둘째, 추정된 모형을 이용하여 h 시점 앞의 예측치를 산출한다. 마지막으로 예측치를 검증용 세트의 실제값과 비교하여 모형의 성능을 평가한다. 평가 지표로는 앞서 <표 2-1>에서 제시한 여러 예측 정확도 측정 지표 중 하나를 선택하여 활용한다. 한편 교차검증법은 훈련용 세트와 검증용 세트를 구분하는 방식에 따라 여러 가지 형태가 존재하며, 본 연구에서는 여러 선행연구에서 활용된 k -Fold 교차검증법과 몬테카를로 교차검증법에 대해 보다 상세하게 기술한다.

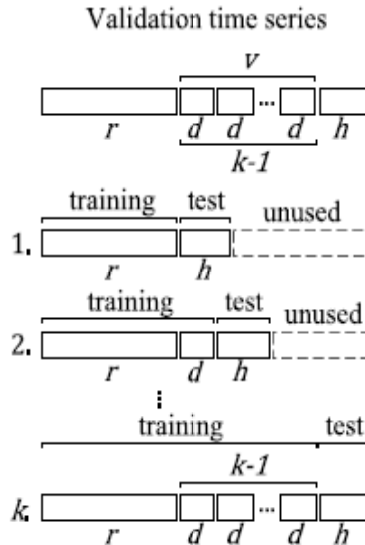
1) k -Fold 교차검증법

k -Fold 교차검증법은 먼저 표본 D 를 k 개의 부분집합 $D_1, D_2, \dots, D_k$ 로 구분한다. 이때 부분집합들은 상호 배타적(mutually exclusive)이며, 거의 동일한 크기로 구성된다. 이후 모형은 k 번 반복적으로 추정되고 검증된다. 각 시행에 있어서 모형은 D_i 는 제외한 D 를 이용하여 모형을 추정하고, 추정된 모형을 통해 예측한 D_i 와 D_i 의 실제값을 비교하여 성능을 검증하게 된다.

3) Fonseca-Delgado & Gómez-Gil(2013), pp. 2-3.

시계열 자료에 대해서도 기본 원리는 유사하나, 시계열이라는 자료의 특수성으로 인해 D 를 구성하는데 일부 차이가 존재한다. <그림 2-6>은 시계열 자료에 대한 k -Fold 교차검증법의 절차를 요약하고 있다. 먼저 모형을 훈련하기 위한 r 개의 관측치를 선정하고, 반복 추정을 위한 $k-1$ 개의 구간과 예측력 검증을 위한 h 구간을 구분한다. 여기서 반복 추정을 위한 활용되는 구간인 v 의 크기는 사용 가능한 전체 관측값인 n 개에서 r 과 h 에 할당된 표본수를 제외한 값이며, $k-1$ 개의 부분으로 나뉘며 각 부분은 d 개의 관측값을 가진다. 각 반복 시행에는 r 개의 초반 관측값이 포함되며, 이후 모형은 다음 h 개의 관측치를 사용하여 평가된다. k 번 반복 시행을 통해 산출된 k 개의 평가 지표 추정치를 평균한 값이 해당 모형의 k -Fold 교차검증 결과이다.

<그림 2-6> k -Fold 교차검증 프로세스

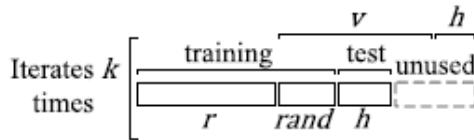


자료: Fonseca-Delgado & Gómez-Gil (2013), p. 3.

2) 몬테카를로 교차검증법

몬테카를로(Monte-Carlo) 교차검증법은 앞서 소개한 k -Fold 교차검증법과 원리는 동일하나 교차검증을 위한 V 구간을 구성하는데 난수(random number)를 활용한다는 점에서 차이가 있다. <그림 2-7>은 몬테카를로 교차검증법의 절차를 도식화하고 있다. k -Fold 교차검증과 마찬가지로 $v = n - r - h$ 이며, n 은 전체 표본수이다. k 번 반복 시행을 통해 예측치와 관측치의 차이가 평가 지표를 통해 평가되고, k 개의 평가 지표 추정치의 평균값이 해당 모형의 몬테카를로 교차검증 결과로 나타난다.

<그림 2-7> 몬테카를로 교차검증 프로세스



자료: Fonseca-Delgado & Gómez-Gil (2013), p. 3.

제3절 소결 및 시사점

이번 장에서는 항만수요예측에 활용될 수 있는 다양한 예측 방법론과 예측모형의 적합도를 평가할 수 있는 지표 및 검증법을 살펴보았다. 항만물동량 예측 연구는 과거 데이터를 활용하여 미래를 전망하는 과정으로 본질적으로 시계열 자료를 활용한다는 점에서 많은 선행연구에서 SARIMA 등 시계열 모형이 활용되었다. 또한 일부 연구에서는 경쟁관계를 파악할 수 있는 L-V 모형, 동태적 관점에서 파악하는 시스템다이나믹스 모형, 유전 알고리즘 모형, 그레이 모형 등도 활용되었다. 반면 회귀모형은 직접적인

예측에 활용되었다기보다는 대부분 타 모형의 성능을 검증하기 위해 비교 대상인 모형으로 활용되었다.

한편 머신러닝 기반 모형은 예측 문제의 목적 및 대상 분야에 따라 다양한 형태로 개발되었다. 특히 전통적 회귀분석 및 시계열분석 기법과의 연계 및 결합을 통해 예측력 및 모형 강건성을 높였다. 특히 외생변수 간 상관관계 도출이 어렵거나 시계열 데이터의 정상성 및 계절성이 부족한 데이터를 활용할 때는 전통적 통계 방법론보다 우수한 성능을 보인 사례도 있었다. 하지만 수요예측 분야에서는 초단기(1개월 이내) 및 단기(1년 이내) 예측을 위해서 많은 학습 데이터(22개월~240개월)가 요구되는 것은 머신러닝의 단점으로 작용한다. 일부 연구에서는 초장기(90년) 항만물동량 예측을 위해 동적 모델링 적용, 직관 및 인과관계를 정량화하기 위한 연구가 진행되기도 하였으나 거시경제 지표를 활용한 인과형 모형을 제시하는데 그쳤다.⁴⁾

이처럼 항만수요예측에는 다양한 방법론의 적용이 가능하며 각 방법론은 자료의 특성과 활용 목적에 따라 장단점이 존재한다. 따라서 예측모형 개발에 있어서 예측모형에 대한 충분한 이해를 바탕으로 다양한 모형들을 실험해 보고, 주어진 자료의 특성과 예측의 목적, 활용성 등을 종합적으로 고려하여야 할 필요가 있다. <표 2-2>는 앞서 살펴본 주요 수요예측 방법론의 특성과 세부 모형에 대해 요약하고 있다.

4) Dorsser et al.(2012), pp. 88-110.

〈표 2-2〉 주요 수요예측 방법론 및 특성

구분	주요 특성	비고
시계열 모형	• 장기 시계열 데이터 확보가 중요 • 일반적으로 단기 예측에 우수한 성능을 보이며, 모형 결합을 통해 예측 성능 향상 가능 • 다양한 투입변수를 활용한 논리적 설명이 가능한 모형 개발에는 한계	AR, MA, ARIMA, SARIMA, VAR, Bayesian VAR, QAR 등
회귀 모형	• 방대한 변수를 활용하여 해석이 용이하고 활용성이 높음 • 다양한 변수 활용에 있어서 다중공선성 등 문제 해결 필요 • 상대적으로 낮은 예측 성능	SLR, MLR, Logistic, Probit, Tobit, Quantile 등
머신러닝 기반 모형	• 폭넓은 변수 활용성 • 상대적으로 높은 예측 성능 • 복잡한 알고리즘으로 투입변수들의 직관적인 설명력은 낮음	SVM, Decision Tree, Random Forest, ANN, XG부스트, ANFIS 등

자료: 저자 작성

또한 예측모형 개발에 있어서 중요한 부분은 제안한 모형의 성능을 과학적인 방법으로 증명하는 것이다. 예측 정확도를 측정하는 지표는 자료의 특성을 면밀히 관측하여 선택해야 하며, 모형 간 지표의 추정치를 비교 검증하기 위해서는 부트스트랩, 교차검증법 등 학술적으로 검증된 방법론을 활용할 필요가 있다.

03

기존 비컨테이너 항만물동량 예측모형 진단

제3장에서는 기존에 활용 중인 품목별 항만물동량 예측모형의 특징 및 한계점을 진단한다. 분석대상 품목은 자동차, 석탄, 고철, 시멘트 등 4개 품목이며, 각 품목별 기존 항만물동량 예측모형 진단 단계는 연관 산업 동향 및 주요 이슈 분석, 기존 예측모형의 활용 변수 및 방법론 검토, 특징과 한계점 도출 순으로 진행된다.

제1절 자동차 항만물동량 예측모형 진단

1. 국내외 자동차산업 동향 및 주요 이슈

1) 세계 자동차산업 동향

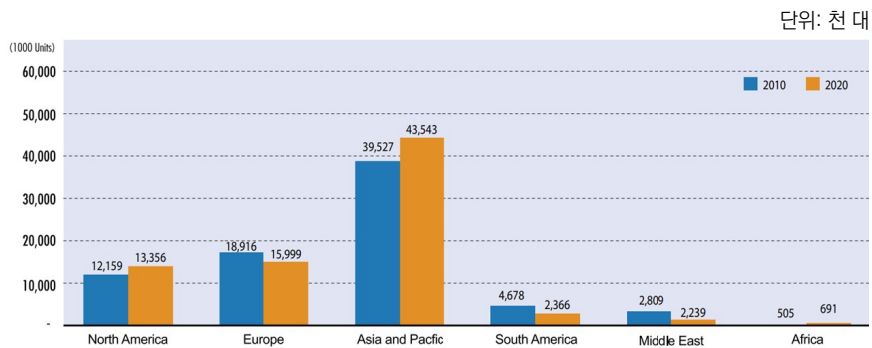
(1) 세계 자동차 생산 및 판매

2021년 글로벌 자동차 생산은 전년도(2020년, -15.5%) 기저효과에도 불구하고 2.0% 증가한 7,978만 대에 그쳤으며, 이는 코로나19 이전인 2019년 9,258만 대에 한참 못 미치는 수준이다. 국가별로는 중국이 전년

대비 3.4% 증가한 2,608만 대를 기록하며 1위를 유지하였고, 이어서 미국이 915만 대 (+3.8%)로 2위, 일본이 785만 대(-2.7%)로 3위를 기록하였다. 인도는 전년대비 29.6% 증가한 440만 대를 기록하며 2020년 6위에서 2021년 4위로 순위가 상승한 반면, 2020년 4위였던 독일은 2021년 8.8%가 감소한 345만 대를 기록하며 6위로 내려앉았다. 한편 2021년 346만 대로 전년대비 1.3% 감소한 우리나라는 2020년에 이어 5위를 유지했다.⁵⁾

지역별 자동차 생산 추이를 살펴보기 위해 2010년과 2020년의 지역별 생산량을 비교한 결과, 2020년 코로나19 영향으로 생산량이 급감한 점을 감안하더라도 아시아·태평양 지역과 북미 지역은 각각 10.2%, 9.8% 증가하면서 견고한 성장세를 보여줬다. 반면 유럽은 -15.4%, 남미는 -49.4%, 중동은 -20.3%를 기록하며 성장세가 정체 또는 하락하였다.⁶⁾

〈그림 3-1〉 2010년과 2020년의 세계 지역별 자동차 생산



자료: 한국자동차산업협회(2021b), p. 4.

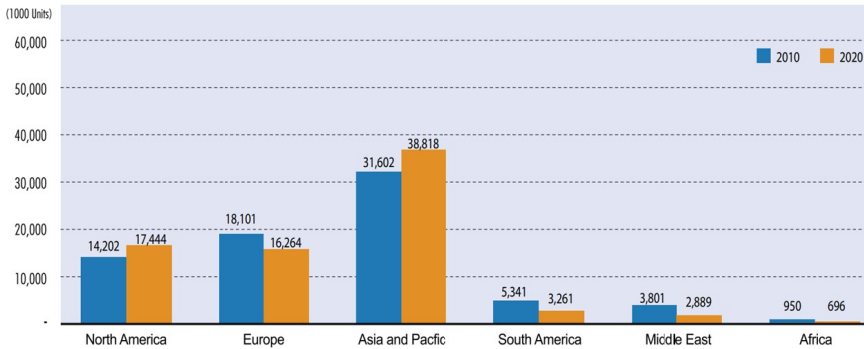
2020년 세계 판매량은 2010년 대비 7.3% 증가한 7,937만 대를 기록하였으며, 지역별로는 생산량 증감 추이와 유사한 형태를 보였다. 아시아·태평양 지역과 북미 지역이 각각 22.8% 증가한 반면, 유럽 지역, 남미 지역, 중동 지역은 각각 -10.1%, -38.9%, -24.0%를 기록하였다.⁷⁾

5) 한국자동차산업협회(2022a), pp. 1-4.

6) 한국자동차산업협회(2021b), p. 4.

〈그림 3-2〉 2010년과 2020년의 세계 지역별 자동차 판매

단위: 천 대

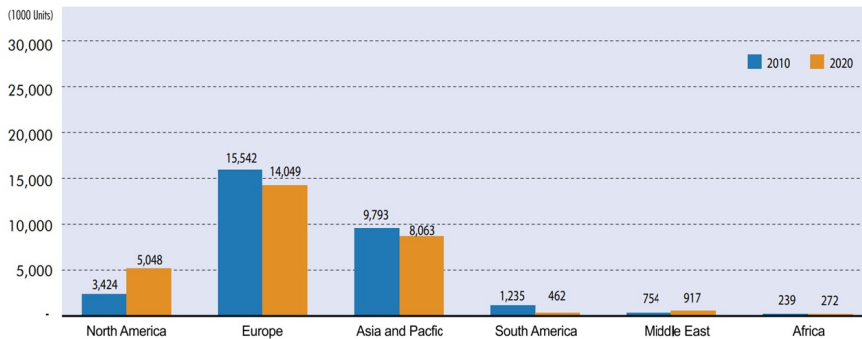


자료: 한국자동차산업협회(2021b), p. 4.

2020년 세계 수출량은 2010년 대비 7.0% 감소한 2,881만 대를 기록하였다. 2020년 기준 지역별 비중은 유럽 지역이 48.8%를 차지하였으며, 이어서 아시아·태평양(28.0%), 북미(17.5%) 순으로 나타났다. 전반적으로 2020년 수출량이 급감한 가운데 북미 지역이 2010년 대비 47.5% 증가를 보여 내수 대비 수출의 급격한 성장세를 나타냈다.⁸⁾

〈그림 3-3〉 2010년과 2020년의 세계 지역별 자동차 수출

단위: 천 대



자료: 한국자동차산업협회(2021b), p. 5.

7) 한국자동차산업협회(2021b), p. 4.

8) 한국자동차산업협회(2021b), p. 5.

(2) 주요 이슈

가. 생산 차질 요인 지속 발생

2021년 세계 자동차 생산은 코로나19 상황이 2년째 지속되고 반도체 등 공급망 위기가 연이어 발생하면서 소폭 반등에 그쳤다. 2021년에도 코로나19는 변이 발생 및 재유행이 반복되면서 자동차 생산 공장 가동을 저하, 수출입 물류 여건 악화, 구매자의 소비 위축 등을 촉발했다. 또한 차량용 반도체 공급 부족 상황은 2020년 반도체 수요예측에 실패한 것에 이어 2021년에는 각종 글로벌 반도체 공급 차질 요인이 발생했다. 미국 NXP사 인피니언 오스틴 공장은 지난 2월 텍사스 한파의 영향으로 생산라인이 한 달 가량 멈췄으며, 말레이시아 인피니언 공장은 6월에 코로나19 확산으로 전국 봉쇄령이 내려지며 반도체 생산이 전면 중단되었다. 또한 대만 TSMC 사는 공업용수 및 생산여력 부족으로 차량용 반도체 생산을 후순위로 미룸으로써 공급 차질을 확산시켰다.

이처럼 공급망 위기 등 악재가 지속되면서 2021년에는 내수 시장 규모에 따른 양극화 현상이 두드러지게 나타났다. 중국, 미국, 인도, 브라질 등 내수시장 규모가 크거나, 생산 중 내수 비중이 큰 국가들은 2020년 코로나19 충격에서 반등에 성공한 반면, 일본, 한국, 독일 등 수출 비중이 높은 국가들은 글로벌 경기 침체, 해운·항만 대란 등으로 수출회복이 제한되며 생산 감소가 지속되었다.⁹⁾

나. 전기차 시장 확대

2021년 세계 자동차 판매는 2020년의 기저효과에도 불구하고 전년대비 4.3% 증가한 8,455만 대에 그쳤다. 반면 전기차 판매는 주요 시장의 환경 규제 강화, 각국 정부의 적극적인 지원정책 추진, 제작사의 전기차 등 고부가

9) 한국자동차산업협회(2022a), pp. 4-6.

가치 차량 생산 집중 등의 영향으로 전년대비 110.7% 증가한 666만 대가 판매되어 역대 최다 판매를 기록하였다. 전체 자동차 판매 대수에서 차지하는 비중도 꾸준히 증가하여 2021년 기준 전기차의 비중은 7.9%에 이른다.¹⁰⁾

2021년 전기차 시장은 아시아, 유럽, 미국 등 세계 전 지역에서 고르게 성장세를 기록하였으며, 특히 전체 전기차 판매의 50% 이상을 차지한 중국의 판매 확대가 전체 시장의 성장을 견인하였다. 중국 전기차 판매는 전년대비 166.6% 증가한 333만 대를 기록하였으며, 중국 전체 자동차 판매에서 차지하는 비중도 전년 4.9%에서 12.7%로 크게 상승하였다. 이밖에 유럽은 전년대비 66.1% 증가한 234만 대, 미국은 101.3% 증가한 67만 대를 판매하였다.¹¹⁾

(3) 주요국 정책 동향

세계 주요국은 2021년에도 전기차 구매 시 보조금 지급 등 강력한 전기차 보급정책을 추진하고 있으며, 이를 통해 전기차 산업육성과 수출 확대를 목표로 하고 있다. 주요국의 자동차산업 정책은 다음의 <표 3-1>에 요약되어 있다.

<표 3-1> 주요국 자동차산업 정책 동향

구분	주요내용
미국	• 바이든 정부, 2030년까지 신차판매의 50%를 전동화하는 방침 발표('21.8월) • EPA(환경보호국), 신차 배기가스 배출량 요건 강화('21.12월) - 이번 조치로 연비 기준을 2026년까지 55mpg로 상향(현재 40mpg) - EPA는 이 규제에 따라 배출량 감소 혜택이 2050년까지 최대 199억 달러에 달하고 운전자가 연료비를 2,100억~4,200억 달러 절약할 수 있을 것으로 기대
독일	• 총선 결과 진보 성향의 연립정부가 공식 출범('21.12월)함으로써 탄소중립 정책 기조 강화 전망 • 전기차 관련 정책으로 내연기관차 신규등록 중단('35년), 전기차 30년까지 1,500만 대 보급(현재 120만 대 수준) 등 추진

10) 한국자동차산업협회(2022c), p. 1.

11) 한국자동차산업협회(2022c), pp. 2-4.

구분	주요내용
중국	• 전기차 시장이 성숙기에 진입한 것으로 판단하여 '21년 전기차 보조금을 전년대비 20% 하향 조정하는 등 전기차 보조금 단계적 삭감 추진 • 각 기업이 생산한 자동차의 평균 연료 소모량에 따라 (-)포인트, 신에너지차 생산에 (+)포인트를 지급하는 '더블포인트 제도' 추진
인도	• 'Make in India' 정책 일환으로 자국내 첨단제품 생산 장려를 위해 생산연동인센티브(PLI) 제도 시행 • 전기차 보급률 증가를 위해 고율의 전기차 수입관세 인하 검토
러시아	• 우크라이나 침공 이후 글로벌 메이커의 러시아 내 생산중단 및 가동 축소 • 원자재 및 부품 수급 차질에 따라 글로벌 완성차 생산 차질 • 전기차 산업육성을 위해 향후 10년간 10조 원 투입 계획('21.8월) 발표

자료: 한국자동차산업협회(2022b), pp. 7-10.

2) 국내 자동차산업 동향

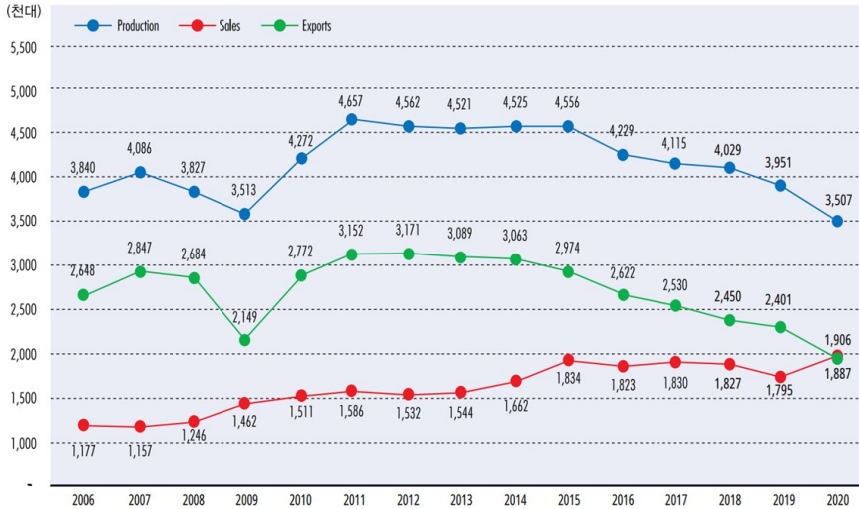
(1) 국내 자동차 생산 및 판매

2021년 우리나라의 자동차 생산은 전년대비 1.3% 감소한 346만 대를 기록하여 2016년 이후 6년 연속 감소하였으나, 글로벌 반도체 공급 부족 등 연이은 악재 속에서 재고관리와 생산조정을 통해 감소폭을 최소화하며 세계 5위를 유지하였다. 내수 판매량은 전년도 내수증가에 따른 선수요 반영으로 역기저 효과가 발생하였으며 차량용 반도체 생산 지연으로 인기차종의 생산 차질이 발생하며 전년대비 9.0% 하락한 173만 대를 기록하였다. 한편 수출은 전년도 기저효과에도 불구하고 국내 제작사의 전략차종 생산 지연 등으로 8.2% 증가한 204만 대에 그쳤다.¹²⁾

12) 한국자동차산업협회(2022a), p. 3.

〈그림 3-4〉 우리나라 자동차 생산·내수·수출 추이

단위: 천 대



자료: 한국자동차산업협회(2021b), p. 172.

(2) 주요 이슈

가. 생산 차질 지속

2021년 국내 자동차산업은 세계 시장과 마찬가지로 코로나19 장기화와 차량용 반도체 부족에 따른 생산 차질이 지속되었다. 생산은 전년대비 1.3% 감소한 346만 대, 내수 판매는 10.6% 감소한 144만 대를 기록한 반면, 수출은 전년도 코로나19로 인한 기저효과와 미국, EU 등 주요시장의 회복세에 힘입어 전년대비 8.2% 증가한 204만 대를 기록하였다.¹³⁾

나. 전기차 시장의 성장

2021년 전기차 판매는 전년대비 103.8% 증가한 12만 8,629대가 판매되어 국내 자동차 시장에서 전기차의 비중이 전년대비 4.1%p 증가한

13) 한국자동차산업협회(2021a), pp. 1-7.

7.4%로 역대 최대 판매를 기록하였다. 차종별로 살펴보면, 승용전기차는 저·무공해차 보급목표 강화, 보조금 규모 확대, 제작사의 신모델 투입 확대 등으로 전년대비 128.7% 증가한 7.1만 대가 판매되었으며, 전기화물차는 전기대비 92.3%, 전기버스는 26.1% 각각 판매가 증가하였다.¹⁴⁾

다. 수입차 시장 지속 성장

국내 자동차 수요의 특징 중 하나는 수입차 시장의 지속적인 성장이다. 국내 자동차 생산이 2016년 이후 6년 연속 감소한 반면 수입차 시장은 2020년까지 꾸준히 성장하였다. 2021년에는 국내생산 차량의 내수 판매량이 전년대비 10.6% 감소한 것에 비해 해외브랜드 수입은 6.8% 감소한 29만 9,263대를 기록하여 상대적으로 시장의 안정적인 성장세를 유지하였다.

(3) 국내 정책 동향

국내 자동차 시장은 2021년 7월부터 본격적으로 반도체 공급 부족에 따른 생산 차질 영향이 나타나고 있으며, 이에 정부는 차량용 반도체 수입 관련 통관 절차를 간소화하고 해외출장 시 자가격리 면제 등의 지원을 통해 업계의 생산 차질을 최소화하기 위한 조치를 취하고 있다.

또한 정부는 코로나19로 인한 민간 소비 위축에 대응하여 2020년에 단행한 개별소비세 인하를 2021년에도 연장하여 내수활성화 정책을 추진하였으나, 반도체 공급 부족에 따른 생산차질로 내수는 감소세를 지속하였다.

한편 친환경차 보급 확대 정책으로 친환경차 취득세 감면(하이브리드차 40만 원, 전기·수소차 140만 원) 및 보조금 지급을 지속적으로 추진하여 소비자 구매 부담을 완화시켰으며, 이는 친환경차의 지속적인 내수 판매 증가로 이어지고 있다.

14) 한국자동차산업협회(2022c), p. 8.

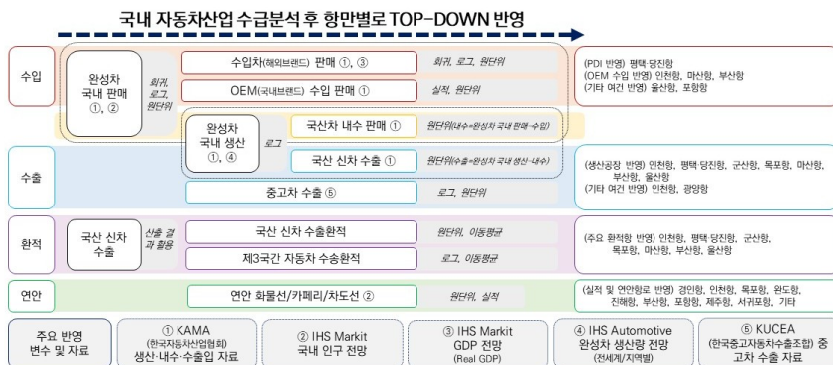
2. 기존 자동차 항만물동량 예측모형 분석

1) 분석 개요

자동차 항만물동량은 수입, 수출, 환적(입, 출항), 연안 물동량으로 구분하고 있으며, 국내 자동차산업의 수급분석을 실시하여 각 물동량을 예측한 후 항만별로 물동량을 배분하는 Top-down 방식과 개별 항만의 특성을 반영하여 항별 물동량을 예측한 후 합산하여 전체 물동량을 산출하는 Bottom-up 방식을 혼용하고 있다.

수입물동량은 국내브랜드의 해외 OEM 생산 중 수입량, 해외브랜드의 수입량을 각각 예측 기간 동안 추정하여 합산한다. 수출물동량은 국내에서 생산한 완성차의 신차 수출량과 중고차 수출로 구성되며, 환적물동량은 국산 신차 수출환적과 제3국 간 자동차 수송환적으로 구성된다. 한편 연안물동량은 연안 화물선, 카페리, 차도선의 수송 실적을 반영하여 예측을 실시한다. 기존 자동차 물동량 예측모형의 대략적인 프로세스는 다음의 <그림 3-5>에서 확인할 수 있다.

<그림 3-5> 자동차 물동량 예측 프로세스



자료: 한국해양수산개발원(2022). p. 106.

2) 예측 방법론

(1) 수입물동량

가. 수입물동량 구분 및 대상 항만 분류

해양수산부 해운항만물류정보시스템(Port-MIS)의 항만물동량 통계를 기반으로 우리나라 무역항 중 자동차 수입물동량이 유의미한 수준으로 발생하는 항만은 인천항, 평택·당진항, 군산항, 목포항, 광양항, 마산항, 부산항, 울산항, 포항항 등 9개 항만이다. 자동차 수입물동량은 해당 9개 항만을 대상으로 예측 기간 동안의 예측을 수행한다. 옥포항, 동해·묵호항, 속초항 등에서도 간헐적으로 수입물동량이 관측되나 해당 항만은 관측 빈도와 물동량 규모 측면에서 자동차 전용부두 건설을 위한 최소 기준에 미달하여 분석에서는 제외한다.

수입물동량은 발생원에 따라 국내브랜드의 해외 OEM 생산량 중 국내역수입량과 해외브랜드 수입량으로 구분되며, 기타 신차수출량 중 해외수출을 위해 항내 이동한 물동량과 중계무역 물동량이 통계상 수입으로 계상됨에 따라 수입물동량에 포함된다. 또한 각 항만은 배후 산업단지와 항로 개설 유무에 따라 특정 성격의 수입물동량을 취급하고 있다. 인천항, 마산항, 부산항의 수입물동량은 대부분 해외 OEM 수입이며, 평택·당진항, 군산항, 목포항, 광양항은 해외브랜드 수입, 울산항은 수출항내이동 물동량, 포항항은 중계무역 물동량이 발생한다.

나. 분류별 전망

해외 OEM 수입이 발생하는 인천항, 마산항, 부산항의 항별 수입물동량 전망은 단순이동평균법을 사용하며, 이동평균 기간은 최근 3년으로 설정한다.

해외브랜드 수입물동량은 한국자동차산업협회의 해외브랜드 수입량을 활용하여 회귀 모형, 로그 추세 모형, 원단위 모형을 각각 적용하여 예측한

후 전망치의 평균을 산출한다. 각 모형을 통해 수입량을 추정하는 과정은 다음과 같다. 먼저 회귀 모형은 해외브랜드 수입 실적을 종속변수, IHS Markit의 우리나라 GDP(real, harmonized)를 독립변수로 활용하여 계수를 추정한 후 예측 기간 동안의 GDP전망치를 산입하여 해외브랜드 수입량을 추정한다. 회귀 모형을 적용한 분석결과는 다음의 <표 3-2>와 같다.

<표 3-2> 해외브랜드 수입 회귀 모형 분석결과

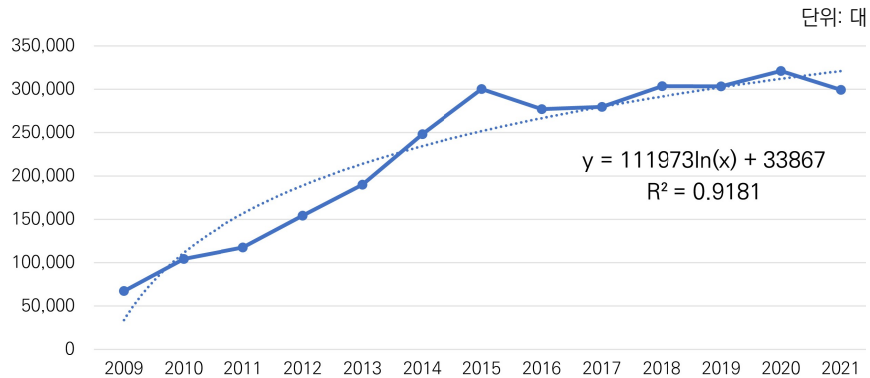
변수	계수	표준오차	t 검정	F 검정	adj. R ²
Intercept	-353539	33960.57	-10.4103***	231.8426***	0.917
GDP	399.5013	26.23745	15.22638***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함

자료: 저자 작성

로그함수를 활용한 추세 분석은 추세가 비교적 안정화되기 시작한 2009년부터 2021년까지의 해외브랜드 수입 실적을 로그함수로 추정하며, 분석 결과는 <그림 3-6>과 같다.

<그림 3-6> 해외브랜드 수입 로그 추세 모형 분석결과

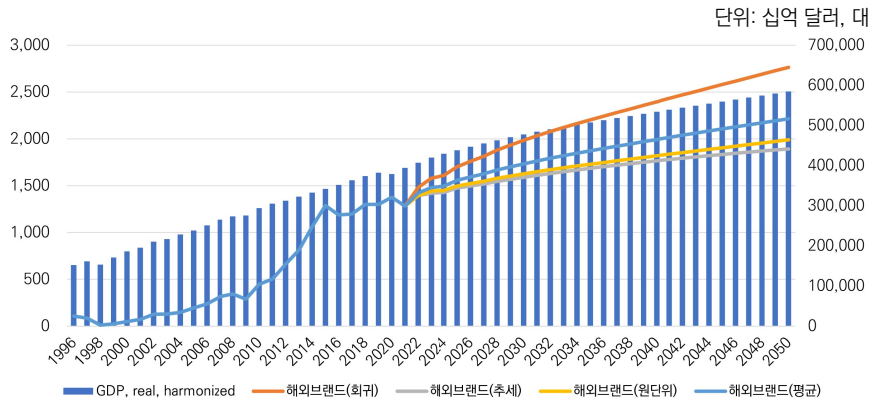


자료: 저자 작성

마지막으로 원단위 모형은 해외브랜드 수입 실적과 우리나라 GDP(real, harmonized)의 비율(ratio)을 산출한 후 단순이동평균법(최근 3년)으로

예측 기간 전망치를 추정한다. 최종적으로는 3개의 모형으로부터 산출된 예측치를 산술 평균하여 해외브랜드 수입량을 전망한다.

〈그림 3-7〉 해외브랜드 수입차 판매량 예측결과



자료: 저자 작성

이후 해외브랜드 수입차 판매량 전망치의 연평균 증가율을 평택·당진항의 수입물동량 실적에 반영하여 예측 기간 동안의 전망치를 산출한다. 한편 기타 해외브랜드 수입물동량이 발생하는 군산항, 목포항, 광양항은 전체 해외브랜드 수입물동량에서 차지하는 비중이 매우 작아 직접적으로 연평균 증가율을 적용하지 않고, 항별 수입물동량 실적에 단순이동평균법(최근 5개년)으로 예측 기간 동안의 전망치를 산출한 후 연평균 증가율을 구해서 직전년도 수입물동량 실적에 적용하여 항별 수입물동량을 전망한다.

울산항의 수출항내이동 물동량은 먼저 국내 생산 신차 수출량을 전망한 후 수출물동량 중 울산항 처리 비중을 항내이동 물동량 비중으로 가정한다. 이후 단순이동평균법으로 예측 기간 비중을 전망하고, 이를 신차 수출물동량에 적용하여 최종적으로 울산항의 수입물동량을 전망한다.

중계무역 중 수입물동량이 발생하는 포항항의 경우는 2013년 이후 대체로 일정 수준의 물동량이 안정적으로 발생하고 있으며, 포항항의 수입물동량

실적을 대상으로 단순이동평균법(최근 5개년)을 적용하여 예측 기간 동안의 전망치를 추정한다.

다. 전체 수입물동량 산출

최종적으로 수입물동량이 발생하는 9개 항만의 수입물동량을 2050년까지 예측한 후 합산함으로써 전체 수입물동량의 2050년까지의 예측치를 산출할 수 있다.

(2) 수출물동량

가. 수출물동량 구분 및 대상 항만 분류

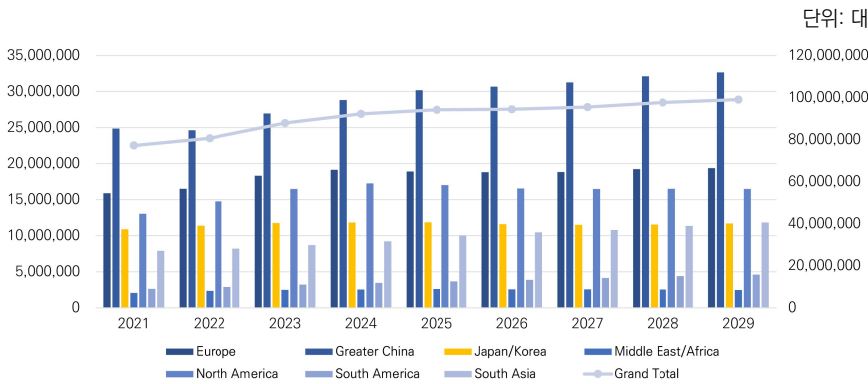
수출물동량은 배후에 자동차생산시설이 위치한 항만을 중심으로 발생하고 있으며, 세계 자동차 생산량 대비 국내 자동차 생산량 실적 및 전망치를 활용하여 예측을 수행한다. 해양수산부 Port-MIS의 항만물동량 통계를 기반으로 우리나라 무역항 중 자동차 수출물동량이 유의미한 수준으로 발생하는 항만은 인천항, 평택·당진항, 군산항, 목포항, 광양항, 마산항, 부산항, 울산항, 포항항 등 9개 항만이다. 자동차 수출물동량은 해당 9개 항만을 대상으로 전망한다. 포항항을 제외한 8개 항만에서는 국산차 수출물동량이 발생하며, 포항항은 중계무역을 통한 수출물동량이 일부 발생한다.

나. 국내 생산 신차 수출량 전망

국내 생산 신차 수출량 전망은 국산차 생산량과 내수판매량을 각각 전망한 후 국내 생산량과 내수판매량의 차이를 국산차 수출량으로 가정하여 국산차 수출량을 전망한다. 이후 국산차 수출량 증가율 전망치를 신차 수출물동량 실적에 적용하여 물동량을 전망하고 연평균증가율로 보정한다. 최종적으로는 인천항에서 발생하는 중고차 수출량을 차감하여 국내 생산 신차 수출물동량을 전망한다.

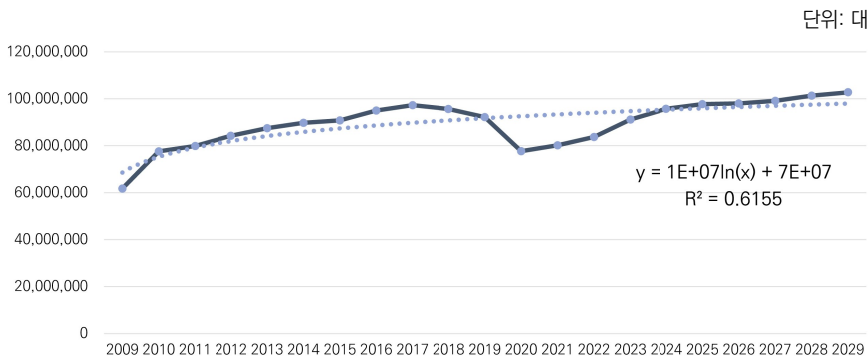
첫 번째 단계인 국산차 생산량 전망에는 IHS Markit의 글로벌 자동차 생산 전망자료 중 2022년에서 2029년 기간의 글로벌 생산량 전망치와 일본/한국 생산량 전망치의 연도별 증감율을 각각 실적치에 적용하여 2022년~2029년 기간 자동차 생산량을 산출한다. 이후 2009년~2029년 기간의 생산량 실적 및 전망치에 로그 추세 함수를 적용하여 예측 모형을 추정하고 추정된 계수를 활용하여 2050년까지의 전망치를 산출한다.

〈그림 3-8〉 2022~2029년 기간 글로벌 자동차 생산 전망



자료: IHS Markit의 글로벌 자동차 생산 전망자료

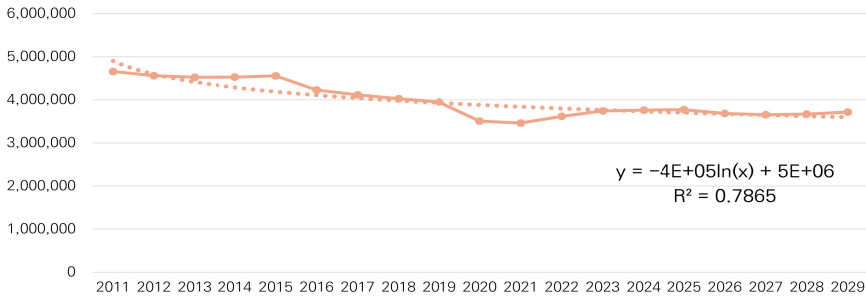
〈그림 3-9〉 2009~2029년 기간 글로벌 자동차 생산 로그 추세 모형 분석결과



자료: 저자 작성

〈그림 3-10〉 2011~2029년 기간 국내 자동차 생산 로그 추세 모형 분석결과

단위: 대



자료: 저자 작성

두 번째 단계인 내수판매량 전망은 국내 자동차 수요가 내수판매량과 수입대수의 합으로 이뤄진다는 가정 하에 국내 수요를 전망한 후 수입량을 차감함으로써 내수판매량을 전망한다.

국내 자동차 수요 전망 방법은 앞서 해외브랜드 수입량과 유사하게 회귀 모형, 로그 추세 모형, 원단위 모형으로 각각 추정한 후 세 가지 전망치를 산술평균하여 산출한다. 각 모형을 통해 국내 수요를 추정하는 과정은 다음과 같다. 먼저 국내 수요 실적치는 국내 생산 차량의 내수판매량 실적과 해외브랜드 수입 및 해외 OEM 수입량을 합산하여 실적치 자료를 작성한다. 이후 회귀 모형은 국내 수요 실적치를 종속변수, IHS Markit의 우리나라 인구를 독립변수로 활용하여 계수를 추정한 후 예측 기간 동안의 인구 전망치를 산입하여 국내 수요를 추정한다. 회귀 모형을 적용한 분석결과는 다음의 〈표 3-3〉과 같다.

〈표 3-3〉 국내 자동차 수요 회귀 모형 분석결과

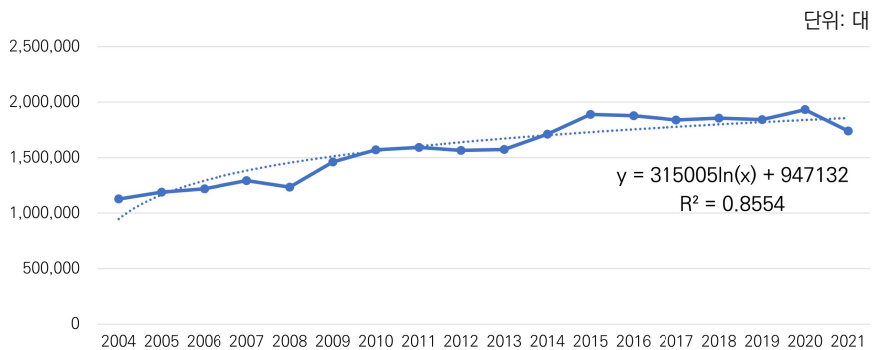
변수	계수	표준오차	t 검정	F 검정	adj. R ²
Intercept	-5655001	1274018	-4.43872***	32.02611***	0.58511
GDP	145381.3	25689.54	5.659161***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함

자료: 저자 작성

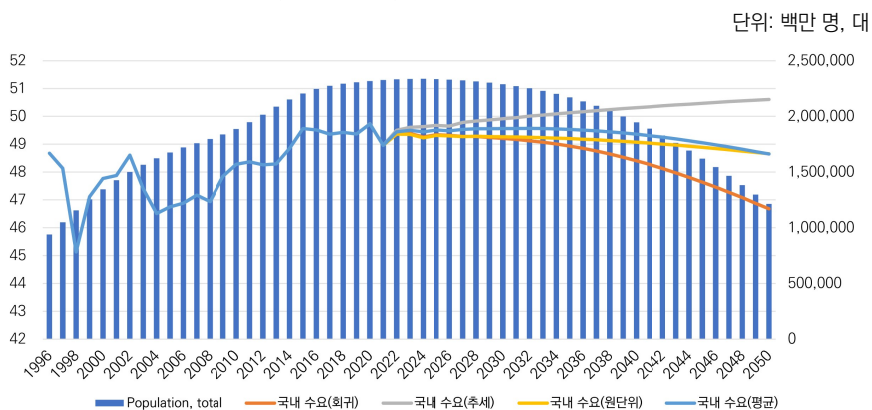
로그함수를 활용한 추세 분석은 추세가 비교적 안정화되기 시작한 2004년부터 2021년까지의 국내 수요 실적을 로그함수로 추정하며, 분석결과는 <그림 3-11>과 같다. 마지막으로 원단위 모형은 국내 수요 실적과 우리나라 인구의 비율을 산출한 후 단순이동평균법(최근 3년)으로 예측 기간 전망치를 추정한다. 최종적으로는 세 가지 모형으로부터 산출된 예측치를 산술 평균한 결과는 <그림 3-12>와 같다.

<그림 3-11> 국내 자동차 수요 로그 추세 모형 분석결과



자료: 저자 작성

<그림 3-12> 국내 수요 예측결과



자료: 저자 작성

이어서 세 번째 단계인 국산차 수출량 전망은 ‘생산량-내수판매량’과 ‘내수판매량 증가율과 수출(생산량-내수판매량) 증가율 중 낮은 증가율을 반영한 수출량’의 평균으로 산출한다. 이후 최종 국내 생산 신차 수출 물동량은 국산차 수출량 증가율 전망치를 적용하여 물동량을 전망한 후 연평균증가율로 보정을 실시한다. 또한 인천항의 수출물동량에는 중고차 수출물동량이 포함되어 있으므로 신차 수출량 전망을 위해서는 중고차 수출량을 별도로 전망하고 전체 수출 물동량에서 차감함으로써 신차 수출물동량을 산출할 수 있다.

다. 중고차 수출물동량 전망

인천항의 중고차 수출물동량 전망은 한국중고차수출조합의 중고차 수출 실적과 한국무역협회의 항만별 중고차 수출입 자료 및 해양수산부 Port-MIS의 인천항 부두별 물동량 처리실적을 활용하여 예측한다. 한국무역협회의 항만별 중고차 수출입 통계는 중고차 관련 주요 HS코드 25개를 활용하여 작성한다. 관련 HS코드는 <표 3-4>와 같다.

<표 3-4> 중고차 관련 주요 HS코드

8701202000, 8701901020, 8702101020, 8702102020, 8702103020,
8702901020, 8702902020, 8702903020, 8703218000, 8703228000,
8703231020, 8703239020, 8703241020, 8703249020, 8703318000,
8703321020, 8703329020, 8703338000, 8704211020, 8704221012,
8704221092, 8704231020, 8704311020, 8704321020, 8704901020.

자료: 한국무역통계진흥원 HS CODE 내비게이션을 활용하여 저자 작성

2015년부터 2021년 기간의 중고차 수출 실적을 로그 추세 함수를 적용하고 추정된 계수를 활용하여 2050년까지의 중고차 수출량을 전망한다. 전체 중고차 수출 실적에서 인천항의 비중을 산출한 후 단순이동평균법(최근 3년)을 적용하여 2050년까지의 인천항 비중을 작성하고 전체 중고차

수출량에 곱해 인천항의 중고차 수출량을 산출한다. 이후 연도별 증감률을 산출하고 인천항 내항 제5부두에서 처리된 자동차 물동량을 중고차로 추정하고 연도별 증감률을 실적에 반영하여 2050년까지의 인천항 중고차 수출 물동량을 전망한다.

라. 항만별 수출물동량 배분 및 합산

신차 수출물동량을 취급하는 인천항, 평택·당진항, 군산항, 목포항, 광양항, 마산항, 부산항, 울산항은 항별 신차 수출물동량 비중을 산출한 후 단순이동평균법(최근 3년)을 적용하여 2050년까지의 비중 전망치를 산출한다. 각 항만의 수출물동량에 비중 전망치를 적용하여 항별 신차 수출물동량을 전망한다.

인천항은 별도로 예측한 중고차 수출물동량을 신차 수출물동량에 합산하여 인천항의 수출물동량을 전망하고, 중계무역 수출물동량이 발생하는 포항항의 경우 단순이동평균법(최근 3년)을 수출물동량 실적치에 적용하여 2050년까지의 수출물동량을 전망한다. 이후 최종적으로 수출물동량이 발생하는 9개 항만의 2050년까지의 예측치를 합산함으로써 전체 수출물동량을 전망할 수 있다.

(3) 환적물동량

가. 환적물동량 구분 및 대상 항만 분류

해양수산부 Port-MIS의 항만물동량 통계를 기반으로 우리나라 무역항 중 자동차 환적물동량이 유의미한 수준으로 발생하는 항만은 인천항, 평택·당진항, 군산항, 목포항, 광양항, 마산항, 부산항, 울산항 등 8개 항만이다. 자동차 환적물동량은 해당 8개 항만을 대상으로 예측 기간 동안의 예측을 수행한다.

환적물동량은 해외에서 생산된 차량이 국내 항만을 거쳐 제3국으로 수송되는 물동량인 제3국 환적물동량과 국내에서 생산된 차량이 해외로 수출되기 전에 국내의 다른 항만으로 이동하였다가 수출되는 수출자동차 환적물동량으로 구분된다.

나. 환적물동량 전망

환적물동량은 해양수산부 Port-MIS에서 제3국 환적물동량과 수출자동차 환적물동량으로 구분이 가능하였던 2012년~2017년 기간의 자료를 기반으로 제3국 환적물동량과 수출자동차 환적물동량을 각각 함수로 추정하고 2050년까지 연장하여 물동량을 추정한다. 각 환적물동량 함수 추정 방법은 로그 추세 함수와 단순이동평균법(최근 3개년), 그리고 2017년 기준 환적물동량 규모 유지의 3가지 방법으로 2050년까지의 물동량을 추정하고 산술평균하여 최종 환적물동량을 산출한다.

환적물동량은 수입환적물동량과 수출환적물동량이 일치한다는 가정하에 일차적으로 추정된 환적물동량을 수입환적과 수출환적에 각각 적용한다. 다만 광양항의 수출환적물동량의 경우 2018년부터 울산항 수출물동량이 광양항에서 환적하여 해외로 수출하는 것이 광양항의 수출환적물동량으로만 계상됨이 확인되어 울산항의 수출물동량 전망치를 광양항 수출환적물동량에 합산하여 전망치를 산출한다.

다. 전체 환적물동량 산출

최종적으로 환적물동량이 발생하는 8개 항만의 환적물동량을 2050년까지 예측한 결과를 합산함으로써 전체 환적물동량의 2050년까지의 예측치를 산출할 수 있다.

(4) 연안물동량

가. 연안물동량 구분 및 대상 항만 분류

해양수산부 Port-MIS의 항만물동량 통계를 기반으로 우리나라 무역항 중 자동차 연안물동량이 유의미한 수준으로 발생하는 항만은 경인항, 인천항, 목포항, 완도항, 진해항, 부산항, 포항항, 제주항, 서귀포항 등 9개 항만이며 이외의 항만에서 발생하는 연안물동량은 기타항으로 별도로 집계한다. 자동차 환적물동량은 해당 9개 항만과 기타항을 대상으로 예측 기간 동안의 예측을 수행한다.

연안물동량은 우리나라 대부분의 항만에서 불규칙적으로 발생하며, 목포항과 완도항의 경우 제주도 및 인근 도서 지역을 왕복하는 Ro-Ro 화물선과 카페리선의 물동량이 연안물동량으로 집계된다. 또한 연안물동량은 Port-MIS에 신고되는 물동량뿐만 아니라 카페리선사의 집계 물동량이 추가되거나 집계 기준이 변경됨에 따라 증가하기도 한다.¹⁵⁾

나. 전체 연안물동량 추정

연안물동량은 Port-MIS의 연안물동량 실적과 IHS Markit에서 발표한 우리나라 인구의 비율을 산출하고, 단순이동평균법(최근 3년)을 활용하여 2050년까지의 비율 전망치를 계산한다. 이후 비율 전망치를 우리나라 인구 전망치에 반영하여 2050년까지의 연안물동량 전망치를 산출한다.

다. 항만별 연안물동량 배분 및 전체 연안물동량 예측치 보정

앞서 산출한 연안물동량 전망치의 연도별 증감률을 연안물동량 발생 9개 항만 및 기타항에 적용하여 항별 연안물동량 전망치를 산출한다. 이어서 최종적으로 항별 연안물동량을 합산함으로써 예측 기간 동안의 전체 연안물동량 예측치를 전망한다.

15) 한국해양수산개발원(2022), p. 111.

3. 기존 예측모형의 특징 및 한계

기존 자동차 항만물동량 예측모형은 단일 수리모형과 과거 실적 데이터만을 활용한 예측치 추정에서 벗어나 항만별 물동량의 특성과 최근 변동사항 등을 적극적으로 반영하여 예측한 것이 특징이다. 또한 제한된 양과 질의 데이터를 바탕으로 2050년까지의 초장기 예측을 수행해야 하는 만큼 가능한 범위 내에서 여러 가지 예측모형을 적용하여 추정한 후 평균치를 활용하는 방식으로 예측치의 편향성이 강화되지 않도록 하였다.

다만 기존 예측모형이 감내할 수밖에 없는 한계점도 존재한다. 먼저 자동차 항만물동량 예측의 기본이 되는 국내 자동차 수급분석에 있어서 단일 변수를 활용한 단순선형회귀 모형을 주로 활용함에 따라 급변하는 대내외 환경변화를 모형에 반영하기 어렵다. 우리나라 인구수를 단일 설명변수로 활용하는 국내 자동차 수요 추정 모형은 직관적인 이해가 쉽고 모형의 단순성 대비 설명력이 높다는 장점이 있다. 다만 복합적인 경제 환경에 따라 발생한 생산 및 수요 차질 영향을 인구수의 변화만으로 반영할 수 없다. 이에 국내 자동차 수요에 영향을 미칠 수 있는 다양한 변수를 발굴하고 모형에 투입함으로써 보다 현실적인 모형으로의 개선이 필요하다.

또한 지속적인 수입차 비중 증가를 반영한 수입물동량 예측모형 고도화가 필요하다. 앞서 살펴본 바와 같이 국내 자동차 시장의 주요 특징은 전기차 시장 확대와 더불어 수입차 시장의 급격한 성장이다. 최근 국내 자동차 생산 및 내수 판매는 정체 또는 소폭 감소하는 추세인 반면 수입은 꾸준히 증가하고 있다. 기존 수입자동차 수요 예측모형에서는 우리나라 연도별 GDP를 단일 설명변수로 하는 단순선형회귀 모형을 활용하고 있으나, 2000년대 이후 급격히 증가한 수입차 수요를 GDP와의 선형관계로 설명하는 것에는 한계가 있다. 이에 추가적인 변수 발굴 및 일반적인 선형회귀 모형에서 벗어난 방법론을 다각적으로 검토할 필요가 있다.

이밖에도 예측모형 고도화를 위해 Port-MIS 상에서 집계되는 통계를 모니터링 함으로써 모형에 투입되는 통계의 품질 향상이 시급하며, 최근 지자체의 보조금 지급 등으로 급격히 증가한 환적물동량의 추이를 지속적으로 관찰하면서 예측모형 개선 시점을 판단해야할 것으로 보인다.

제2절 석탄 항만물동량 예측모형 진단

1. 국내외 석탄산업 동향 및 주요 이슈

1) 세계 석탄산업 동향 및 주요 이슈

(1) 세계 석탄산업 동향

국제에너지기구(IEA)는 2022년 상반기에 코로나19 발생 이후 경기회복 가속화에 따라 전 세계 석탄 수요가 증가하여 2021년 기준 약 79억 6,000만 톤을 기록할 것으로 추산하였다.¹⁶⁾ 이는 소비 정점을 기록했던 2014년과 유사한 수준으로, 2024년까지 전 세계의 석탄 수요는 지속적인 증가 추이를 보일 것으로 전망된다.

코로나19가 발병하기 시작한 2020년 당시 세계 석탄 수요는 두 자릿수 이상 급감할 것으로 예상되었으나, 실제로는 74억 5,600만 톤으로 전년 대비 4.4% 감소에 그쳤다. 한편, 세계 경제 성장률은 예상보다 빠르게 회복하여 2021년 기준 5.8%의 성장률을 기록하였으며, 일부 지역의 덩거나 추운 날씨로 인해 전력수요가 증가하였다. 또한 러시아-우크라이나 전쟁으로 인한 천연가스 공급이 원활하지 않아 가스 가격은 사상 최고 수준으로 인상

16) 에너지경제연구원(2022), p. 3.

되고 있으며, 신·재생에너지를 이용하여 생산되는 발전량은 예상보다 낮아 상대적으로 석탄화력발전 수요가 증가하였다.

이에 국제에너지기구는 세계 석탄 소비가 지속적으로 증가하여 2024년에는 80억 3,100만 톤에 달할 것으로 전망하였다. 이는 중국, 인도, 동남아시아 등 석탄을 주요 에너지원으로 사용하는 국가들의 전력수요 증가가 주요 원인이다. 특히 2021년부터 2024년까지 석탄화력 발전량 비중은 중국의 경우 4.1%, 인도 11%, 동남아시아 12% 수준으로 증가할 것으로 전망되는 한편, 미국과 EU는 각각 21%, 30%로 석탄화력 발전량 비중이 감소할 것으로 전망된다.¹⁷⁾

(2) 주요 이슈

가. 석탄발전 단계적 감축

세계 주요국들의 탄소중립 선언에 이어 석탄발전의 단계적 감축 합의(COP26) 등 전 세계적으로 청정에너지로의 전환 수요가 확대되고 있다. 이에 우리나라도 2050 탄소중립 및 석탄발전 Zero를 목표로 석탄발전 감축을 추진하고 있으며, 지난 2021년 10월 대통령소속 2050 탄소중립녹색성장위원회는 ‘2050 탄소중립 시나리오’를 확정 및 발표하였다.

또한 같은 시기에 영국 글래스고에서 개최된 제26차 유엔기후변화협약 당사국총회(COP26)에서 세계 각국은 석탄발전을 ‘단계적으로 감축(phase down)’하기로 합의하였다. 이에 ‘글래스고 기후합의(Glasgow climate pact)’ 대표 결정문에는 탄소저감장치가 없는 석탄발전소 단계적 감축 및 비효율적인 화석연료 보조금 단계적 폐지 등을 촉구하는 문구가 포함되었다.¹⁸⁾ 우리나라 역시 석탄발전의 단계적 감축을 발표하는 등 전 세계적인 청정에너지로의 전환 움직임에 동참 의사를 밝혔다.

17) 에너지경제연구원(2022), p. 4.

18) 뉴시스(검색일: 2022.6.1.)

나. 연료탄(석탄) 가격의 상승

석탄은 용도에 따라 크게 두 가지로 분류할 수 있다. 하나는 조강생산을 하는데 있어 철광석에서 철(Fe) 성분만 뽑아낼 수 있도록 하는 환원제로 쓰이는 원료탄(coking coal)이며, 다른 하나는 화력발전을 하는데 사용되는 연료탄(thermal coal)이다.¹⁹⁾

최근 러시아~우크라이나 분쟁으로 인한 공급망 충격과 하절기 발전수요 확대 등의 영향으로 2022년 3월 기준 연료탄 가격은 톤당 246달러로 전월 대비 약 61% 급등하는 등 공급충격이 발생하였다. 그러나 미 연준의 금리 인상 가속화 우려와 더불어 중국의 경기 둔화 및 석탄증산 정책으로 4월 연료탄 가격은 톤당 199달러로 하락하였다.²⁰⁾ 이렇듯 연료탄 가격은 국제 정세와 밀접한 관련이 있는데, 이에 따라 석탄의 생산 및 수출입 물동량이 변동되므로 지속적인 모니터링이 필요하다.

다. 석탄화력발전소 가동률 지속 감소

2019년 기준 전 세계 석탄발전소의 평균 가동률은 51%로 사상 최저치를 기록하였다.²¹⁾ 중국을 제외한 대부분의 국가들은 석탄발전소의 발전량이 감소세를 보이고 있다. 특히 미국과 유럽의 감소세가 뚜렷하며, 미국의 경우 2018년 대비 약 16%, 유럽은 약 24%의 발전량 감소를 기록하였다. 이러한 가동률 감소는 탄소중립 등의 전세계적 정책 추진에 따른 재생에너지 확대에 의해 석탄의 전력수요가 급격히 감소하고 있기 때문인 것으로 분석된다.

19) 장원익(2010), p. 346.

20) 한국광해광업공단(2022), p. 3.

21) 그린피스(검색일: 2022.6.1.)

2) 국내 석탄산업 동향 및 주요 이슈

(1) 국내 석탄산업 동향

우리나라는 석탄 매장량이 부족하여 대부분 수입에 의존한다.²²⁾ 우리나라의 석탄 소비는 대부분 유연탄으로, 1990년 초반까지만 하더라도 무연탄과 유연탄의 소비율이 비슷하였으나, 이후 무연탄 소비가 급감하면서 2020년 기준 무연탄 점유율은 6.2% 수준을 기록하고 있다.²³⁾

〈표 3-5〉 우리나라 석탄소비 비중 추이

단위: 천 톤

구분	무연탄	유연탄	합계
2009년	9,586	98,790	108,376
2010년	9,956	111,162	121,118
2011년	11,072	119,870	130,942
2012년	10,315	117,872	128,187
2013년	10,506	119,078	129,584
2014년	9,925	123,399	133,324
2015년	10,539	124,547	135,086
2016년	10,801	118,509	129,310
2017년	8,260	131,490	139,750
2018년	9,272	131,779	141,051
2019년	7,923	125,084	133,007
2020년	7,190	109,382	116,572

자료: KESIS 국가에너지통계종합정보시스템(검색일: 2022.6.5.)

유연탄의 소비는 주로 발전용, 제철용, 시멘트용, 연료용 카바이드 등에서 발생한다. 2020년 기준 국내에서는 발전용 사용 비중이 63.8%로 가장 높고, 제철용은 30.9%, 시멘트 생산 및 기타용이 5.3%이며, 2017년부터 2020년 기간동안 발전용은 다소 탈석탄 정책으로 인해 다소 감소추이를 보인 반면, 제철용 사용 비중이 증가하는 추세를 보였다.

22) 인천항만공사(2010), p. 19.

23) KESIS 국가에너지통계종합정보시스템(검색일: 2022.6.5.)

발전용 유연탄의 경우 주로 화력발전소의 원료로 활용되므로 발전소의 신설 및 폐쇄 등에 영향을 받으며, 이러한 내용이 포함된 전력수급기본계획에 따라 유연탄의 수입물동량의 추이가 변화한다. 한편 제철용의 경우 고로를 이용하여 철재를 생산²⁴⁾하는 포스코와 현대제철에서 소비가 대부분 이루어지며, 시멘트 생산용의 경우 소성공정을 통해 클링커를 제조하기 위해 사용된다.

〈표 3-6〉 부문별 유연탄 소비 추이

단위: 천 톤

구분	제철용	발전용	시멘트 생산 및 기타	합계
2017년	36,290	88,309	6,891	131,490
2018년	34,619	90,764	6,395	131,779
2019년	34,959	83,637	6,488	125,084
2020년	33,797	69,772	5,814	109,382

자료: KESIS 국가에너지통계종합정보시스템(검색일: 2022.6.5.)

(2) 주요 이슈

가. 제9차 전력수급기본계획 수립

석탄 관련 주요 이슈는 제9차 전력수급기본계획 수립과 최근 LNG로의 전환 등을 들 수 있다. 전력수급기본계획은 국내 전력수급의 기본방향 및 장기전망·전력설비 시설 계획·전력수요관리 등 우리나라의 종합적인 전력 정책으로서 2년 단위로 수립 및 시행된다.²⁵⁾ 현재는 2020년 12월 수립된 제9차 전력수급기본계획(2020~2034)을 적용하고 있다.

제9차 전력수급기본계획에 따르면, 2034년까지 석탄발전 총 60기 중 30기(15.3GW)를 폐지한다. 이 중 24기(12.7GW)는 LNG 발전으로 대체

24) 유연탄은 고로방식(Blast Furnace)의 철강제조에 있어 원료탄으로 이용, 즉 고로에 유연탄과 철광석을 함께 넣어 가열하면 유연탄이 타면서 발생하는 열에 철광석이 녹아 쇳물이 만들어짐. 보통 고로에서 쇳물 1톤을 생산하는데 유연탄 0.75톤이 사용됨

25) 에너지정책소통센터(검색일: 2022.6.5.)

한다는 내용이 포함되어 있다.²⁶⁾ 한편, 현재 건설중인 신규 석탄화력은 7기이다. 신서천화력(한국중부발전, 1009MW), 고성하이화력 1,2호기(고성그린파워, 2080MW), 강릉안인화력 1,2호기(강릉에코파워, 2080MW), 삼척화력 1,2호기(삼척블루파워, 2100MW)는 계획에 따라 준공될 예정이다. 한편, 2022년 내 제10차 전력수급기본계획이 수립될 예정이다. 정권교체에 따른 정부정책 방향이 다소 달라짐에 따라 석탄화력발전 관련 내용이 어떻게 달라지느냐에 따라 석탄의 물동량도 영향을 받을 것으로 전망된다.

〈표 3-7〉 제9차 전력수급기본계획의 주요 내용

구분	주요내용
수요관리	기준수요 대비 최대전력 12.6% 절감 추진 • '34년 최대전력 102.5GW 도출
발전설비 구성	석탄발전 60기 중 30기 폐지, 그 중 24기는 LNG로 전환 • 석탄설비: (20년)35.8GW → (30년)32.6GW → (34년)29.0GW(6.8GW ↓) • 원전설비: (20년)23.3GW → (30년)20.4GW → (34년)19.4GW(3.9GW ↓) • LNG설비: (20년)41.3GW → (30년)55.5GW → (34년)59.1GW(17.8GW ↑) • 신재생설비: (20년)20.1GW → (30년)58.0GW → (34년)77.8GW(57.7GW ↑)
온실가스 배출량 목표 달성	전원 Mix 전환, 석탄발전 운영제한 등 온실가스 감축방안 제시 • 30년 온실가스 배출량 목표(1.93억 톤) 달성
미세먼지 감축	석탄발전 축소, 재생에너지 및 LNG 발전 확대 등 추진 • 발전부문 미세먼지 배출량: (19년)2.1만 톤 → (30년)0.9만 톤(57% ↓)

자료: 에너지경제연구원(2021b), p. 24.

나. 포항제철소 1고로 폐쇄

제철방식에는 고로방식과 전기로 방식이 있는데 고로방식은 용광로에 철광석, 코크스, 석회석 등을 넣어서 쇳물을 만들어내는 반면, 전기로는 고철을 전기적인 방식으로 녹여서 철강을 제조한다. 이 중 고로는 철광석을 녹여 쇳물을 생산하는 공정시설을 의미한다. 고로를 통한 철강의 제조는 제선→

26) 산업통상자원부(2020), p. 40.

제강→압연의 과정을 거친다. 이 중 제선공정에서 유연탄이 사용되는데, 제선은 고로에 철광석과 코크스, 석회석 등을 넣고 뜨거운 열기를 가하여 쇳물을 만드는 공정이다. 유연탄은 코크스 공정을 위해 사용된다.

고로방식은 대용량 설비를 필요로 하며, 높은 건설비와 대기오염물질 배출이 많다. 따라서 업체의 진입장벽이 높아 국내에서는 포스코와 현대제철만이 고로로 철강을 생산하고 있다. 1970년 4월 착공을 시작한 포항제철소는 1973년 1고로 운영을 시작하여 거의 50여년의 시간동안 약 5,520만 톤의 쇳물을 생산해왔다. 그러나 2021년 12월 포스코 포항제철소의 1고로가 폐쇄됨에 따라 일부 유연탄 사용량의 감소는 불가피할 것으로 예상된다.

다. 국내 에너지전환 정책 추진

우리나라 정부는 2017년 10월 「에너지전환 로드맵」을 발표하였다.²⁷⁾ 탄소중립 추진 등 전 세계적인 탈석탄 정책 트렌드에 따라 우리나라도 본격적인 에너지전환 정책을 추진하기 시작하였다. 「에너지전환 로드맵」은 재생에너지 비중을 확대하는데 초점이 맞춰져 있다. 재생에너지 발전량 비중을 2030년까지 20%까지 확대하겠다는 목표를 설정하였다. 이러한 기조는 2017년 12월 발표된 「재생에너지 3020 이행계획」과 「제8차 전력수급기본계획」을 통해 구체화되어 있다.

이후 「제9차 전력수급기본계획」에서는 2018년 7월 발표된 「온실가스 감축로드맵 수정안」을 통해 강화된 전환 부문의 감축목표를 어떻게 달성할지 구체화하였다. 특히 ‘석탄은 가동 후 30년 도래 발전기 모두 폐지 및 LNG 연료 전환’²⁸⁾을 달성하는 설비용량의 계획을 제시하였다는 점에서 기존 다른 계획과 차별성을 갖는다. 또한 탈석탄 측면에서 이전에 비해 강화된 정책목표를 제시하였다고 평가할 수 있다.

27) 에너지경제연구원(2021a), p. 2.

28) 산업통상자원부(2020), p. 31.

석탄발전은 「제8차 전력수급기본계획」 및 「제9차 전력수급기본계획」을 통해 노후 발전설비 폐지, LNG 등 청정연료로의 전환, 노후석탄의 봄철 가동 중지 및 석탄발전 상한제약 등 설비용량과 발전량을 축소하는 목표가 제시되어 있다.²⁹⁾ 석탄발전 설비용량은 계획기간인 2034년 내 신규설비 추가가 계획되어, 2023년까지 증가하다가 이후 감소하는 추세를 갖는다.

〈표 3-8〉 국내 에너지전환 정책 발표 현황

시기	계획 및 주요내용
2017.10.24	「에너지전환로드맵」 • 원전의 단계적 감축: 2017년 24기 → 2038년 14기 • 재생에너지 비중 확대: 2016년 7% → 2030년 20%
2017.12.20	「재생에너지 3020 이행계획」 • 재생에너지 비중: 2016년 7% → 2022년 10.5% → 2030년 20% • 재생에너지 보급: 2016년 13.3GW → 2022년 27.5GW → 2030년 63.8GW
2017.12.29	「제8차 전력수급기본계획」 • 원전: 신규 6기 건설 백지화, 노후 10기 수명 연장 중단 • 석탄발전: 2022년까지 노후 석탄발전소 10기 폐지 • 재생에너지: 2030년까지 재생에너지 발전설비 58.5GW 확충
2019.06.04	「제3차 에너지기본계획」 • 2040년 최종에너지 소비: BAU 대비 18.6% 절감 • 최종소비 원단위(TOE/백만원): 2017년 0.113 → 2040년 0.087 • 2040년 재생에너지 발전량 비중: 30~35%
2020.07.14	「한국판 뉴딜 종합계획」(한국판 그린뉴딜) • 총사업비 73.4조원(국비 42.7조원 투자), 일자리 65.9만개 창출(~2025년) • 2025년 재생에너지 보급목표: 42.7GW
2020.12.28	「제9차 전력수급기본계획」 • 원전: 신규 6기 건설 백지화, 노후 10기 수명 연장 중단 • 석탄발전: 2020년 35.8GW → 2022년 38.3GW → 2030년 32.6GW → 2034년 29.0GW • 신재생 발전: 2020년 20.1GW → 2034년 77.8GW

자료: 에너지경제연구원(2021a), p. 7.

29) 에너지경제연구원(2021a), p. 13.

2. 기존 석탄 항만물동량 예측모형 분석

1) 분석 개요

석탄은 화물품목코드인 HS Code상 무연탄과 유연탄으로 구분되며, 중장기 항만물동량 예측시 품목 특성에 따라 무연탄과 유연탄으로, 운송방식에 따라 수입, 수출, 연안물동량으로, 수요처에 따라 제철용, 발전용, 시멘트 생산용, 기타로 구분하여 각 항만별로 물동량을 예측한 후 합산하는 Bottom-up 방식을 이용한다.

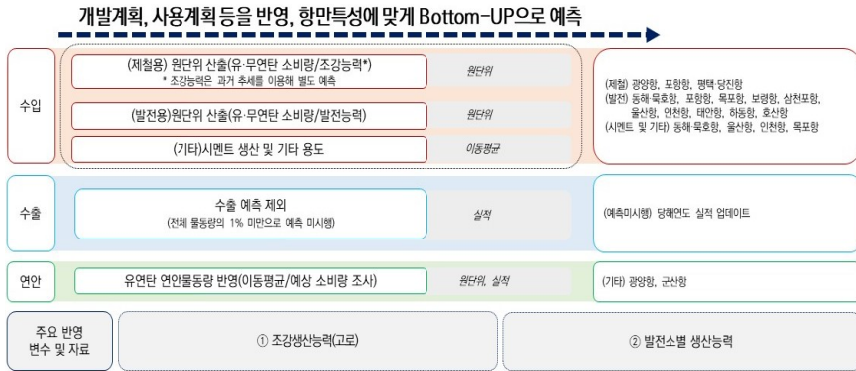
〈표 3-9〉 무연탄과 유연탄의 HS코드

HS코드	국문 품명
무연탄 (270111)	석탄과 석탄으로 제조한 연탄·조개탄과 이와 유사한 고체 연료(석탄(잘게 부수었는지에 상관없으며 응결시킨 것은 제외한다), 무연탄)
유연탄 (270112)	석탄과 석탄으로 제조한 연탄·조개탄과 이와 유사한 고체 연료(석탄(잘게 부수었는지에 상관없으며 응결시킨 것은 제외한다), 유연탄)

자료: 해양수산부, 해운항만물류정보시스템(Port-MIS) 코드집, p. 331.

수입물동량은 총 석탄 물동량 중 95.4%(2021년 기준) 비중을 차지하는 가장 중요한 예측 대상이다. 따라서 제철용과 발전용, 시멘트 생산 및 기타 용도 등 활용처별로 각각 구분하여 예측하며, 각 활용처별 예측치를 합산하여 최종 수입물동량 예측치를 도출한다. 연안물동량의 경우 유연탄의 연안운송이 이루어지고 있는 광양항, 군산항 등의 물동량의 최근 실적 및 추이를 반영하여 예측을 실시한다. 수출물동량은 석탄 총 물동량의 1% 미만으로 미미한 비중을 차지하고 있어 별도로 예측은 하지 않고 제외한다.

〈그림 3-13〉 석탄 물동량 예측 프로세스



자료: 한국해양수산개발원(2022). p. 44.

2) 예측 방법론

(1) 수입물동량

가. 주요 투입변수 선정

석탄 수입물동량은 화력발전소 증설 및 국내 조강생산능력 증가 등의 직접적인 영향을 받는다. 따라서 기존 예측 방법론은 화력발전소별 발전용량과 제철소의 조강생산능력을 수입물동량 예측을 위한 주요 변수로 활용한다. 석탄화력발전소의 발전용량은 전력수급기본계획에 의해 발표된다. 전력수급기본계획은 전력분야 최상위 국가계획으로서, 법적 근거는 전기사업법 제25조이다.³⁰⁾ 전력수급기본계획은 향후 15년 동안 우리나라 전력산업이 나아가야 할 방향을 제시하고 있다. 전력은 2차 에너지(전기 등)로서 1차 에너지(석탄 등)를 원료로 사용하여 생산될 뿐만 아니라 소비에 있어서 1차 에너지와 대체 관계에 놓여 있기 때문에 전력수급기본계획은 1차 에너지 소비에도 많은 영향을 미친다. 따라서 발전용 수입물동량을 예측하기 위해서

30) 유승훈(2021), pp. 23-27.

항만 배후에 위치한 화력발전소의 발전용량과 전력수급기본계획에 따른 화력 발전소의 증설 및 폐지를 투입변수로 선정한다.

조강이란 일반적으로 한 국가가 보유하고 있는 강철 생산능력을 말할 때 흔히 사용된다. 사전적으로 강괴라는 고체를 의미하기도 하나 통상적으로 쇳물이라고 불린다. 조강생산은 쇳물 생산량을 의미하며, 조강생산은 철강 산업에서 가장 상위공정이므로 국가별 철강생산량의 기준이 되는 용어라고 할 수 있다. 이러한 조강생산은 고로 방식과 전기로 방식으로 구분된다. 고로는 철광석을 녹여 철강을 만들기 위해 석탄(유연탄)을 활용하는 용광로로서, 석탄 물동량 예측에 활용하는 조강생산능력은 고로의 생산능력을 말한다. 고로를 사용하는 국내 대표 철강 제조사는 포스코와 현대제철이 있으며, 포스코는 광양항과 포항항, 현대제철은 평택·당진항을 주로 이용하여 석탄을 수입한다. 두 기업의 조강 생산량이 곧 우리나라의 조강 생산량으로 볼 수 있다.

조강생산능력은 실적 데이터만 수집 가능하여 예측변수로 활용하기 위해서 2050년까지의 장래 데이터를 자체 전망하고 있다. 포스코 광양은 로그 추세모형, 포스코 포항은 선형 추세모형이 예측에 적합하여 두 추세 모형으로 장래 조강생산능력을 전망한다. 한편 현대제철 평택·당진항은 확보 가능한 데이터의 부족으로 포스코 포항 및 광양 제철소의 조강생산예측치의 연평균 증가율을 적용하여 전망한다.

나. 수입물동량 구분 및 대상 항만 분류

해양수산부 해운항만물류정보시스템(Port-MIS)의 항만물동량 통계를 기반으로 우리나라 무역항 중 석탄 수입물동량이 유의미한 수준으로 발생하는 항만은 인천항, 평택·당진항, 대산항, 태안항, 보령항, 군산항, 광양항, 하동항, 삼천포항, 울산항, 포항항, 호산항, 동해·묵호항, 옥계항 등 14개 항만이다. 이 중 유연탄 수요가 발생하는 항만은 인천항, 평택·당진항, 광양항,

울산항, 포항항, 동해·목포항 등 6개 항만이다. 광양항과 포항항, 평택·당진항은 각각 포스코, 현대제철 등 제철소에서 사용하는 무연탄을 주로 처리하고 있으며, 동해·목포항은 인근에 위치한 동해화력발전소의 발전용 수요가 지속적으로 발생하고 있다. 이를 제외한 나머지는 모두 유연탄 물동량으로 포스코, 현대제철 등 제철소 수요와 화력발전소의 발전용 원료탄으로 사용된다. 항만별로 살펴보면, 포항항, 광양항, 평택·당진항은 제철용 수요, 인천항, 보령항, 태안항, 삼천포항, 호산항 등은 화력발전소의 발전용 수요라고 할 수 있다. 항만별 석탄의 주요 수요처를 정리하면 다음의 <표 3-10>과 같다.

<표 3-10> 항만별 석탄 주요 수요처

구분	무연탄				유연탄				
처리항	주요 수요처	발	제	기	주요 수요처	발	제	시	기
광양항	포스코				포스코/호남, 여수				
군산항					군장에너지 등				
대산항					당진화력, 당진에코				
동해·목포항	동해				시멘트/동해				
목포항					시멘트/발전/기타				
보령항					보령화력, 신서천화력				
삼천포항					삼천포화력				
옥계항					시멘트				
울산항	민간등 기타				대구열병합발전소/기타				
인천항	기타				영흥화력				
태안항					태안화력				
평택·당진항	현대제철				현대제철				
포항항	포스코				포스코				
하동항					하동화력				
호산항					삼척그린파워				

주: ‘발’은 발전용, ‘제’는 제철용, ‘시’는 시멘트 생산용, ‘기’는 기타 용도를 의미
 자료: 한국해양수산개발원(2022). p. 45.

다. 분류별 예측방법

수입물동량 예측은 앞서 언급한 바와 같이 석탄의 용도별 수요처를 고려하여 예측한다. 먼저 제철용 유·무연탄 예측을 위해서 주요 처리항의 실제 수입량 대비 수요처의 조강생산능력의 원단위를 산출한다. 처리항별 원단위가 산출되면 예측시점 t_0 와 과거 $t_{-1}, \dots, t_{-n}$ 간의 가중치를 달리 적용한 가중평균값에 로그 및 선형 추세 모형으로 자체 추정한 장래 조강생산능력을 적용하여 제철용 유·무연탄 예측치를 도출한다.

또한 발전용 유·무연탄은 주요 처리항의 수입물동량 실적 대비 수요처의 유·무연탄 발전소 능력의 원단위를 산출한다. 수입물동량 대비 발전능력의 원단위가 산출되면 무연탄 전력수급기본계획에 따라 증설 및 폐지되는 장래의 화력발전소의 발전능력이 인근 처리항에 미치는 영향을 파악한다. 마지막으로 제철용 유·무연탄 수입물동량 예측과 동일하게 처리항의 수입물동량 대비 발전능력의 원단위 값에 가중평균값을 구하고 전력수급기본계획에 고시된 처리항 인근의 장래 화력발전소 발전능력을 적용해 유·무연탄 수입량을 예측한다. 또한 유연탄과 기타 에너지를 혼용으로 사용하는 열병합 발전소의 경우, 원단위 편차가 심하여 전국 원단위 평균치와 업체별 예상 소비량을 조사하여 활용한다.

(2) 연안물동량

국내 석탄 연안물동량의 약 98% 이상은 유연탄이 차지하고 있다. 2021년 기준 석탄의 연안물동량은 총 569만 1천 톤을 기록하였으며, 이 중 광양항이 129만 3천 톤으로 가장 많은 비중을 차지하였다. 다음으로는 보령항 152만 톤, 군산항 101만 8천 톤, 대산항 24만 8천 톤 순이다. 광양항과 군산항에서 발생하는 연안물동량은 군산일반산업단지 내 위치한 열병합발전소인 GE4 물량이다. 발전소 가동을 위해 필요한 유연탄 물량을 광양항을 통해 수입한 후, 연안운송을 통해 군산항에서 받는 물류구조를 갖고 있다.

또한 2021년부터 보령항에서 연간 150만 톤 규모의 연안물동량이 발생하고 있는데 이는 신서천화력발전소가 운영을 시작함에 따라 발생하는 물량으로 파악되었다. 신서천화력발전소에 사용되는 유연탄은 보령항을 통해 수입되어 마량진항의 항만구역 밖 항만시설로 연안운송을 통해 받는데, 통계상 보령항의 물동량으로 집계되고 있다. 광양항과 군산항, 보령항 등과 같이 연안운송 물동량의 수요처가 명확한 물량은 유사한 수준의 물동량이 지속적으로 발생할 가능성이 높다. 다만 석탄 연안물동량은 국내 비중이 적고 물동량 발생의 변동성이 강해 원단위법을 적용하기 어렵다. 따라서 먼저 연안 물동량 관련 발전소별 예상 소비량을 조사하고 이동평균법을 적용해 중장기 예측에 활용하고 있다.

3. 기존 예측모형의 특징 및 한계

기존 석탄물동량 예측모형은 석탄을 종류와 사용 용도에 따라 구분한 후 항만별로 예측하고, 합산을 통해 석탄의 총 항만물동량을 예측하는 Bottom-up 방식을 이용한다. 그 과정에서 원단위법, 이동평균평활법 등을 이용하여 중장기 물동량을 예측하고 있다. 석탄의 경우 다른 품목에 비해 비교적 수요처가 명확하고, 시설의 증설 및 감축 등이 발생하지 않는 이상 매년 발생하는 수요의 변동이 크지 않다. 이에 단순한 모형 설계를 통해 직관적인 이해와 예측 프로세스의 효율성을 추구한 것이 기존 모형의 특징이라고 할 수 있다.

하지만 향후 석탄산업은 국내외 에너지전환 정책이 가속화됨에 따라 전력 공급이 변화하고 그에 따른 석탄화력발전소 가동률에 대한 변동이 있을 것으로 예상된다. 또한 포항제철소의 고로 폐쇄와 같은 조강생산능력의 변화가 석탄 물동량 예측의 불확실성을 높일 것으로 예상된다. <표 3-11>은 제철용

석탄 처리량의 원단위 값이 발전용 석탄 처리량의 원단위 값보다 변동이 크음을 보여준다. 원단위 방식은 단순한 모형 설계와 높은 직관성이 장점이지만 변동이 큰 데이터에는 적용하기 어렵다. 이에 향후 발생할 수요 변동 및 불확실성에 대비해 제철용 수입물동량 예측을 위한 원단위 방식 적용에 대한 타당성을 분석해볼 필요가 있다.

〈표 3-11〉 주요 석탄 수입항 별 원단위

단위: 천 R/T

원단위 평균	주요 발전용 석탄 처리항			주요 제철용 석탄 처리항		
	태안항	인천항	대산항	광양항	포항항	평택·당진항
5년	2.3	2.7	2.1	10.8	8.9	10.0
10년	2.7	2.8	2.6	10.3	8.7	9.6
표준편차	0.5	0.3	0.6	1.2	0.6	0.7

자료: 해양수산부 Port-MIS(검색일: 2022.9.12.)

이를 위해 제철용 수입물동량 예측모형 개선을 위한 대내외 변수 발굴이 우선적으로 필요하다. 과거 실적인 물동량만을 고려한 모형은 개선의 한계(원단위 혹은 추세분석)가 있다. 따라서 새로운 모형 적용을 위해서는 제철용 수입물동량과의 상관관계가 높은 설명변수를 우선적으로 발굴해야 한다.

또한 기존 예측모형은 표본 내 검정(in-sample validation)에서 비교적 안정적인 성능을 보여주고 있으나, 표본 외 검정(out-of-sample validation)에서는 상대적으로 예측력이 떨어지는 것으로 나타났다. 특히 기존에는 단일 모형만을 활용하고 있어 해당 모형의 성능을 객관적으로 평가할 수단이 존재하지 않았다. 이에 다중회귀분석, 머신러닝 기반 모형 등 다양한 통계 방법론을 적용한 모형 개발을 통해 기존 모형을 비교·검증하고 예측 결과의 신뢰성을 제고할 필요가 있다.

제3절 고철 항만물동량 예측모형 진단

1. 국내외 고철산업 동향 및 주요 이슈

고철은 철광석과 함께 철강산업의 원료로 사용되는 품목으로, 국내 고철 수요의 20% 정도가 항만을 통해 수입된다. 고철은 철재 생산에 필요한 원료이자 철재의 부산물로 생기는 재활용품이며, 철재의 생산과정, 철강 수요 산업의 가공, 철강 소재 제품의 폐기물 등에서 발생한다. 철재 생산 방식은 크게 고로 방식과 전기로 방식의 두 가지로 나눌 수 있으며, 고로 내 코크스를 태워 발생하는 열로 철광석을 녹여 철재를 생산하는 고로 방식과 달리, 전기로 방식은 전열을 통해 고철을 녹여 철재를 생산한다.

최근 고철 품목은 전 세계적으로 철강산업에 대한 탄소 규제가 강화됨에 따라 수요가 급격하게 증가하고 있다. 전기로 방식은 고로 방식과 비교해 탄소 배출량이 적으며, 최근 고로 방식에서도 탄소 배출량을 줄이기 위해 고철을 생산과정에 투입하는 추세이기 때문이다. 글로벌 수요 증가로 인해 국내외 고철 가격은 2020년 말부터 급격하게 상승했으며, 2022년에도 높아진 가격이 유지되고 있다.

국내 고철 수요 중 약 80%가 국내에서 공급되고 나머지 20%는 해외에서 수입되고 있으며, 총공급량 대비 1% 미만의 수출이 발생하고 있다. 국내 고철 총 물동량은 2021년 기준 523.2만 R/T이며, 수입 물동량은 그중 419.1만 R/T(80.0%), 수출 물동량은 25.0만 R/T(4.8%), 연안 물동량은 80.1만 R/T(15.3%)를 차지한다.

〈표 3-12〉 국내 고철 수급 현황

단위: 톤

구분	구입			수요	판매	
	국내구입	수입	자가발생	자가소비	국내판매	기타
2016년	15,507,039	5,739,133	6,096,607	27,404,373	8,279	0
2017년	17,611,463	6,403,363	6,594,497	30,664,774	15,516	0
2018년	17,586,660	6,091,442	6,254,179	29,955,572	24,500	2
2019년	16,446,772	5,864,540	6,352,109	28,600,745	9,161	0
2020년	16,098,897	3,876,144	5,891,180	25,830,911	7,953	1,098
2021년	18,608,802	4,128,060	5,68,897	28,295,633	25,994	0

자료: 한국철강협회(검색일: 2022.9.12.)

〈표 3-13〉 국내 고철 항만물동량 현황

단위: 천 R/T

구분		2017	2018	2019	2020	2021
고철	전체	6,583	6,523	6,360	4,601	5,242
	수입	5,166	5,274	5,300	3,747	4,191
	수출	563	253	219	153	250
	연안	854	997	841	702	801

자료: 해양수산부 Port-MIS(검색일: 2022.9.12.)

한편 고철을 활용하는 국내 제강사는 현대제철, 동국제강, 한국철강, 와이케이스틸, 환영철강공업, 세아베스틸, 세아창원특수강, 포스코, 대한제강 등이며, 인천항, 포항항, 평택·당진항, 군산항, 마산항, 부산항, 울산항 등 다양한 항만을 통해 고철을 수입한다.

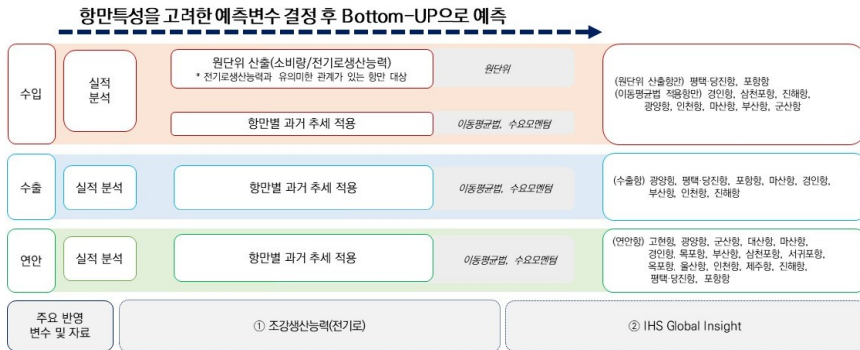
2. 기존 고철 항만물동량 예측모형 분석

1) 분석 개요

기존 고철 항만물동량 예측 프로세스는 각 항만의 특성을 고려한 예측변수를 결정해 항만별로 고철 물동량을 예측하는 상향식(Bottom-up) 접근

법을 사용한다. 수입, 수출, 연안 물동량을 각각 구분하여 예측하며, 수입 물동량의 경우 항만 인근 전기로 공장의 유무로 구분하여 예측한다. 주요 방법론으로 회귀분석, 이동평균법, 수요모멘텀 등을 적용하여 항만별 중장기 물동량을 예측한다.

〈그림 3-14〉 고철 물동량 예측 프로세스



자료: 한국해양수산개발원(2022). p. 86.

2) 예측 방법론

(1) 수입 물동량

수입 물동량 예측은 먼저 전기로를 활용하여 철재를 생산하는 대형 철강 업체가 위치한 항만과 그렇지 않은 항만을 분류한다. 또한, 인근에 전기로가 위치한 항만 중 고철 수입 물동량이 전기로 생산능력과 유의미한 관계가 있는 항만을 분류한다. 이중 유의미한 관계가 도출된 항만은 인근 전기로 공장의 생산능력을 단일 독립변수로 사용한 회귀모형과 이동평균법을 적용한 후, 두 예측치를 조합하여 최종 예측값을 산정한다. 항만 주변에 전기로가 없거나 항만의 수입 물동량이 인근 전기로 생산능력과 유의한 관계가 없는 경우는 이동평균법을 적용하여 예측값을 산정한다. 이때 이동평균법으로 도출된 예측값을 평활하기 위한 방법으로, 예측값 간 차이에 이동평균법을 다시 적용하는 수요모멘텀 방법을 함께 적용한다.

2021년 예측에는 전기로가 위치한 항만으로 광양항, 포항항, 평택·당진항, 인천항, 마산항, 군산항, 부산항 등이 조사되었으며, 이중 평택·당진항과 포항항만이 수입 물동량과 전기로 생산능력 간의 유의미한 선형 관계가 있는 것으로 파악되었다. 따라서 두 항만에는 이동평균법과 회귀분석을 각각 적용한 후 두 예측치에 8:2의 비율을 적용하여 최종 예측값을 도출하였다. 전기로 생산능력과 유의미한 관계가 없는 나머지 항만에 대해서는 이동평균법 및 수요모멘텀 방법을 적용하여 예측값을 산정하였다. 이동평균법으로 적용할 연도의 개수에 대한 선택은 항만마다 평균제곱오차(MSE)가 최저가 되는 값으로 적용하여 결정하였다.

(2) 수출 및 연안 물동량

고철 수출 물동량과 연안 물동량은 그 비중이 적거나 변동성이 큰 특징을 가진다. 이에 따라, 수출 및 연안 물동량은 별도의 설명변수를 사용하는 대신 항만별 과거 실적 추세를 이용하여 물동량을 예측한다. 5년 가중이동평균법과 수요모멘텀을 적용하여 예측하며, 가중치 선택은 MSE가 최저가 되는 가중치를 선택한다.

3. 기존 예측모형의 특징 및 한계

국내 공장의 고철 수요는 80% 이상이 국내 자급을 통해 공급되므로, 공장의 생산능력만으로 수입 물동량을 예측하기엔 어려움이 있다. 고철 수입 물동량과 전기로 공장 생산능력 단일회귀분석 결과 대부분의 항만에서 유의하지 않은 결과가 도출되었다.

전기로 공장 인근 항만 중 평택당진항과 포항항의 수입 물동량은 전기로 생산능력과 관련이 있는 것으로 도출되었으나, 인천항, 광양항, 부산항, 마산항, 군산항의 경우 수입 물동량과 전기로 생산능력 간 유의미한 관계가

없는 것으로 확인되었다. 또한, 매년 증가와 감소를 반복하는 고철 수입에 비해 전기로 생산능력은 유연하게 변할 수 없으므로 단일 예측변수로 사용하기에는 어려움이 있었다.

〈표 3-14〉 고철 수입 물동량과 인근 항만 전기로 생산능력 관계

항만	전기로 생산능력 회귀계수	전기로 생산능력 p-value
포항항	227.8	0.0002
평택당진항	213.4	0.0048
인천항	79.7	0.8189
광양항	51.7	0.5276
부산항	46.2	0.1555
군산항	47.9	0.6766
마산항	20.4	0.8506

자료: 저자 작성

또한, 2021년의 경우 철재 전방 산업의 수요 증가와 탈탄소 규제로 인해 고철 수요가 급격하게 증가했음에도 고철 자급률의 증가로 인해 고철 수입량이 증가하지 않는 패턴이 나타났다. 이에 따라 고철 수요의 변화뿐 아니라 고철 수요의 약 80%를 차지하는 자급률의 변화를 고철 수입량 예측모형에 반영해야 더 정확한 예측 결과를 얻을 수 있을 것으로 보인다.

따라서 국내 전기로 공장의 생산능력뿐 아니라 철강산업의 경기와 국내 고철 자급률 등 다양한 변수를 종합적으로 고려할 수 있는 방법론 개발이 요구된다. 고철 물동량은 철재 수요에 따라 발생하는 파생수요로써 철재가 사용되는 전방 산업 경기의 영향을 받는다. 고철 물동량은 철재 생산에 있어 철광석과 경쟁 구도에 있고, 고철 가격뿐 아니라 철광석 가격의 하락에 따라 수요가 증가한다. 고철 물동량은 국내 자급률에 영향을 받으며 고철은 생산되기보다는 폐기 철에 의해 발생하므로 발생량을 예측하는 것이 필요하다. 즉, 고철 수입 물동량은 국내 철 발생량과 음의 관계, 전방 산업 경기와 양의 관계이다. 제4차 전국무역항 기본계획 및 수정계획 지원에 대비하여 고철 항만물동량 예측모형 개선 및 추가적인 변수 확보가 필요하다.

제4절 시멘트 항만물동량 예측모형 진단

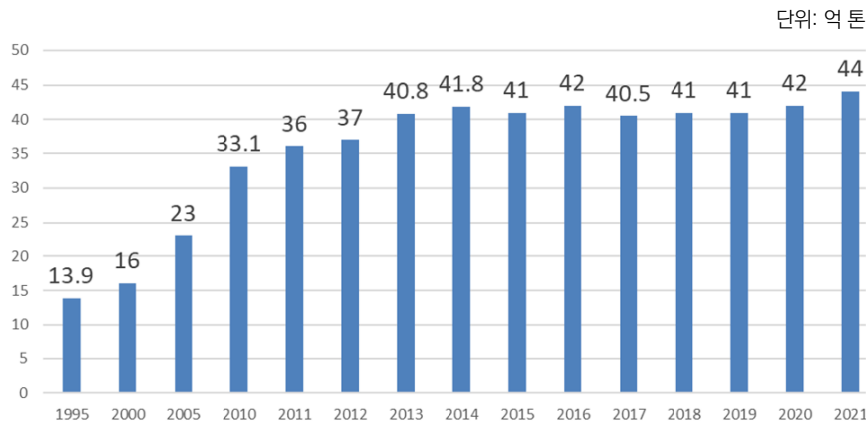
1. 국내외 시멘트산업 동향 및 주요 이슈

1) 국외 시멘트산업

(1) 세계 시멘트 수급 동향

2021년 세계 시멘트 생산량은 44억 톤이며 2020년(42억 톤) 대비 4.8% 증가했으며 2010년 이후 연평균 2.6% 증가했다. 시멘트 최대 생산국은 중국이며 2021년 기준 연간 25억 톤(전 세계 생산량의 56.8%)을 생산했다. 한편 우리나라는 세계 13위의 시멘트 생산국이며 2021년 기준 연간 5,044만 톤(1.1%)을 생산했다.

〈그림 3-15〉 세계 시멘트 생산량



자료: US Geological Survey(2022), p. 4, 저자 재작성

〈표 3-15〉 2021년 국가별 시멘트 생산량 및 비중

단위: 백만 톤, %

순위	국가	생산량	비중
1	중국	2,500	56.8
2	인도	330	7.5
3	베트남	100	2.3
4	미국	92	2.1
5	터키	76	1.7
6	인도네시아	66	1.5
7	브라질	65	1.5
8	이란	62	1.4
9	러시아	56	1.3
10	사우디아라비아	55	1.3
11	일본	52	1.2
12	멕시코	50	1.1
13	대한민국	50	1.1
14	이집트	40	0.9
-	그 외	808	18.4

자료: US Geological Survey(2022), p. 5; 한국시멘트협회(2021a), pp. 185-188, 저자 재작성

2021년 세계 시멘트 소비량은 44억 2,000만 톤이며 2020년 대비 6% 증가했다. 세계 시멘트 소비량은 2018년 이후 연평균 2.9% 증가했다. 시멘트 최대 소비국인 중국은 연간 24억 6,300만 톤을 소비하며 이는 전 세계 소비량의 55.7% 비중을 차지한다. 우리나라는 전 세계 11위 소비 국가이며, 연평균 4,936만 톤(1.1%)을 소비한다.

코로나19 사태가 안정화되면서 시멘트 수급 상황은 나아지고 있지만 글로벌 공급망 약화와 인플레이션과 같은 새로운 문제가 발생하고 있다. 2021년 3월 수에즈 운하에 에버그린호가 좌초되면서 상반기 시멘트와 클링커 수출입 실적에 부정적인 요소로 작용할 수 있었지만, 다행히 그 영향은 크지 않은 것으로 분석된다. 2021년 하반기에는 유연탄 가격 상승과 건설경기 회복과 같은 요인이 세계 시멘트 수급 변화에 큰 영향을 미친 것으로 판단된다.

〈표 3-16〉 2021년 국가별 시멘트 소비량 및 비중

단위: 백만 톤, %

순위	국가	소비량	비중
1	중국	2,463	55.7
2	인도	315	7.1
3	미국	110	2.5
4	인도네시아	65	1.5
5	브라질	65	1.5
6	베트남	65	1.5
7	이란	65	1.5
8	터키	61	1.4
9	사우디아라비아	57	1.3
10	러시아	56	1.3
11	대한민국	49	1.1
12	이집트	49	1.1
-	그 외	999	22.6

자료: 한국시멘트협회(2021a) pp. 185-188, 저자 재작성

(2) 주요 이슈

가. 탄소중립

2016년 전 세계는 지구온난화를 방지하기 위해 파리기후변화협정을 체결하고 온실가스 감축 및 탄소 배출량 감소를 위한 국가별 기본 대응 계획을 수립했다. 그에 따라 온실가스 다배출 업종인 시멘트 산업도 탄소중립을 목표로 많은 노력을 기울이고 있다.

2021년 미국시멘트협회는 탄소중립 로드맵을 개발했다. 본 로드맵에는 협회의 회원사들과 함께 시멘트 가치사슬 전반에 걸친 탄소중립 목표를 달성하기 위한 계획이 포함되어 있다. 2021년 일본의 스미모토 오사카 시멘트는 2050년까지 일본 정부의 목표에 맞춰 탄소중립을 달성하기 위한 중장기 목표와 전략을 수립했다. 탄소배출 저감 기술개발 및 도입, 무탄소 전력

공급, 저탄소 시멘트 및 콘크리트 제품 개발, 친환경 결합재료 및 탄소포집 활용 및 저장기술 개발 등을 도입할 계획을 밝혔다. 인도의 라파즈홀심은 인도의 탄소중립 목표를 가속화하기 위해 1,226억 원을 투자해 폐열 회수 시스템을 확대했다. 멕시코를 기반으로 하는 시멕스는 제로탄소 콘크리트를 개발하여 품질 저하 없이 탄소배출량을 기존 대비 70% 이상 줄였다.³¹⁾

나. 유연탄 가격 상승

2021년 하반기에 세계 시멘트 산업에 가장 큰 영향을 미친 요인은 원자재 위기이다. 미국, 유럽 및 아시아의 전력 생산업체의 유연탄 사용량이 많아지면서 전기 및 유연탄 가격이 모두 상승했다.

중국은 환경 및 안전상의 이유로 탄광 폐쇄, 반부패 조치 및 국가 행사를 위한 광산 폐쇄로 유연탄 생산량을 줄였다. 또한 정치적인 이유로 호주 석탄 수입 금지, 인도네시아 홍수, 몽골의 코로나19 재유행으로 유연탄 수급에 문제가 있었다. 유연탄 가격 상승으로 인해 2021년 9월의 중국 시멘트 생산량은 2억 3,330만 톤으로 2020년 대비 12% 감소했다.

2) 국내 시멘트산업

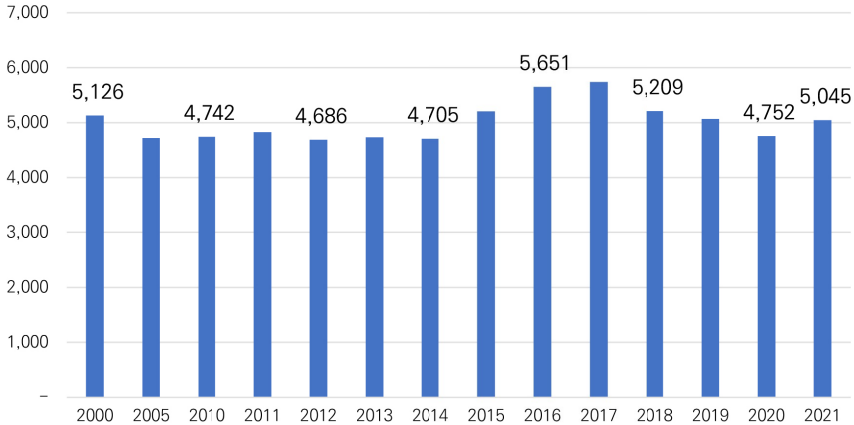
(1) 국내 시멘트 수급 동향

2021년 국내 시멘트 총생산량은 5,045만 톤이며 2020년(4,752만 톤) 대비 6.2% 증가했다. 국내 시멘트 생산량은 2010년 이후 연평균 0.56% 증가했다. 국내 시멘트 생산량 및 재고량 대부분은 내수로 소비되며 수출입 물량은 매우 적다. 2021년 국내 소비량은 4,936만 톤이며 이는 2021년 생산량의 97.8%를 차지한다.

31) 한국시멘트협회(2021b), pp. 31-34.

〈그림 3-16〉 국내 시멘트 생산량

단위: 만 톤



자료: 한국시멘트협회(2021a), pp. 10-11, 저자 재작성

〈표 3-17〉 국내 시멘트 수출입 및 내수량

단위: 만 톤, %

연도	2000	2005	2010	2015	2020	2021
내수	4,800	4,629	4,549	5,074	4,716	4,936
수출	488	597	752	504	528	461
수입	68	340	77	116	51	55

자료: 한국시멘트협회(2021a) pp. 10-11, 저자 재작성

2021년 기준 내수량의 43.9%(2,165만 톤)가 수도권에서 소비되며 경북(11.0%), 전남(8.7%), 충남(8.1%), 경남(7.9%) 순으로 내수량이 높다. 수요 형태별로는 민간 수요가 99.6%로 4,916만 톤이며 정부(공공) 수요는 0.4%로 20만 톤이다. 수출은 79.4%(357만 톤)가 시멘트의 원자재인 클링커이며, 완제품인 시멘트는 20.6%(104만 톤)를 차지한다. 수입의 경우는 100%(55만 톤) 시멘트이다. 한편 국내 시멘트 주요 제조업체는 총 9개(삼표, 쌍용, 한일·현대, 아세아, 성신, 한라 등)로 2021년 기준 9개 기업의 연간 총생산능력은 6만 147만 톤이며 평균 가동률은 70.6%이다.

〈표 3-18〉 국내 주요 시멘트 제조사의 생산능력 및 가동률

단위: 만 톤, %

구분	삼표	쌍용	한일	한일·현대	아세아	성신	한라	고려	합계
생산능력	1,004	1,475	713	686	414	968	818	660	6,147
가동률	69.1	81.7	66.9	66.3	76.4	60.0	66.9	96.1	70.6

자료: 한국시멘트협회(2021a), pp. 22-23, 저자 작성

(2) 주요 이슈

가. 탄소중립

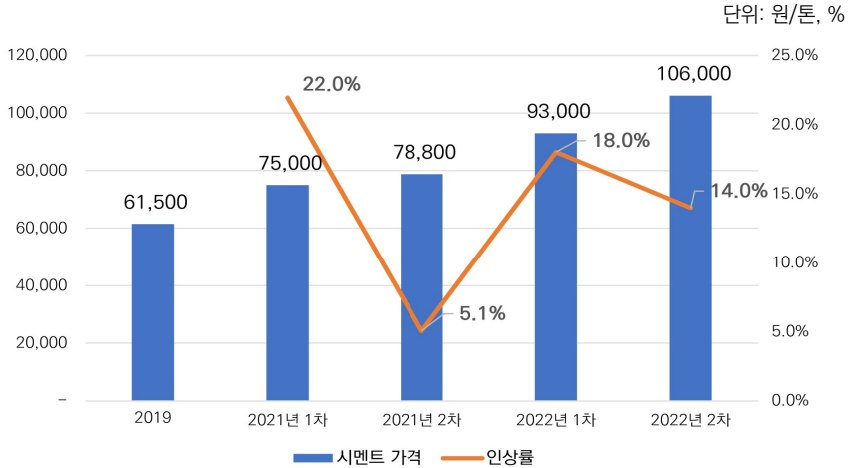
2020년 정부의 ‘2050 탄소중립 비전’ 선언에 동참하기 위해 국내 시멘트 업계는 ‘시멘트그린뉴딜위원회’를 발족하고 탄소중립을 위한 혁신기술 개발 로드맵을 제시했다. 주요 7개 시멘트 제조사(삼표, 한일현대, 아세아, 성신양회 등)와 산·학·연 전문가들은 ‘2050 탄소중립 도전 공동선언문’을 통해 시멘트 산업의 탄소중립 신산업 모델을 제안했다. 선언문에는 순환자원 및 대체연료 사용 확대, 저탄소 원료 활용 및 공정 에너지 효율 향상을 통한 온실가스 감축, 저탄소 시멘트 생산, 탄소 포집 및 전환 기술 개발 등 탄소순환형 시멘트 생산공정 구현에 대한 계획이 포함되었다.³²⁾

나. 지속적인 시멘트 가격 인상

국내 시멘트 제조사는 세계적인 유연탄 가격, 유가 및 물류비 상승, 전기료 인상 등으로 연이어 가격을 인상했다. 2019년 톤당 6만 1,500원이었던 시멘트 가격은 2021년에 7년 만에 22.0% 인상되었으며, 이후 지속적으로 인상되어 2022년 톤당 10만 6,000원을 기록했다. 정부는 유연탄 가격 상승 및 유가 상승으로 인해 시멘트 수급이 어려워지자 시멘트 제조사에 시멘트 생산량을 증가시키고 내수 비율을 높일 것을 권고했다. 하지만 2021년부터 지속된 시멘트 가격 인상은 레미콘 업계의 파업 등과 같이 시멘트 수급의 불확실성을 높이는 변수로 작용하고 있다.

32) 한국시멘트협회(2021b), pp. 6-7.

〈그림 3-17〉 국내 시멘트 가격 인상 추이



자료: 저자 작성

2. 기존 시멘트 항만물동량 예측모형 분석

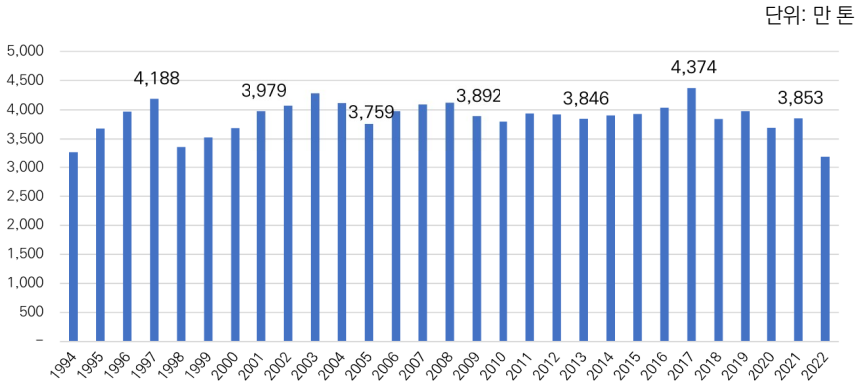
1) 분석 개요

2021년의 시멘트 총 물동량은 3,853만 톤이며, 2001년 이후 연평균 0.2% 감소하였다. 가장 큰 비중을 차지하는 연안 물동량은 3,355만 톤으로 전체 물동량 중 87.1%를 차지한다. 같은 기간 수출입 비중은 각각 12.0% (463만 톤)와 0.9%(34만 톤)이다. 장치산업의 특성상 항만 배후 단지에 있는 시멘트 제조사의 전용부두를 통한 국내 운송이 물동량 대부분을 차지한다.

시멘트 제조사는 삼척, 동해, 영월, 단양 및 제천 등 강원도와 충청도에 분포하고 있다. 연안 물동량은 동해·묵호항, 옥계항과 삼척항에 인접해 있는 쌍용C&E, 한라시멘트 및 삼표시멘트에서 생산된 제품이 대규모 해상운송을 통해 각 지역 항만에 공급되는 구조이다. 시멘트 수요가 발생하는 주요 항은 인천, 부산, 광양, 마산, 울산 및 군산이다. 수출 물량은 2021년

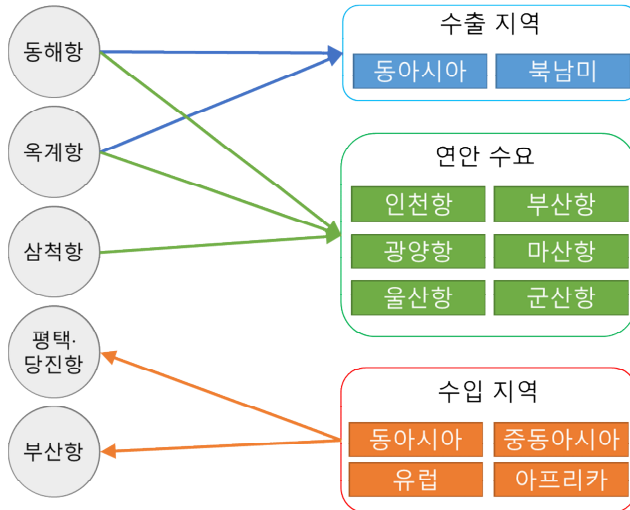
기준 79.4%가 클링커이며 동해·묵호항과 옥계항을 통해 동아시아, 북남미 등으로 수출된다. 수입은 전량 시멘트이며 일본을 포함한 동아시아, 중동, 유럽 및 아프리카에서 평택·당진항과 부산항을 통해 수입된다.

〈그림 3-18〉 연간 시멘트 총 물동량



자료: 해양수산부 Port-MIS(검색일: 2022.9.12.), 저자 재작성

〈그림 3-19〉 2021년 기준 시멘트 물동량 발생 구조



자료: 저자 작성

기존 모형의 예측 프로세스는 <그림 3-20>에 나타나 있다. 기존 모형은 국내 시멘트 산업의 수급분석 후 전국 시멘트 수요를 추정하여 항만별로 물동량을 분배하는 하향식(Top-down)을 사용한다. 본 모형은 수입, 수출 및 연안 물동량 예측으로 구분되며 사용되는 변수는 한국은행의 건설수주액, IHS Markit의 GDP 및 국내 인구 전망치, 한국시멘트협회의 시멘트 생산, 내수, 수입 및 수출량이 있다.

예측모형에서 사용되는 방법론은 보간법, 로그 변환, 선형 및 다중회귀 분석, 원단위 분석 및 이동평균법이다. 본 절에서는 기존 모형을 수입, 수출 및 연안 물동량 예측으로 구분하여 각 구분별 물동량 예측 프로세스에서 사용하는 변수 및 방법론을 설명한다.



자료: 한국해양수산개발원(2022). p. 25.

2) 예측 방법론

(1) 수입물동량

수입물동량 예측은 지속적인 물동량이 발생하는 평택·당진항과 부산항을 대상으로 한다. 1996년 이후 꾸준한 수입 물동량이 발생한 인천항과 광양항은 2019년부터 수입물동량이 발생하지 않고 있으며 목포항과 동해·

목호항은 간헐적으로 수입물동량이 발생하여 수입 물동량 예측 대상에 포함하지 않는다.

수입물동량 예측은 과거 수입 물동량 실적을 입력변수로 사용한다. 수입 물동량이 발생하는 항만별 수입 실적의 추세를 반영하기 위해 단순이동평균법(SMA, Simple Moving Average)과 연평균 증감률(CAGR, Compounded Annual Growth Rate) 보정을 사용한다.

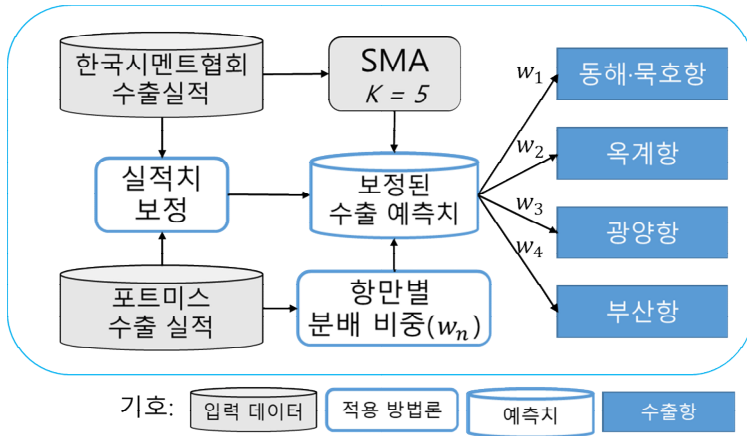
$$Y_{t+1} = Y_t^{SMA} \times (1 + R_{t_s, t_e}) \quad (\text{III-1})$$

여기서 Y_{t+1} 은 $t+1$ 시점의 수입물동량, $Y_t^{SMA} = \frac{1}{n} \sum_{i=t-n}^t Y_i$ 는 t 시점에 서 단순이동평균법으로 산출한 수입물동량, n 은 단순이동평균의 기간이며, $R_{t_s, t_e} = \left(\frac{Y_{t_s}}{Y_{t_e}}\right)^{\frac{1}{t_s - t_e}} - 1$ 은 시점 t_s 부터 t_e 기간 동안 수입물동량의 연평균 증감률이다.

(2) 수출물동량

수출물동량의 예측 대상은 동해·목호항, 옥계항, 광양항과 부산항이다. 시멘트 수출은 클링커와 시멘트로 구분되는데 한국시멘트협회에서 집계하는 클링커 및 시멘트의 수출 실적과 Port-MIS 수출 실적을 입력변수로 사용한다. 먼저 한국시멘트협회의 수출 실적과 Port-MIS의 수출 실적의 보정 값을 도출한다. 한국시멘트협회의 수출 실적에 단순이동평균법($k=5$)으로 예측치를 도출한다. 이후 사전에 도출된 보정 값을 적용하여 보정된 수출 예측치를 정의한다. Port-MIS 수출 실적을 기반으로 항만별 수출 비중을 적용하여 항만별 수출물동량을 분배한다. 예측을 위한 항만별 수출 비중은 단순이동평균법($k=3$)을 사용한다.

〈그림 3-21〉 수출 물동량 예측 방법



자료: 저자 작성

(3) 연안물동량

연안물동량 예측은 시멘트 생산 거점인 동해·목호, 삼척, 옥계, 광양항과 시멘트 수요 지역인 인천, 부산, 광양, 마산, 울산항 등 총 15개 항만을 대상으로 한다. 연안물동량 예측은 전국 시멘트 수요 산정, 연안 입출항 분배로 구분할 수 있는데, 전국 시멘트 수요 산정을 위한 입력 변수는 한국의 행의 건설수주액, IHS Markit의 GDP 및 국내 인구 전망치, 한국시멘트협회의 시멘트 내수와 해상 비율 데이터를 사용한다. 연안 입출항 분배를 위한 입력 변수는 한국시멘트협회의 내수 데이터와 Port-MIS의 연안 실적 데이터를 사용한다.

가. 전국 시멘트 수요 산정

전국 시멘트 수요 산정은 회귀분석 기반 수요 추정(이하 회귀 모형), 원단위 기반 수요 추정(이하 원단위 모형), 단순이동평균법 기반 수요 추정(이하 이동평균 모형) 결과의 평균값으로 한다. 먼저 회귀분석 모형은 건설

수주액 추정을 위해서 IHS Markit의 실질GDP(PPP조합)를 활용한 단순선형회귀분석을 실시한다. 분석 결과, F 검정값이 456.869($p < .001$)로 본 모형이 적합하다고 할 수 있으며 $\text{adj.}R^2 = 0.9214$ 로 높은 설명력을 나타냈다. 본 모형을 바탕으로 IHS Markit에서 제공하는 GDP 예측치를 기반으로 건설수주액을 예측한다.

〈표 3-19〉 건설수주액 추정 회귀분석 모형 요약

변수	계수	표준오차	t 검정	F 검정	$\text{adj.}R^2$
상수	-190,962	44,574.186	-4.284***	456.869***	0.9214
GDP	754.913	35.318	21.374***		

주: *는 $p\text{-value} < .05$, **는 $p\text{-value} < .01$, ***는 $p\text{-value} < .001$ 을 의미함
자료: 저자 작성

회귀분석 모형은 전국 시멘트 수요 산정을 위해 전 단계에서 추정된 건설수주액과 IHS Markit에서 제공하는 국내 인구 전망치가 시멘트 수요에 영향을 미치는지 파악하기 위해 다중회귀분석을 실시한다. 데이터 스케일링을 위해 독립변수를 로그로 변환한다. 분석 결과에 따르면, F 검정값이 4.725($p < .05$)로 모형이 적합하다고 판단되나 $\text{adj.}R^2 = 0.271$ 로 다소 낮은 설명력을 보였다. 마지막으로 회귀분석 모형은 수요 예측값을 수요 실적을 기준으로 재보정하여 최종 수요 추정 결과를 도출한다.

〈표 3-20〉 시멘트 수요 추정 회귀분석 모형 요약

변수	계수	표준오차	t 검정	F 검정	$\text{adj.}R^2$
상수	685184109.7	217518141.2	3.150**	4.725*	0.271
총인구수	-211993550.4	69943437.01	-3.031**		
건설수주액	14039574.38	4875418.842	2.880**		

주: *는 $p\text{-value} < .05$, **는 $p\text{-value} < .01$, ***는 $p\text{-value} < .001$ 을 의미함
자료: 저자 작성

원단위 모형은 인구 대비 시멘트 소비량을 도출하여 IHS Markit의 국내 인구 전망치 대비 시멘트 소비량을 예측하는 방법이다. 인구 대비 시멘트 소비량의 예측은 단순이동평균법($k = 5$)을 사용한다. 이동평균 모형은 과거 시멘트 소비의 추세를 반영하기 위해 과거 물동량에 $k = 5$ 인 단순이동평균법을 적용하여 예측한다. 최종 시멘트 수요 산정값은 회귀분석 모형, 원단위 모형, 이동평균 모형의 결과를 평균하여 사용한다.

나. 연안 입출항 분배

시멘트 수요 산정 결과는 전국의 시멘트 수요를 의미한다. 전국의 시멘트 수요 산정 값에서 연안 물동량을 추정하기 위해서는 내수 출하량 중 해송으로 운반되는 시멘트 물동량을 추측해야 한다.

첫 번째로 시멘트 수요 산정 결과를 바탕으로 내수 출하량을 추정한다. 시멘트 내수 출하량을 추정하기 위해서는 수요 산정 결과에서 예측 수입량을 제외한다. 본 모형에서 예측 수입량을 도출하기 위해서 단순이동평균법($k = 3$)을 적용한다. 시멘트 물동량은 크게 해송, 육송, 철도 운송으로 구분된다. 두 번째로 연안 물동량을 추정하기 위해서는 내수 추정값에서 해송으로 운반되는 물동량만을 고려해야 한다. 따라서 포장 형태로 운반되는 시멘트 물동량을 제외한 벌크 물동량을 도출하고 과거 해송 실적 비중을 고려한 연안 물동량을 추정한다.

포장 출하량은 자기회귀(AR) 모형과 보정비를 적용하여 도출하였다. 포장 출하량의 추정 식을 아래와 같다. t_0 는 기준 연도이며 본 모형에서는 2000년을 기준으로 정의했다. t_{n+1} 은 예측연도이다. 내수 연안 비중은 내수 출하 대비 연안 물동량 발생 기준을 의미하며 본 모형에서는 최근 1년의 내수 연안 비중이 변화가 없다는 것을 가정하여 예측에 사용하였다.

$$\text{포장 출하량} = -2.842 \log(t_0 - t_{n+1}) + 13.044 \quad (\text{III-2})$$

세 번째, 연안 출하량이 정의되면 시멘트 생산 거점 항인 동해·묵호, 삼척, 옥계, 광양항에 최근 1년간 출하 비중을 계산하여 출하량을 분배한다. 본 방법에서도 미래에 출하 비중에 변화가 없다는 것을 가정한다. 연안 물동량은 출하량과 입하량이 같다. 따라서 수요 항인 인천, 부산, 마산 및 울산 등 12개 항만에 입하 비중에 따라 분배한다. 분배방식은 출하와 동일하다.

3. 기존 예측모형의 특징 및 한계

국내외 시멘트 산업의 주요 이슈인 탄소중립과 원자재가격 상승으로 인한 생산량 감소, 내수 우선 정책 및 시멘트 가격 상승은 단·중·장기 시멘트 물동량에 큰 영향을 줄 것으로 예상된다. 먼저 탄소중립을 위한 국내 및 글로벌 정책에 따라 친환경 재료 사용, 지속가능성 제조 및 탄소포집 기술 적용 등은 국내외 중장기 시멘트 수요에 변화를 줄 것으로 예상된다. 따라서 국내 및 글로벌 경기 상황을 예측모형에 반영하기 위한 거시경제변수 발굴이 필요하다. 또한 원자재 및 시멘트 가격 상승은 단기 시멘트 수요에 영향을 줄 것으로 예상된다. 따라서 단기 시멘트 수요예측을 위한 새로운 접근이 필요하다. 마지막으로 시멘트 생산량이 감소하게 되면 국가는 국내 건설경기를 유지하기 위해 내수 우선 정책을 적용한다. 이러한 시멘트 수급 구조 변화는 항만물동량 분배에 대한 개선방안을 요구한다.

기존 시멘트 물동량 예측 프로세스는 변수 선정, 전국 시멘트 수요량 기반 물동량 추정, 항만별 물동량 분배로 나눌 수 있다. 먼저 변수 선정 단계에서는 한국은행의 건설수주액, IHS Markit의 GDP 및 국내 인구 전망치, 한국시멘트협회의 시멘트 생산, 내수, 수입 및 수출량 데이터를 정의한다. 시멘트 생산, 내수, 수집 및 수출량에 영향을 미칠 수 있는 설명변수로 건설수주액, GDP 및 국내 인구 전망치를 사용하고 있으나, 설명변수 간의 상관관계 및 종속변수와의 인과관계에 대한 검증은 거치지 않았다. 따라서

시멘트 수요 기반 전국 물동량 추정을 위해서 독립변수 검증을 통한 변수 발굴 및 분석이 필요하다.

두 번째 전국 시멘트 수요량 기반 물동량 추정 단계에서의 한계점은 연안 출하 비중 선정이다. 시멘트 물동량은 전국 수요를 충족시키기 위해 주요 생산 거점 항만에서 생산된 시멘트를 출하하는 방식이다. 따라서 주요 생산 거점 항만의 출하량을 정확하게 예측하는 것이 무엇보다 중요하다. 그리고 주요 생산 거점 항만의 예측 시점 t_x 의 출하 물동량 예측치는 전년 도(t_{x-1})의 해당 항만의 연안 출하 비율을 통해 추정한다. t_{x-1} 시점의 물동량은 SMA, GART 및 원단위법을 사용하는 모형 특성상 $t_x, t_{x+1} \dots t_n$ 에 가장 큰 영향을 미친다. 따라서 t_{x-1} 의 물동량이 일시적으로 과대·과소 산정되었을 경우 장기 물동량 예측치에도 많은 변화를 준다. 따라서 t_x 와 t_{x+1} 과 같은 단기 물동량 예측을 위해서 t_{x-1} 의 비중을 활용한 연안 출하 비중 선정이 아닌, 계절성 및 추세를 포함할 수 있는 정밀한 단기 물동량 예측 방법론 적용이 필요하다.

세 번째 항만별 물동량 분배 단계에서의 한계점은 항만별 연안 입하 가중치 설정이다. 시멘트 물동량 예측모형의 예측 시점 t_x 의 분배 가중치는 t_{x-1} 의 전국 물동량 대비 항만별 연안 입하 물동량 비중을 사용한다. $t_x, t_{x+1} \dots t_n$ 의 분배 가중치는 t_{x-1} 과 같다고 가정한다. 이는 출하 물동량 비중을 선정하는 문제점과 같다. 이를 해결하기 위해서는 연안 입하 비중을 추정하기 위해 지역별 거시경제 지표 등을 활용한 지역별 수요를 기반으로 입하 물동량 분배 규칙을 재정의할 필요가 있다.

제5절 소결 및 시사점

기존 비컨테이너 항만물동량 예측모형은 품목별로 특성을 반영하여 예측 프로세스를 설계하고 효율을 극대화하는 방향으로 개발되었다. 품목과 연관된 산업과 물류 흐름을 직관적으로 설명할 수 있도록 핵심 변수를 선정하고 가능한 단순화된 모형을 사용하였다. 기존 예측모형의 표본 내 검정 결과는 모형의 간결성에 비해 높은 설명력을 보여줬으나, 표본 외 검정 시 예측 성능이 확연히 떨어지는 결과를 보였다. 특히 최근 급변하는 국내외 환경변화는 품목별 연관 산업의 구조를 변화시키고 있으며, 이는 항만물동량에도 영향을 미칠 것으로 예상된다. 이에 따라 각 품목별로 항만물동량 예측모형을 진단한 결과를 반영하여 개선 방향을 도출할 필요가 있다. 품목별 항만물동량 예측모형 진단결과와 개선 방향을 요약하면 다음과 같다.

자동차 항만물동량은 국내 자동차 수요에 기반하여 추정하고 있으며, 단순선형회귀 모형을 활용하고 있다. 이에 최근의 대내외 환경변화를 반영하기 어려우며 표본 외 검정 시 예측 정확도가 떨어진다. 특히 수입물동량의 경우 증가하는 수입차 수요를 반영하지 못하고 있다. 이에 다양한 설명변수의 활용을 검토하고 신규 예측모형을 개발함으로써 예측 정확도 제고가 필요하다.

석탄과 고철 항만물동량도 각각 조강생산능력과 전기로 생산능력에 의존하여 원단위법 또는 단순선형회귀 모형을 활용하고 있으나, 추가 설명변수 발굴 및 예측모형 개선이 시급한 것으로 나타났다. 이에 석탄의 경우 상관분석을 통해 설명변수를 추가 발굴하고 다중선형회귀 모형으로 예측 정확도 개선을 시도할 필요가 있으며, 고철은 국내 고철 자급률을 고려한 고철 수급분석을 실시함으로써 보다 정교한 모형 개발이 가능할 것으로 판단된다.

또한 기존 시멘트 항만물동량 예측모형은 주요 시멘트 생산 거점 항만의 연안 출하 비중 산정과 항만별 연안 입하 가중치 산정 방식에서 개선이 필요한 것으로 진단되었다. 이에 우선적으로 시멘트 수요 기반 전국 물동량 추정을 위해 추가 설명변수 발굴이 필요하며, 계절성 및 추세를 반영한 단기 물동량 예측 방법론을 적용하여 연안 출하 비중 산정 시 활용할 필요가 있다. 이밖에 지역 수요 기반 입하 물동량 분배 비중 산정 방식 개선을 통해 시멘트 연안물동량 예측의 정확도 제고가 기대된다.

한편 각 개선 방향은 4장에서 실증분석을 통해 개선이 실질적으로 어느 정도의 효과를 가지는지 계량적으로 측정한다. 특히 교차검증법을 도입하여 표본 외 검정 시 모형의 성능을 제시하고 기존 모형 대비 신규 모형의 우수성을 증명하고자 한다.

〈표 3-21〉 품목별 예측모형 진단 결과 및 개선 방향

품목	진단 결과	개선 방향
자동차	• 자동차 수요예측 시 단순선형회귀 모형 사용	• 대내외 환경변화를 반영하는 모형 개발을 통해 예측 정확도 제고
	• 수입물동량 예측 시 최근의 추세 반영 미흡	• 시계열 모형과 비교검토를 통해 최근 추세를 반영한 예측모형 개발
석탄	• 투입변수의 다양성 부족	• 상관분석을 통한 투입변수 발굴
	• 원단위법 및 단순 추세분석에 의존	• 다양한 분석방법론 도입으로 비교·검증 및 예측 정확도 개선
고철	• 기존 고철 수요예측 모형의 예측 정확도 저하	• 추가 설명변수 발굴 및 예측모형 개발
	• 국내 고철 자급률 반영 필요	• 국내 고철 수급분석을 통한 수입 물동량 예측 정확도 개선
시멘트	• 경험에 의한 변수 선정	• 상관분석을 통한 변수 선정
	• 단기 예측 결과에 따른 장기 예측치 변화가 큼	• 머신러닝을 적용한 단기 물동량 예측
	• 입하 물동량 분배 시 전년도 비중 사용	• 지역 수요 기반 입하 물동량 분배

자료: 저자 작성

04

비컨테이너 항만물동량 예측모형 고도화 방안

제4장에서는 앞서 3장에서 진단한 품목별 예측모형의 한계점을 보완하여 모형의 예측력 및 활용도를 제고할 수 있는 고도화 방안을 제시한다. 각 품목의 특성 및 가용한 데이터를 고려하여 적절한 예측모형을 선별하고, 실증분석을 통해 예측치를 전망한다. 또한 교차검증법을 활용하여 모형 간 예측력을 계량적으로 비교·검토하고, 최종적으로 모형의 적합도, 활용도, 효율성 등을 기반으로 모형의 고도화 방안을 제시한다.

제1절 자동차 항만물동량 예측모형 고도화

기존 자동차 항만물동량 예측모형은 국내 자동차산업의 수급분석을 기반으로 항만별 수출입 및 환적 물동량을 Top-down 방식으로 예측한다. 이에 국내 자동차 수요에 대한 예측모형 개선이 자동차 항만물동량 예측모형 고도화의 핵심이라고 볼 수 있다. 본 절에서는 (1) 국내 자동차 수요 예측모형 고도화와 (2) 수입자동차 수요 예측모형 고도화를 통해 자동차 항만물동량 예측모형의 예측 정확도를 개선하고자 한다.

첫 번째, 국내 자동차 수요예측 모형 고도화는 앞서 2장에서 소개한 머신러닝 기반 모형 중 하나인 ANFIS 모형을 적용하여 국내 자동차 수요예측의 정확도 제고가 목적이다. 기존 모형에서는 우리나라의 연도별 인구수를 단일 독립변수로 활용하는 단순선형회귀(Simple Linear Regression) 모형을 통해 수요를 추정하였으며, 이에 모형 추정 이후 시점에 대한 전망치의 전반적인 정확도가 낮았다. ANFIS 모형을 활용하여 인구 외 다양한 경제 변수들을 모형에 투입하여 학습시킨다면 전망치의 예측 정확도를 향상시킬 수 있을 것으로 기대된다.

두 번째는 해외브랜드 수입자동차 수요 예측모형 고도화이다. 기존 모형은 우리나라 GDP를 단일 독립변수로 하는 SLR 모형을 활용하고 있으나, 2000년대 이후 수입차 수요가 급격히 증가하며 기존 모형의 예측 정확도가 낮아지고 있다. 이에 국내 수입차 수요를 국내 연도별 GDP 및 인구수 변수와의 상관관계를 고려한 VAR 모형으로 대체하여 예측 정확도를 제고하고자 한다.

1. 국내 자동차 수요량 예측 정확도 개선

1) 데이터 수집 및 변수 선정

ANFIS 모형은 머신러닝 기반의 인공지능망 모형의 한 종류로 기존의 회귀 모형이 독립변수의 활용에 많은 제약이 있는 것과 달리 종속변수에 영향을 미칠 것으로 판단이 되는 다양한 변수들을 동시에 모형에 포함시킬 수 있으며, 비선형적 관계도 파악할 수 있다는 장점이 있다. 본 연구에서는 국가별 인구·사회 및 경제 관련 지표들의 장기 전망치를 제공하는 IHS Markit의 통계 자료를 활용하였으며, 분석 대상 변수 선정에는 선행연구를 참고하였다.

Vahabi et al.(2016)은 이란의 자동차 판매량 예측을 위해 환율, 인플레이션율, 1인당 소득, 대출 이자율, 수입 관세, 수입금액, 신규주택 착공금액 등을 변수로 활용하였다.³³⁾ Islam et al.(2016)은 말레이시아 자동차 판매량에 영향을 미치는 요인으로 소비자 가격 지수(CPI, Customer Price Index), GDP, 이자율, 환율, 실업률, 인플레이션율 등을 선정하였다,³⁴⁾ 또한 Zhang et al.(2017)은 중국의 전기차 수요예측 모형에 소비자 가격 지수, 소비자 신뢰지수, 생산자 가격 지수, 연료 소매 가격, 차량 가격 등의 변수를 활용하였다.

한편 IHS Markit은 세계 200여 국가의 500개 가량의 인구·사회 및 경제 지표 실적치 및 증장기 전망치를 연도별, 분기별, 월별로 제공하고 있다.³⁵⁾ 대표적인 지표로는 인구, 실업률, 임금, 국민 소득 계정, 환율 및 물가 등이다. 회귀 모형 또는 머신러닝 기반의 모형을 활용하여 증장기 예측을 수행하기 위해서는 모형에 투입된 외생변수들의 예측 기간 내 전망치들이 요구되며, 이에 예측모형 개발에 있어서 IHS Markit 자료의 활용도가 매우 높다고 할 수 있다.

본 연구에서는 기존 모형에서 고려한 우리나라 인구수 외에 선행연구에서 제안한 변수들과 IHS Markit에서 실적치 및 전망치 입수 가능성 등을 종합적으로 고려하였다. 소비자 가격 지수, 단기 이자율, GDP, 실업률, 임금 지수 등을 설명변수로 선정하였으며, 종속변수는 연도별 국내 국산 및 수입(해외 OEM생산 포함) 신차 판매량을 국내 연도별 자동차 수요로 설정하였다. 각 변수는 1996년부터 2021년까지 연도별로 관측되었으며, 기초 통계량은 다음의 <표 4-1>과 같다.

33) Vahabi et al.(2016), p. 27.

34) Islam et al.(2016), pp. 874-876.

35) IHS Markit(검색일: 2022.9.12.)

〈표 4-1〉 국내 자동차 수요 예측모형 변수 및 기초통계량

변수	단위	표본수	평균	표준편차
소비자 가격 지수	2015년=100	26	85.63	16.02
단기 이자율	%	26	4.59	3.79
GDP	10억 달러	26	1092.76	440.72
실업률	%	26	3.75	1.00
임금 지수	2015년=100	26	80.41	24.89
인구수	백만 명	26	48.89	2.03
자동차 판매량	천 대	26	1526.43	284.29

자료: 저자 작성

2) 모형 설계

ANFIS 모형은 데이터를 훈련용과 검증용 그룹으로 구분하여 훈련용 그룹의 데이터만을 사용하여 모형을 학습하고 검증용 그룹으로 학습된 모형을 평가한다. 본 연구에서는 2장 2절에서 소개한 k -Fold 교차검증법을 활용하여 개발된 모형에 대한 검증을 강화하고자 하였다.

학습 단계에서는 퍼지추론 함수로 Zadeh 함수를 활용하였으며, 경사 하강법(gradient descent)의 크기(step size)는 0.01로 설정하였다. 또한 접속 연산자(conjunction operator)는 일반적인 최소값으로 설정하였으며, 최대 반복 횟수는 1,000회로 반복 학습하였다. 분석에는 통계프로그램 R(ver.3.6.3)의 ‘frbs’ 패키지를 활용하였다.

10-Fold 교차검증을 위해 첫 번째 훈련용 데이터는 초반 16개의 자료를 사용하여 학습하고 17번째 자료의 외생변수들의 실측치로 17번째 자동차 판매량을 예측하였다. 예측치는 해당 시점의 실제값과 비교하여 RMSE와 MAPE를 계산하였다. 이후 훈련용 데이터를 한 시점씩 확대하면서 최종 26번째 시점의 예측치와 실제값의 RMSE와 MAPE를 계산하고, 최종적으로 10개의 RMSE와 MAPE를 각각 평균하여 모형의 성능을 평가하였다. 또한

모형 간 성능을 비교하기 위해 기존 방식인 SLR 모형과 시계열 모형 인 ARIMA 모형을 이용하여 10-Fold 교차검증을 하였다.

3) 분석결과 및 검증

ANFIS 모형과 ARIMA 및 SLR 모형으로 2012~2021년 기간 국내 자동차 수요를 예측한 결과는 다음의 <표 4-2>와 같다. 대부분 기간에서 ANFIS 모형으로 예측한 결과가 실제값에 근사한 값을 제공하고 있으며, SLR 모형은 인구수 변수의 계수값에 따라 매년 유사한 크기로 증가하는 추세를 보이는 반면 ANFIS 모형은 실제값의 변동과 유사한 추세를 보였다.

<표 4-2> ANFIS 모형과 SLR 모형의 예측값 비교

단위: 천 대

연도	실제값	ANFIS 예측값	ARIMA 예측값	SLR 예측값
2012	1565.168	1623.648	1449.201	1395.079
2013	1573.232	1594.516	1436.670	1451.995
2014	1712.075	1567.459	1461.250	1502.278
2015	1889.730	1629.215	1546.120	1581.458
2016	1877.164	1743.332	1641.891	1682.985
2017	1839.782	1801.925	1601.505	1745.193
2018	1855.798	1873.561	1631.595	1779.747
2019	1842.112	1809.311	1639.697	1805.693
2020	1932.305	1854.129	1644.285	1815.536
2021	1740.049	1907.784	1727.266	1830.397

자료: 저자 작성

모형의 성능 평가는 예측값과 실제값의 차이인 오차를 <표 2-1>의 예측 정확도 지표로 산출하고, k -Fold 교차검증을 활용하여 각 모형을 비교하였다. <표 4-3>은 ANFIS 모형과 ARIMA 및 SLR 모형에 대해 10-fold 교차검증으로 산출한 예측 정확도 결과를 보여준다. ANFIS 모형의 RMSE와 MAPE가 SLR 모형보다 더 작게 나타나 ANFIS 모형이 모형의 예측 정확도

를 향상시킨 것으로 나타났다. 특히 MAPE 값을 비교하면 ANFIS 모형이 SLR 모형 대비 오차율을 2.7%p 감소시킨 것으로 나타났다.

〈표 4-3〉 ANFIS 모형과 SLR 모형의 예측 정확도 비교

모형	RMSE	MAPE
ANFIS	95.3059	5.3
ARIMA	204.7935	11.3
SLR	141.775	8.0

자료: 저자 작성

ANFIS 모형을 활용한 국내 자동차 수요예측은 기존 모형이 인구수 단일 변수에 의존함에 따라 발생하는 문제점을 보완하여 다양한 변수의 영향력을 모형 내에서 고려하고 있으며, 이를 통해 정교한 예측 결과를 제공한다. 특히 표본 외 검정(out-of-sample validation)으로 예측력을 비교할 시 ANFIS 모형의 예측 정확도가 기존 모형 대비 월등히 우수한 결과를 보였다. 다만 ANFIS 모형은 다른 머신러닝 기반의 모형들과 마찬가지로 투입 변수들의 종속변수에 대한 영향을 직관적으로 설명하기 어렵다는 단점이 있다. 이에 예측 결과에 대한 논리적인 설명을 뒷받침하기 위해 영향요인 분석을 추가로 고려할 수 있다.

2. 수입자동차 수요 예측모형 개발

1) 데이터 수집 및 변수 선정

VAR 모형은 대표적인 시계열 자료 분석 기법인 ARIMA 모형의 장점과 다양한 외생변수의 영향을 고려하는 회귀 모형의 장점을 동시에 만족하는 모형이다. ARIMA 모형처럼 미래의 전망치가 과거 관측치의 특정 규칙에 따라 예측될 수 있다고 가정하고 있으며, 회귀 모형과 유사하게 변수 간의

상호작용을 분석할 수 있다. 특히 회귀 기반의 모형을 사용하여 미래 전망치를 예측하기 위해서는 개별 외생변수의 전망치 확보가 필수적인 반면, VAR 모형을 활용하면 외생변수와의 관계를 고려하되 별도로 외생변수의 전망치가 요구되지 않는다.

한편 본 연구에서 VAR 모형에 활용한 변수로는 기존 모형에서 활용한 우리나라 연도별 GDP와 앞서 전체 우리나라 자동차 수요 예측모형에서 핵심 변수로 활용한 우리나라 인구수를 고려하였다. GDP와 인구수 변수는 각각 수입차 판매량과 높은 상관관계를 보였으며, 수입차 판매량 전망에 있어서 논리적 설명이 다른 외생변수에 비해 쉬운 장점이 있다. 각 변수는 1996년부터 2021년 기간 연도별로 기록되었으며, 각 변수의 기초통계량은 다음의 <표 4-4>와 같다.

<표 4-4> 수입 자동차 수요 예측모형 변수 및 기초통계량

변수	단위	표본수	평균	표준편차
GDP	10억 달러	26	1092.76	440.72
인구수	백만 명	26	48.89	2.03
수입차 판매량	천 대	26	137.11	127.45

자료: 저자 작성

2) 모형 설계

VAR 모형의 일반식은 앞서 2장 1절에서 소개한 바와 같으며, 본 절의 수입자동차 수요 예측모형은 다음과 같다.

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + e_t \quad (\text{IV-1})$$

여기서 $y_t = (y_{1,t}, \cdots, y_{k,t})$ 는 k 개의 내생변수로 이루어진 벡터이며, 본 연구에서는 GDP, 인구수, 수입차 판매량으로 구성되었다.

한편 VAR 분석에 앞서 실시한 정상성(stationary) 검정에서 각 변수들은 정상성을 만족하지 못했다. 이는 비교적 짧은 시계열 자료 내에 IMF 외 환위기, 금융위기 등 사회 전반적으로 충격이 발생한 시점이 포함된 영향이 큰 것으로 판단되며, 로그 변환(logarithm transformation)과 차분(differencing) 후에도 정상성을 확보하지 못했다. 이는 경제변수를 활용한 시계열 자료의 실증분석에서 흔하게 발생하는 문제로 모형의 활용에 추가적인 주의가 필요함을 의미한다.

모형의 성능 검증을 위해 10-Fold 교차검증법을 활용하였으며, 이를 위해 10개의 시계열 자료를 구축하였다. 이후 각 시계열 자료에 대해 VAR 모형으로 예측한 전망치를 해당 시점의 실제값과 비교하여 RMSE와 MAPE를 계산하였다. 최종적으로 10개의 RMSE와 MAPE를 각각 평균하여 모형의 성능을 평가하였으며, 모형 간 성능 비교를 위해 기존 방식인 SLR 모형을 이용하여 10-Fold 교차검증을 실시하였다. 분석에는 통계프로그램 R(ver. 3.6.3)의 'vars' 패키지를 활용하였다.

3) 분석결과 및 검증

VAR 모형과 SLR 모형으로 2012~2021년 기간 국내 수입자동차 수요를 예측한 결과는 다음의 <표 4-5>와 같다. SLR 모형은 예측 기간 내 전반적으로 과소 추정하는 경향을 보인 반면, VAR 모형은 과거 데이터의 패턴에 따라 증감하며 대체로 실제값에 유사한 추세를 보였다. 또한 SLR 모형이 대체로 연도별 안정적인 변화를 보인 반면 VAR 모형의 예측값은 상대적으로 다이내믹한 변화를 보였다. VAR 모형의 경우 보다 장기간의 시계열 자료가 확보된다면 예측값의 변동폭이 확연히 감소할 수 있을 것으로 예상된다.

〈표 4-5〉 VAR 모형과 SLR 모형의 예측값 비교

단위: 천 대

연도	실제값	VAR 예측값	SLR 예측값
2012	154.407	183.534	102.365
2013	191.068	147.909	125.881
2014	266.431	163.515	157.910
2015	331.858	282.425	179.092
2016	304.378	419.447	214.103
2017	297.238	349.037	259.235
2018	323.607	273.860	291.691
2019	324.580	301.214	278.641
2020	338.135	288.385	282.879
2021	311.480	319.515	335.031

자료: 저자 작성

〈표 4-6〉은 VAR 모형과 SLR 모형에 대해 10-Fold 교차검증으로 산출한 예측 정확도 결과를 보여준다. VAR 모형의 RMSE와 MAPE가 SLR 모형보다 더 작게 나타났으며, 이는 VAR 모형이 모형의 예측 정확도를 향상시켰음을 의미한다. 특히 MAPE 값을 비교하면 VAR 모형이 SLR 모형 대비 오차율을 5.5%p 감소시킨 것으로 나타났다.

〈표 4-6〉 VAR 모형과 SLR 모형의 예측 정확도 비교

모형	RMSE	MAPE
VAR	52.240	19.0
SLR	66.346	24.5

자료: 저자 작성

VAR 모형을 활용한 국내 수입자동차 수요예측은 기존 모형이 국내 GDP 단일 변수를 활용한 단순선형회귀 모형을 활용함에 따라 발생하는 문제점을 보완하였다. 단순선형회귀 모형은 표본 내 검정(in-sample validation)에서는 우수한 성능을 보였으나, 2000년대 이후 급격히 증가한 수입자동차 판매량을 선형 관계로 설명하기는 어려우며 다른 설명변수의 영향을 고려

하지 않는 한계가 있다. 이에 비해 VAR 모형은 추가적인 설명변수의 영향을 고려하고 있으며, 특히 표본 외 검정으로 예측력을 비교할 시 VAR 모형의 예측 정확도가 기존 모형 대비 향상된 결과를 보인 점에서 개선이 이뤄졌다고 판단된다. 다만 VAR 모형은 회귀 모형이나 ARIMA 모형 대비 설명변수들의 영향을 직관적으로 설명하기 어려우며, 머신러닝 기반의 모형과 비교해서는 활용 가능한 투입변수의 수에 제약이 있다. 이에 투입변수들의 월별 또는 분기별 자료를 활용하여 관측치수를 확대하거나, 모형의 활용도를 제고하기 위해 정상 시계열을 만족시키는 방안을 추가로 고려할 필요가 있다.

제2절 석탄 항만물동량 예측모형 고도화

기존 석탄 항만물동량 예측모형 진단을 통해 발굴한 예측모형 고도화 방안은 (1) 석탄 수요예측을 위한 신규 변수 발굴, (2) 제철용 유연탄 수입물동량 예측모형 다각화이다. 첫 번째 개선 과제인 석탄 수요예측 정확도 개선을 위한 신규 변수 발굴에 대해서는 기존 모형에서 고려하지 않았던 국내외 거시경제 지표와 석탄 수요 간의 상관분석을 통해 석탄 수요에 유의미한 영향을 미치는 변수를 발굴하고자 한다. 두 번째 제철용 유연탄 수입물동량 예측모형은 기존에 원단위법만을 활용한 단일 모형 접근법에서 벗어나 다중회귀 모형과 머신러닝 기반 모형 등 다양한 통계방법론을 적용하고 모형 간 예측력을 비교·검토함으로써 예측 결과의 신뢰성을 제고하고자 한다.

1. 석탄 수요예측을 위한 신규 변수 발굴

본 연구에서는 석탄 물동량 예측을 위한 추가 변수 발굴을 위해 기존 모형에서 고려한 조강생산능력 외에 선행연구에서 제안한 변수들과 IHS Markit에서 제공하는 인구·사회 및 경제 관련 지표들을 고려하였다.

조상호(2020)는 석탄 물동량에 영향을 주는 요인을 공급 관련과 시장·경제 관련으로 구분하여 제시하였으며, 케이프선 운임지수, 입항선박 척수 및 톤수, 석탄 및 철광석 가격, 시장전망지표, 광공업생산지수, 대미환율 등 총 7개의 관련 요인 중 입항선박 톤수, 석탄가격 및 대미환율을 최종변수로 선정하였다. IHS Markit의 통계자료 중에서는 석탄 물동량과 유의성이 있는 것으로 판단되는 소비자물가지수와 노동인구수, 명목GDP지수, 유연탄 가격, 한국은행 정책금리 등을 추가 고려 가능한 변수로 선정하였다.

선정된 변수들은 석탄 수입물동량과 연관관계를 분석하는 상관분석을 실시하였으며, 양(+) 관계를 가지는 긍정요인과 음(-)의 관계를 가지는 부정요인으로 구분하였다. 상관계수는 변수 간 선형 관계를 계량하기 위한 피어슨 상관계수를 활용하였으며, 분석결과는 <표 3-7>과 같다. 긍정요인으로는 노동인구수가 0.8874로 가장 높았으며 부정요인으로는 중앙은행 정책금리가 가장 높았다. 기존 연구에서 선정한 대미달러 환율보다 실질거래 환율이 더 높은 음의 상관관계를 보였으며, 유연탄 및 철광석 가격은 각각 0.6129와 0.4605로 상관계수가 상대적으로 약한 양의 관계를 보였다. 조강생산능력의 경우 물동량이 적은 포항의 조강생산능력이 전체 유연탄(석탄) 물동량 수입에 더 높은 상관관계를 보였다.

〈표 4-7〉 석탄 수입물동량 예측용 변수 상관분석 결과 요약

유형	변수	단위	상관계수
긍정요인	포항항 조강생산능력	천 톤	0.9168
	노동인구수	명	0.8874
	명목GDP지수	인덱스	0.8659
	소비자 물가지수	인덱스	0.8550
	광양항 조강생산능력	천 톤	0.7983
	유연탄 가격	달러/톤	0.6129
	철광석 가격	달러/톤	0.4605
부정요인	중앙은행 정책금리	%	-0.8218
	실질 거래 환율	인덱스	-0.4984

자료: 저자 작성

2. 제철용 유연탄 수입물동량 예측모형 다각화

석탄의 제철용 수입물동량은 국내에서 고로방식을 활용하는 제철소인 포스코와 현대제철 인근에 위치한 광양항과 포항항, 평택·당진항을 대상으로 예측한다. 다만 평택·당진항의 경우 2009년 개장하여 실적 데이터가 많지 않고 변동성이 심해 본 연구의 분석대상에서 제외하였다. 예측모형으로는 단순선형회귀(SLR), 다중선형회귀(MLR) 및 서포트벡터회귀(SVR, Support Vector Regression) 모형을 적용하였으며, 제철용 유연탄 수입물동량 실적치와 각 모형을 통한 예측치의 차이를 MAPE로 평가하였다. 또한 분석 자료는 1994년~2021년 기간의 변수별 연간 데이터이며, 2012년~2021년 기간의 예측치를 k -Fold 교차검증법($k = 10$)을 적용해 모형간 비교·검정을 실시하였다.

먼저 포스코 포항과 포스코 광양의 조강생산능력과 수입되는 제철용 유연탄 물동량에 대한 단순선형회귀분석을 실시하였다. 분석 결과, 조강생산능력과 유연탄 수입 물동량의 조정결정계수($adj. R^2$)는 0.732로 나타났다.

또한 F 검정과 t 검정 결과 총 조강생산능력이 제철용 유연탄 수입물동량의 변동성을 설명하기에 유의한 것으로 분석되었다.

〈표 4-8〉 유연탄 수입물동량 SLR 분석결과 요약

변수	계수	표준오차	t 검정	F 검정	adj. R^2
상수	-7,574,492.6	1,083.4	-1.8	74.7***	0.732
총 조강생산능력	1083.4	125.2	8.6***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
자료: 저자 작성

다중선형회귀분석을 위한 변수 선정에는 앞서 실시한 상관분석의 결과를 참고하였으며, 설명변수 간 상관관계가 높아 회귀분석 시 다중공선성 문제가 발생할 가능성이 큰 변수는 제외하였다. 결과적으로 중앙은행 정책금리, 실질거래환율, 총조강생산능력이 설명변수로 선정되었으며, 분석결과 조정 결정계수값은 0.77로 단순회귀분석 결과보다 0.038 상승했으며 중앙은행 정책금리와 총조강능력이 통계적으로 유의한 결과를 보였다.

〈표 4-9〉 유연탄 수입물동량 MLR 분석결과 요약

변수	계수	표준오차	t 검정	F 검정	adj. R^2
상수	-6635.8	13964.9	-0.4	31.2***	0.770
중앙은행 정책금리	-2491.0	1101.7	-2.2*		
실질거래 환율	143.8	85.6	1.6		
총 조강생산능력	0.8	0.2	3.6***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
자료: 저자 작성

모형 간 성능 비교를 위해 k -Fold 교차검증법($k = 10$)을 적용해 MAPE를 비교한 결과는 〈표 3-10〉과 같다. 다중회귀모형은 조강생산능력만을 변수로 하는 단순회귀 모형 대비 1.3%p 예측력이 개선되었다. 서포트벡터 회귀모형(SVR)의 경우는 실적 데이터(학습 데이터 수)의 부족으로 인해 예측

력의 개선 효과가 없는 것으로 나타났다. 또한 별도로 분석을 진행한 인공 신경망 분석 결과도 SVR 대비 낮게 나와 석탄물동량 예측모형 고도화를 위해서는 더 많은 변수를 포함한 데이터 확보와 모형 설계에 대한 연구가 진행되어야 할 것으로 판단된다.

〈표 4-10〉 유연탄 수입물동량 모형별 예측 정확도 비교

단위: %

예측연도	SLR	MLR	SVR
2012	25.1	24.8	41.4
2013	10.2	13.8	34.1
2014	11.5	19.8	39.9
2015	21.0	26.0	47.8
2016	20.5	18.1	49.5
2017	17.7	12.8	49.2
2018	14.7	10.3	48.7
2019	12.5	7.0	48.3
2020	9.2	2.7	43.4
2021	14.7	8.3	46.5
평균	15.7	14.4	44.9

자료: 저자 작성

한편 기존의 Bottom-up 방식의 예측 정확도를 개선하기 위해서는 항만별 유연탄 수입물동량 예측이 필요하다. 이에 항만별로 기존의 원단위법과 SLR 및 MLR 분석을 별도로 실시하여 각 모형의 성능을 비교·분석하였다. 먼저 광양항 수입물동량의 경우 포스코 광양제철소의 조강생산능력을 활용한 SLR 모형 대비 MLR 모형의 설명력과 유의성이 더 높은 것으로 나타났다. 반면 포항항의 경우 추가된 변수를 제외한 SLR 모형의 설명력이 더 높은 것으로 나타났다.

〈표 4-11〉 광양항 유연탄 수입물동량 MLR 분석결과 요약

변수	계수	표준오차	t 검정	F 검정	adj. R^2
상수	-3648.5	7382.9	-0.4	44.9***	0.830
중앙은행 정책금리	-2293.5	627.5	-3.6***		
실질거래 환율	139.5	52.5	2.6**		
광양 조강능력	0.8	0.2	4.3***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
 자료: 저자 작성

〈표 4-12〉 포항항 유연탄 수입물동량 MLR 분석결과 요약

변수	계수	표준오차	t 검정	F 검정	adj. R^2
상수	-5081.6	7774.4	-0.6	10.1***	0.503
중앙은행 정책금리	-110.7	562.1	-0.1		
실질거래 환율	0.8	44.1	0.0		
포항 조강능력	1.1	0.4	3.0***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
 자료: 저자 작성

항만별 수입 물동량 예측모형 간 예측 정확도를 비교·평가하기 위해 기존 원단위 모형(수입물동량 대비 조강생산능력 비중)과 SLR 모형 및 MLR 모형의 성능을 비교했다. 2011년~2021년 기간 광양항과 포항항의 실제 물동량과 모형별 예측값을 비교한 결과 광양항은 MLR 모형이 7.9%로 가장 좋은 성능을 보였다. 포항항의 경우 MLR 모형과 원단위 모형의 MAPE 평균값이 동일하지만 편차는 다중회귀모형이 더 컸다. 이는 기존 원단위 법이 상대적으로 안정적인 예측값을 제공한다고 볼 수 있다.

〈표 4-13〉 항만별 유연탄 수입물동량 모형별 예측 정확도 비교

단위: %

예측연도	광양항			포항항		
	SLR	MLR	원단위	SLR	MLR	원단위
2012	7.1	19.7	17.2	7.4	6.7	0.1
2013	12.8	3.5	11.8	2.3	1.9	12.3
2014	11.1	10.2	7.8	1.8	1.2	2.5
2015	11.7	5.1	11.4	2.3	2.2	12.0
2016	15.2	6.4	5.3	1.4	1.3	4.9
2017	16.3	8.6	7.2	0.2	0.1	4.3
2018	15.7	10.2	7.1	1.2	0.9	7.0
2019	6.8	1.8	4.7	16.1	15.1	6.7
2020	5.0	10.5	6.2	12.8	11.2	1.2
2021	12.6	3.5	8.3	16.9	15.9	5.2
평균	11.4	7.9	8.7	6.2	5.6	5.6

자료: 저자 작성

분석결과를 종합하면 제철용 유연탄 수입물동량 예측에 있어서 MLR 모형이 Top-down과 Bottom-up 방식 양 측면에서 모두 가장 우수한 예측력을 보였다. 다만 성능의 차이가 기존 모형 대비 월등히 우수한 정도는 아니며, 기존 모형의 효율성과 활용성을 고려하면 모형 교체에 따른 이익이 크지 않을 것으로 판단된다. 그럼에도 앞서 논의한 바와 같이 향후 석탄 수요에 변화가 예견된 만큼 다모형 접근법을 통해 물동량 변화를 다각적으로 검토하고 선제적으로 대비할 필요가 있다. 특히 현재 확보된 연간 데이터만으로는 성능이 불확실한 SVR 모형 등 머신러닝 기반 모형을 월간 또는 분기별 자료 학습을 통해 활용하는 등의 추가적인 연구가 필요할 것으로 판단된다.

제3절 고철 항만물동량 예측모형 고도화

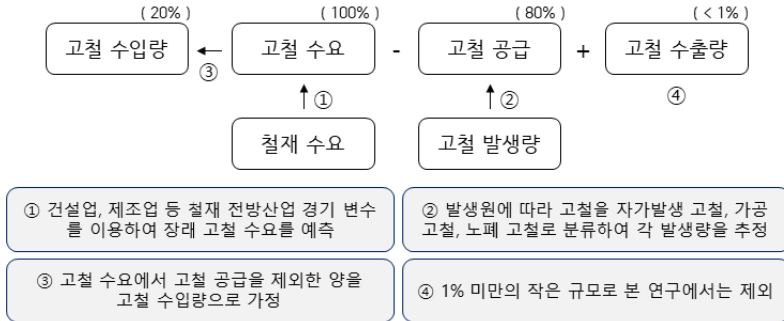
앞서 3장에서 진단한 바와 같이 기존 모형은 고철 국내 자급률이 증가함에 따라 수입 물동량이 감소하는 최근 추세를 반영하지 못하고 있다. 이에 따라 본 연구에서는 고철 수요와 공급을 각각 추정하여, 이를 토대로 수입 물동량을 예측하고자 하였다. 분석에 앞서, 국내에서 발생하는 고철은 99% 이상 국내에서 소비되고 있기에 고철 수입량은 국내 고철 수요에서 공급을 제외한 양으로 가정하였다.

먼저, 본 절의 첫 번째 부분에서는 고철 수요 추정모형을 개발하고자 하였다. 고철은 철강의 원료이고 고철 수요는 건설, 제조, 조선, 자동차 등 철재 전방산업의 수요에 따라 발생한다. 따라서 해당 산업과 관련된 거시 경제변수를 활용하여 장래 고철 수요를 파악하고자 한다.

두 번째 부분에서는 고철 공급 추정모형을 개발하고자 한다. 고철은 발생원에 따라 철강사의 조강생산공정에서 발생하는 자가발생 고철, 건설·제조업 등 수요산업의 가공공정에서 발생하는 가공 고철, 내용연수가 끝난 건물이나 장비가 폐기되어 발생하는 노폐 고철으로 분류할 수 있다. 이에 따라, 각 고철의 종류에 따라 발생량을 추정하여 합산하면 고철 공급량을 산정할 수 있다.

마지막 세 번째 부분에서는 앞서 추정한 수요와 공급 모형을 반영하여 고철 수입량 예측모형을 개발하고자 한다. 또한, 위 과정을 통해 개발된 새로운 모형과 기존 모형의 예측 정확도를 비교하여 최종 모형을 선택하고자 한다. <그림 4-1>은 본 연구에서 제안하는 고철 수입 물동량 예측모형의 개념도이다.

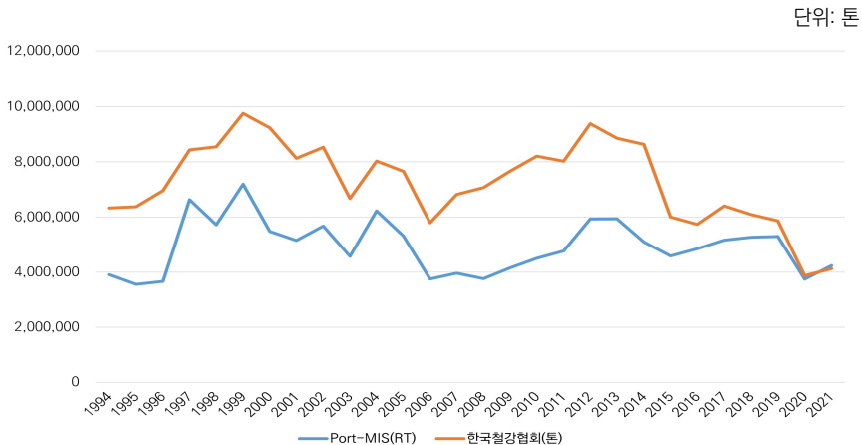
〈그림 4-1〉 고철 수입 물동량 고도화 방안 도식



자료: 저자 작성

한편, 고철 수입 자료는 해양수산부의 Port-MIS 항만통계와 한국철강협회를 통해서 얻을 수 있다. Port-MIS는 운임톤(RT) 기준으로 고철 수입량을 집계하며, 한국철강협회는 금액(달러) 및 중량(톤) 기준으로 집계된 관세청 통계자료를 활용한다. 그러나 고철 국내 구매량, 자가 발생량 등의 고철 수급 관련 세부 데이터는 현재 한국철강협회에서 유일하게 제공하고 있기에, 본 연구에서는 분석의 정합성을 위해 한국철강협회 자료를 활용하였다.

〈그림 4-2〉 Port-MIS와 한국철강협회의 고철 수입 물동량 차이



자료: 저자 작성

1. 국내 고철 수요 예측모형 개발

철강 수요와 관련된 변수 중 2040년 이상의 장기 예측치가 제공되는 거시경제변수는 GDP, 인구, 건설업 GDP, 제조업 GDP, 제조업 생산지수(Production Index; PI) 등이 있다. 먼저 중장기 고철 수요를 예측하기 위한 모형 선택을 위해 stepwise 방법으로 변수를 선택한 회귀 모형과 다양한 추세 분석기법을 적용하여 비교하였다.

〈표 4-14〉 철강 수요 관련 변수 추이 및 전망

단위: 십억 달러, 백만 명, 2015=1.00

연도	GDP(PPP)	건설 GDP	제조 GDP	인구	제조업 PI
2018	2,099.5	91.74	483.53	51.68	1.07
2019	2,146.6	89.37	488.93	51.80	1.06
2020	2,131.4	88.19	483.69	51.84	1.07
2021	2,219.8	85.90	517.02	51.83	1.15
2030	2,668.8	119.55	708.90	51.29	1.36
2040	2,995.4	127.76	799.64	49.32	1.51

자료: IHS Markit(검색일: 2022.9.12.), 저자 재작성

회귀모형을 기반으로 한 수요 추정 함수 도출 결과, 제조업 생산지수가 단일 독립변수로 들어간 모형이 선택되었다. 변수 선택을 위해서는 AIC와 각 변수의 t 검정을 기준으로 한 stepwise 방법을 적용하였다. 1991년부터 2021년 데이터를 바탕으로 시행한 회귀분석 모형의 결과는 〈표 4-15〉와 같다. 제조업 생산지수를 독립변수로 한 회귀분석 결과, 제조업 생산지수가 1 증가할 때 고철 수요는 평균 1,710.5만 톤만큼 증가하는 것으로 분석되었다. 고철 수요는 2022년부터 2040년까지 평균 0.88%로 증가하여, 2030년에는 3,603만 톤, 2040년에는 3,859만 톤에 이를 것으로 예측되었다.

〈표 4-15〉 고철 수요예측 회귀 모형 요약

변수	계수	표준오차	<i>t</i> 검정	<i>F</i> 검정	adj. <i>R</i> ²
상수	1274.7	119.6	10.659	117.3***	0.802
제조업 생산지수	1710.5	157.9	10.831***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
자료: 저자 작성

추세모형 중에는 로그 추세모형이 다른 추세모형보다 우수한 적합 결과를 보였으며, 제조업 생산지수를 독립변수로 둔 회귀모형보다도 설명력이 높게 나타났다. 1991년부터 2021년 데이터를 바탕으로 시행한 로그 추세모형의 결과는 〈표 4-16〉와 같다. 로그 추세모형을 통해 고철 수요를 예측한 결과, 2022년부터 2040년까지 평균 0.50%로 증가하여, 2030년에는 3,226만 톤, 2040년에는 3,372만 톤에 이를 것으로 예측되었다.

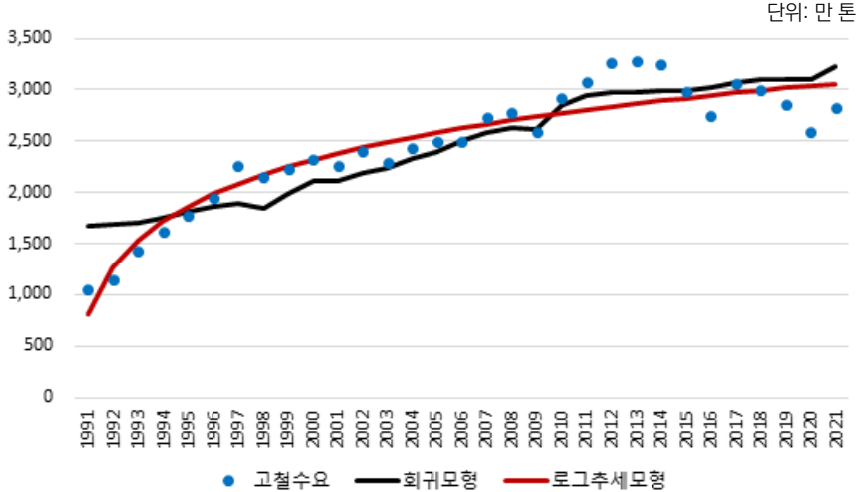
〈표 4-16〉 고철 수요예측 로그 추세모형 요약

변수	계수	표준오차	<i>t</i> 검정	<i>F</i> 검정	adj. <i>R</i> ²
상수	810.5	114.5	7.079	230.6***	0.888
로그 추세	654.8	43.1	15.186***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
자료: 저자 작성

회귀모형과 로그 추세모형 추정 및 예측 결과는 〈그림 4-3〉과 같다. 장기적으로는 회귀모형이 로그 추세모형보다 더 높은 고철 수요예측치를 나타냈다.

〈그림 4-3〉 고철 수요예측 모형 비교



자료: 저자 작성

2. 국산 고철 공급량 예측모형 개발

고철은 발생하는 방식에 따라 자가발생 고철, 가공 고철, 노폐 고철의 세 가지로 분류할 수 있다. 자가발생 고철은 철강업체의 고로 및 전기로 제강 공정 중 발생하며, 대부분 별도 처리 없이 회수되어 다시 제강공정에 활용된다. 이에 따라 국내 조강생산량에 비례하는 특징을 가진다. 가공 고철은 주로 기계, 자동차, 조선 등 철강을 통해 중간제품 및 완제품을 생산하는 과정에서 발생하는 철강 자투리와 불량품으로 발생한다. 따라서 철강 수요업체 생산량과 밀접한 관련이 있다. 노폐 고철은 내용연수 초과 등의 원인으로 폐기된 고철을 의미한다. 자가발생 고철과 가공 고철이 발생 후 즉시 재활용되는 것과 다르게, 노폐 고철은 철재 완제품이 제조된 후 제품 수명 등 충분한 시간이 흘러야 발생한다. 이에 따라, 노폐 고철의 발생량은 건축물, 선박, 기계 등 국내에서 사용되고 있는 철강 제품의 총량인 철강 축적

량과 비례한다. 또한, 노후화 선박·장비·차량 교체, 건물 재건축 등 정책에 따라서도 단기적인 추이가 바뀔 수 있다. 고철의 국내 공급은 이 세 가지 고철 사용량의 합으로 산정할 수 있다.

먼저, 자가발생 고철 발생량은 한국철강협회 조강생산량 자료를 활용하여 추정하였다. 1991~2021년 데이터를 바탕으로 시행한 회귀분석 모형 결과는 <표 4-17>과 같다. 한편, 3장에서 소개한 기존 고철 예측모형에서도 전기로 조강생산량 예측치를 활용하고 있다. 본 연구에서는 분석방법론 비교를 위해 기존 고철 예측모형에서 적용한 모형을 동일하게 활용하여 조강생산량을 예측하였다. 기존 고철 예측모형에서는 전국 고로 조강생산량과 전기로 조강생산량에 대해 각각 국내 GDP를 단일 변수로 활용한 로그 선형회귀모형을 적용하고 있다.

<표 4-17> 자가발생 고철 발생량 추정모형 결과 요약

변수	계수	표준오차	<i>t</i> 검정	<i>F</i> 검정	adj. <i>R</i> ²
상수	-67.799	30.029	10.659	323.4***	0.918
조강생산량	0.100	0.006	10.831***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
자료: 저자 작성

가공 고철의 발생량은 세계철강협회 자료 중 ‘Apparent consumption of finished steel(이후, 국내 철재 소비량)’ 항목을 활용하여 예측하였다. 해당 항목은 강재, 판재 등 철재 완제품의 국내 출하량에서 수입량을 더하고 수출량을 뺀 값을 의미한다. 즉, 국내 철강업체의 제조·가공 공정에 투입된 철재의 총량을 나타내기에 가공 고철 발생량의 우회변수로 활용할 수 있다. 이 항목의 예측치는 기존 고철 수급과 관련한 연구의 방법론을 따라, 1인당 국내 철재 소비량과 1인당 GDP 간 logistic growth curve 모형을 활용하여 예측하였다(Hatayama et. al, 2010). Logistic growth curve 모형의 형태는 (IV-2)와 같으며, APFPC와 GDPPC는 각각 1인당 국내 철재

소비량과 1인당 GDP를 의미한다. 모수 π, α, β 의 추정에는 통계프로그램 R(ver.3.6.3)의 'stats' 패키지 nls 함수를 활용하였다. 가용한 가장 최근 자료인 1991~2019년 데이터를 통해 분석한 결과, 모형 추정 결과는 <표 4-18>, <그림 4-4>와 같다.

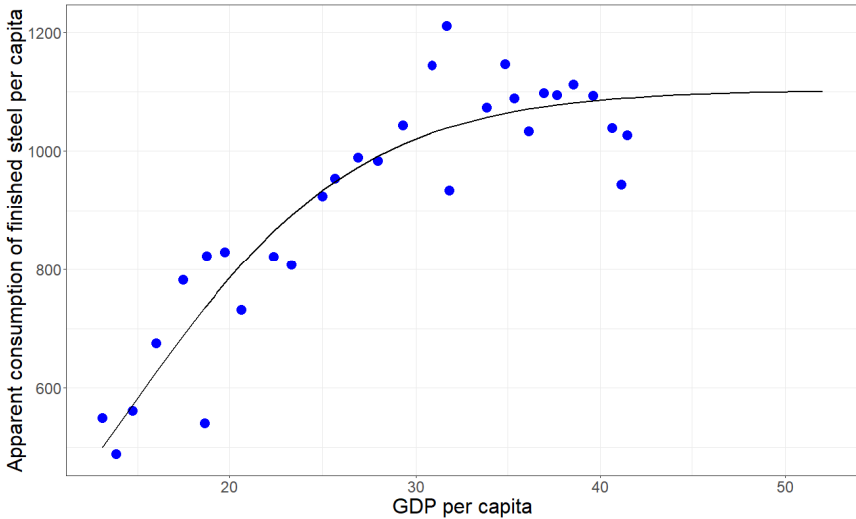
$$APFPC = \frac{\pi}{1 + \exp(\alpha + \beta \times GDPPC)} \quad (IV-2)$$

<표 4-18> 1인당 국내 철재 소비량 추정모형 결과 요약

모수	추정값	표준오차	t 검정
π	1125.59	40.188	28.008***
α	-2.17	0.451	-4.816***
β	0.15	0.029	5.178***

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
자료: 저자 작성

<그림 4-4> Logistic growth curve를 통한 국내 철재 소비량 예측



자료: 저자 작성

노폐 고철의 발생량은 국내 철강 축적량과 비례하여 나타나므로, 한국철강협회의 국내 철강 축적량 자료를 활용하여 산정하였다. 국내 철강 축적량은 1991년부터 2020년까지 매년 약 2천만 톤 정도씩 안정적으로 증가하고 있으며, 해당 기간의 선형 추세선의 결정계수도 0.996으로 매우 높다. 따라서 국내 철강 축적량의 예측치는 선형 추세선을 활용하여 도출하였다. 예측 결과는 다음과 같다.

〈표 4-19〉 국내 철강 축적량 추정모형 결과 요약

변수	계수	표준오차	<i>t</i> 검정	<i>F</i> 검정	adj. <i>R</i> ²
상수	105092	4628.7	22.705	6866.1***	0.996
선형 추세	21064.3	260.7	82.862***		

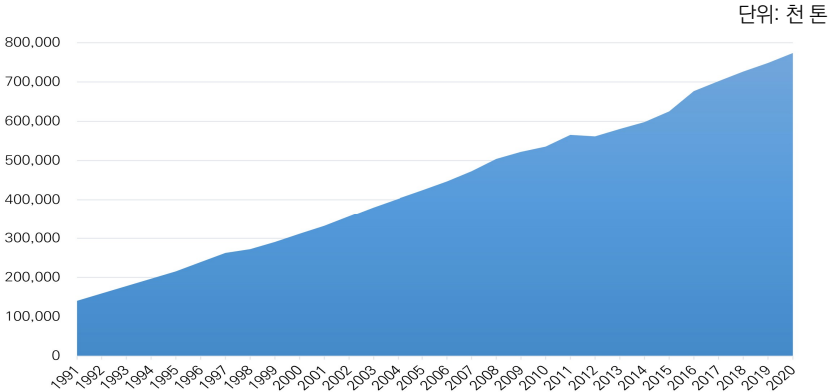
주 : *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
자료: 저자 작성

〈표 4-20〉 국내 철강 축적량

연도	2000	2010	2019	2020	2030	2040
누계축적량	311,389	535,381	748,322	773,543	969,265	1,185,308
전년대비 증가율	1.07	1.03	1.03	1.03	1.02	1.02

자료: 저자 작성

〈그림 4-5〉 국내 철강 축적량



자료: 저자 작성

한편, 한국철강협회에서 제공하는 고철 수급자료는 가공 고철과 노폐 고철의 발생량을 구분하지 않고 ‘국내 구입’ 항목으로 묶어서 제공하고 있다는 한계점이 존재한다. 이는 자가발생 고철이 철강업체에서 발생한 뒤 시중에서 거래되지 않고 철강업체에 의해 직접 활용되는 반면, 가공 고철과 노폐 고철은 고철업체에 의해 수집되어 철강업체에 판매되기 때문이다. 이에 따라, 본 연구에서도 국내 고철 구입량을 종속변수, 철강축적량과 국내 철재 소비량을 독립변수로 둔 회귀모형을 적용하여 분석하였다.

〈표 4-21〉 국내 고철 구입량 추정모형 결과 요약

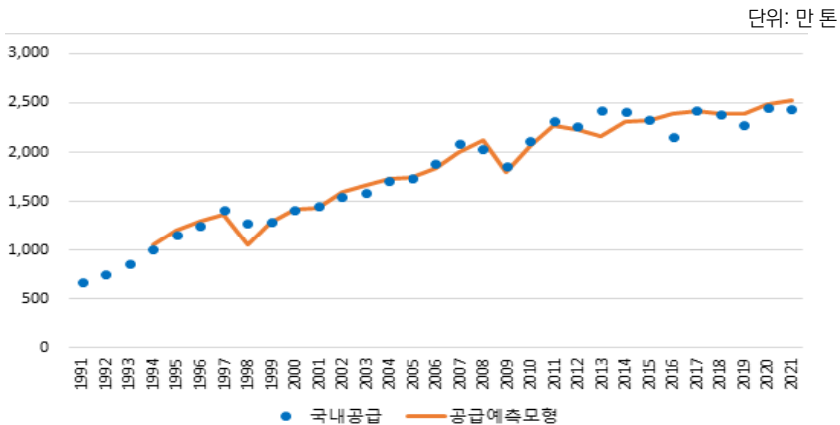
변수	계수	표준오차	<i>t</i> 검정	<i>F</i> 검정	adj. <i>R</i> ²
상수	-46.0589	87.9389	-0.5238	234.4***	0.947
철강축적량	0.0009	0.0002	3.7976***		
국내 철재 소비량	0.0216	0.0037	5.9034***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함

자료: 저자 작성

최종적으로, 자가발생 고철, 가공 고철, 노폐 고철의 공급예측 모형을 통합한 결과는 〈그림 4-6〉과 같이 도출되었다.

〈그림 4-6〉 고철 공급예측 모형



자료: 저자 작성

3. 수급분석을 반영한 고철 수입물동량 예측모형 개발

본 절에서 고도화하고자 하는 고철 수급 반영 수입 예측모형은 앞서 추정된 수요에서 공급을 뺀 차이를 구해 수입 물동량을 예측하는 모형이다. 수요 모형으로는 고철 수요를 제조업 생산지수로 설명하는 회귀모형과 로그 추세선을 사용한 추세모형의 두 가지 후보를 탐색했다. 공급 모형은 자가발생 고철, 가공 고철, 노폐 고철의 발생량을 추정하는 각각의 모형을 개발하여 그 예측값을 합치는 형태로 구성했다. 이 중 수요 모형의 경우, 두 후보 중 예측 정확도가 높은 모형을 최종 모형으로 선택하고자 한다.

또한 고도화된 모형과 기존 모형의 예측 정확도를 비교하여 모형의 개선 정도를 파악하고자 한다. 편의를 위해 고철수급 반영모형 중 고철 수요를 제조업 생산지수로 설명한 모형을 ‘수급회귀모형’, 로그 추세분석을 통해 설명한 모형을 ‘수급추세모형’으로 명명하겠다. 또한 기존 모형 중 전기로 생산능력을 사용한 모형을 ‘기존회귀모형’, 5년 이동평균법을 활용한 모형을 ‘기존추세모형’이라고 지칭하겠다.

총 4가지 고철 물동량 예측모형을 검증하는 방법으로는 k -Fold 교차검증법($k = 10$)을 적용하였다. 먼저 1991~2011년 기간의 연간 자료를 통해 각 모형을 훈련하고 2012년의 고철 물동량을 예측하였다. 예측치는 해당 시점의 실제 관측치와 비교하여 RMSE와 MAPE를 계산하였다. 이후 훈련용 데이터의 연도를 한 시점씩 순차적으로 늘려가며, 1991~2020년 자료를 통해 2021년에 대한 RMSE와 MAPE를 계산하는 데까지 반복하였다. 최종적으로 10개의 RMSE 및 MAPE를 평균한 값이 가장 낮은 모형을 최종 모형으로 선택하였다.

기존 모형과 고철수급 반영모형의 5-fold 교차검증 시행별 예측 결과는 <표 4-22>와 같다. 고철 수요가 급격하게 증가했던 2012~2014년에는 수급

회귀모형을 제외한 모든 모형에서 고철 수입량을 과소 추정했다. 특히 고철수급모형에서 오차가 가장 크게 나타났다. 그러나, 주요 철강사의 전기로 폐쇄 등으로 고철 수요가 감소했던 2015년부터는 수급추세모형이 다른 모형에 비해 가장 낮은 오차를 보였다. 한편, 코로나19의 영향 등으로 고철 경기의 급격한 변화가 있었던 2020년부터는 모든 모형에서 100만 톤 이상의 오차가 발생했다. 특히 수급회귀모형과 기존회귀모형은 300만 톤 이상의 오차가 나타났음을 확인할 수 있었다.

〈표 4-22〉 고철수급 반영모형과 기존 모형의 예측값 비교

단위: 만 톤

연도	실제값	고철수급 반영모형 (수요: 회귀)	고철수급 반영모형 (수요: 추세)	기존 모형 (회귀분석)	기존 모형 (추세분석)
2012	938.1	960.1	572.0	776.9	755.6
2013	885.7	924.4	590.3	808.5	806.8
2014	863.6	869.4	578.7	824.4	842.6
2015	600.4	802.1	577.3	831.8	862.1
2016	573.9	781.5	572.5	818.0	818.0
2017	640.3	792.9	584.2	805.0	772.3
2018	609.1	766.2	576.3	792.4	712.8
2019	586.5	714.0	573.0	776.3	657.5
2020	387.6	703.4	602.8	763.1	602.1
2021	412.8	727.0	523.0	751.5	559.5

자료: 저자 작성

기존 모형과 고철수급 반영모형에 대해 10-fold 교차검증으로 산출한 예측 정확도 비교 결과는 〈표 4-23〉과 같다. 수급추세모형의 RMSE 및 MAPE가 다른 모든 모형보다 작게 나타났다. 특히 MAPE의 경우 기존회귀모형 및 기존추세모형 대비 오차율을 각각 17.0%p, 4.9%p 감소시킨 것으로 나타났다. 수급회귀모형의 경우, 기존회귀모형보다는 우수한 예측 정확도를 보였으나 5년 이동평균법 모형에 비해 낮은 정확도를 나타냈다.

〈표 4-23〉 고철수급 반영모형과 기존 모형의 예측 정확도 비교

모형	RMSE	MAPE
고철수급 반영모형(수요: 회귀분석 적용)	154.3	30.6
고철수급 반영모형(수요: 로그 추세분석 적용)	139.9	20.8
기존 모형(회귀분석: 수입~전기로 생산능력)	200.5	37.9
기존 모형(수입 물동량 5년 이동평균법 적용)	145.6	25.8

자료: 저자 작성

분석 결과를 종합하면, RMSE 및 MAPE 기준으로 최종 선택된 수급추세 모형은 기존 모형에서 반영되지 않았던 수급 구조를 모형에 반영함에 따라 더 높은 예측 정확도를 나타냈다. 특히 고철 수요 증가세가 둔화하기 시작한 2015년 이후부터는 수급추세모형의 예측 정확도가 기존 모형 대비 월등히 우수한 결과를 보였다. 다만 고철 수요 추정에 적용한 추세모형은 이후 고철 수요의 급격한 추세 변화가 나타날 시 큰 예측 오차가 나타날 수 있기에 추후 지속적으로 검토할 필요가 있다.

제4절 시멘트 항만물동량 예측모형 고도화

기존 시멘트 항만물동량 예측모형 진단을 통해 발굴한 모형의 개선점은 (1) 시멘트 수요예측을 위한 변수 선정, (2) 주요 출하항의 단기 물동량 예측 정밀화, (3) 지역별 수요 기반 입하 물동량 분배 방안 개선이다. 〈그림 4-9〉는 본 연구에서 제안하는 시멘트 항만물동량 예측모형의 개념도이다. 본 연구에서는 방법론적 관점에서 세 가지 개선방안을 제안한다.

첫 번째는 시멘트 수요예측을 위한 변수 발굴이다. 기존 예측모형에서는 GDP, 인구 및 건설수주액을 통한 국내 시멘트 수요를 예측한다. 개선 모형

에서는 기존 모형에서 고려되지 않았던 국내외 거시경제 지표와 시멘트 수요 간의 상관분석을 실시한다. 상관분석 결과를 통해 선정된 변수는 시멘트 내수 추정을 위한 변수로 선정된다.

두 번째는 주요 출하항의 단기 물동량 예측 정밀화이다. 주요 출하항의 물동량은 전국 연안 물동량에 많은 영향을 미치는 중요한 요인이기 때문에 정밀한 예측이 필요하다. 기존의 출하량 예측은 예측 시점의 물동량에 따라 크게 변화해 예측 신뢰도를 낮출 수 있는 단점이 있다. 제안된 방법론은 기존 장기 예측치 변화에 큰 영향을 주는 단기 물동량 예측을 정밀화하여 장기 물동량 예측 신뢰도를 높일 수 있을 것으로 기대한다. 본 연구에서는 주요 출하항인 동해·묵호, 삼척 및 옥계항을 대상으로 머신러닝 기반 단기(1년) 물동량을 예측 방법을 제안한다.

세 번째는 지역별 경제 지표 기반 입하 물동량 분배 방안 개선이다. 기존 모형의 입하 물동량 분배는 과거의 물동량 발생 비중을 기반으로 한다. 분명 과거의 연안 물동량 발생 비중은 미래의 연안 물동량 분배를 위해 중요한 요인은 틀림이 없다. 하지만 시멘트 물동량은 수요지역과 공급지역이 분명하게 구분되어 있어 입하항 주변의 수요 추정 지표를 활용하면 연안 물동량 분배의 정확도를 높일 수 있을 것으로 예상된다.

1. 시멘트 수요예측을 위한 신규 변수 발굴

기존 모형은 전국 시멘트 수요를 추정하기 위해서 실질 GDP(PPP 조합)를 변수로 활용하여 국내 총건설수주액을 먼저 추정한다. 실질 GDP(PPP 조합)는 IHS Markit에서 2050년까지의 전망치를 제공하지만, 건설수주액은 전망치를 제공하지 않아 모형에서 자체적으로 전망한다. 기존 모형에서 건설수주액은 전국 시멘트 수요를 예측하기 위한 주요 변수이기 때문에 건

설수주액을 추정하는 것은 매우 중요한 단계다. 따라서 본 장에서는 건설수주액과 시멘트 수요예측의 정확도를 높이기 위해 기존 모형에서 고려되지 않았던 거시경제지표에 대한 변수 활용성을 검토하고자 한다. 본 연구에서는 상관분석을 통해 건설수주액, 시멘트 수요 및 거시경제지표 간 선형 관계를 분석해 추가적인 변수를 선정하고자 한다.

먼저 장기 건설수주액을 추정하기 위해 활용 가능한 2,161개의 연간 거시경제지표 데이터를 활용하여 국내 총건설수주액과의 상관관계를 분석했다. 상관분석을 통해 도출된 변수는 ‘자산 유동성’, ‘인구수(15~24세)’, ‘도시 인구 증가율’과 ‘건설업 명목 GDP’이다. 기존 변수인 ‘총인구수’와 ‘실질 GDP(PPP 조합)’도 각각 0.8560과 0.8647로 높은 양의 상관관계를 가지는 것으로 나타났다.

〈표 4-24〉 건설수주액 추정 변수 상관분석 결과 요약

유형	변수	단위	상관계수
긍정요인	건설업 명목 GDP	달러	0.9425
	총인구수	명	0.8560
	실질 GDP(Harmonized PPP)	달러	0.8647
부정요인	화폐유통속도	포인트	-0.8947
	인구수(15~24세)	명	-0.8927
	도시 인구 증가율	%	-0.8695

자료: 저자 작성

상관분석을 통해 도출된 변수와 기존 변수를 활용하여 총건설수주액을 회귀분석을 통해 추정했을 때의 $\text{adj.}R^2$ 값은 0.9784로 기존 변수만을 사용했을 때($\text{adj.}R^2$:0.919) 대비 0.0594 상승하여 회귀 모형의 설명력을 높였다. 하지만 기존 변수인 ‘총인구수’와 동일한 성격의 지표인 ‘인구수(15~24세)’는 변수 간 높은 다중공선성을 보여 제외했으며 변수 선정 이유에 대한 설명이 어려운 ‘화폐유통속도’를 제외하여 ‘총인구수’, ‘실질 GDP(PPP 조합)’, ‘도시 인구 증가율’ 및 ‘건설업 명목 GDP’를 총건설수주액을

예측하기 위한 회귀분석을 실시했다. 상관분석을 통해 선정된 변수로 회귀 분석을 실시한 결과 조정 R^2 값이 기존 모형보다 감소했으며 t 검정 결과 도 변수의 통계적 유의성을 설명하지 못했다. 따라서 단계적 탐색을 통해 ‘도시 인구 증가율’과 ‘건설업 명목 GDP’를 최종 변수로 선정하여 회귀 분석을 실시했다. 회귀분석 결과 adj. R^2 값은 0.919로 기존 모형과 동일했으며 선정된 두 변수 모두 통계적으로 유의한 것으로 평가되었다.

〈표 4-25〉 총건설수주액 추정 회귀분석 모형 요약

변수	계수	표준오차	t 검정	F 검정	adj. R^2
상수	546090.8	241,273.3	2.2	120.9***	0.919
도시 인구 증가율	-461874.9	162,858.6	-2.8*		
건설업 명목 GDP	9790.3	1,812.6	5.4***		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함
자료: 저자 작성

최종 선정된 회귀분석 모형의 정확도 검증을 위해 2000년~2021년의 실제 건축수주액과 기존 모형 및 제안된 모형의 건설수주액 예측치를 MAPE로 검증했다. 두 가지 모형을 비교·검증하기 위해서 k -Fold 교차검증법($k = 10$)을 적용했으며 2012년부터 2021년까지의 10년의 예측치를 실적치와 비교하고 평균하여 평가하였다. 평가 결과를 보면 제안된 모형이 기존 모형보다 예측 정확도 지표에서 개선된 결과를 보여준다. MAPE로 해석해 보면 제안된 모형의 총건설수주액 예측력이 기존 모형보다 7.0%p 개선됐으며 2017년, 2018년을 제외한 모든 예측연도에서 제안된 모형이 기존 모형보다 높은 정확도를 보였다.

〈표 4-26〉 총건설수주액 예측 결과 요약

단위: %

예측연도	기존 모형	제안된 모형
2012	29.4	14.1
2013	44.6	36.7
2014	18.3	9.2
2015	21.4	21.1
2016	20.7	3.4
2017	12.0	15.3
2018	6.8	14.0
2019	9.9	6.4
2020	21.9	7.1
2021	22.2	10.1
평균	20.7	13.7

자료: 저자 작성

건설수주액 예측값이 도출되면 기존 모형은 로그 변환된 총인구수와 건설수주액으로 시멘트 수요를 예측한다. 기존 모형의 전국 시멘트 수요는 회귀분석, 원단위법과 이동평균법의 예측치 평균으로 정의된다. 본 연구에서 제안하는 개선 모형은 기존 모형과 동일하게 회귀분석, 원단위 및 이동평균법을 평균하여 최종 예측치를 제안한다. 하지만 기존 모형은 시멘트 수요예측을 위한 회귀분석 시 다소 낮은 설명력($\text{adj. } R^2=0.271$)을 보여준다. 따라서 회귀모형은 총인구수와 건축수주액 외에 예측력을 높일 수 있는 추가적인 변수 설정이 반드시 필요하다.

시멘트 수요 예측 모형을 위한 추가 변수 선정을 위해서 2,161개의 거시경제지표와 피어슨 상관계수를 활용한 상관분석을 실시했다. 상관분석을 통해 도출된 외생변수는 ‘건설업 명목 GDP(비중)’, ‘국내 원유가격’, ‘순상품교역조건지수(TOT, Trade of Terms)’와 ‘GDP 대비 경상수지’이다.

〈표 4-27〉 건설수주액 추정 변수 상관분석 결과 요약

유형	변수	단위	상관계수
긍정요인	건설업 명목 GDP(비중)	%	0.7578
	순상품교역지수	포인트	0.6547
	총건설수주액	원	0.0977
부정요인	GDP 대비 경상수지	%	-0.6488
	국내 원유가격	원	-0.6533
	총인구수	명	-0.2052

자료: 저자 작성

상관분석을 통해 도출된 변수를 사용하여 시멘트 수요에 대한 회귀분석을 했을 때의 $\text{adj.}R^2$ 값은 0.8257로 기존의 0.2713 대비 크게 설명력이 상승했다. 하지만 추가된 변수를 모두 시멘트 수요 예측에 사용하였을 때 수요가 크게 발산하는 형태를 보여 현실성이 떨어지는 예측 결과를 도출했다. 따라서 변수 간 적정성을 확인하여 예측 결과가 발산하는 원인인 ‘국내 원유가격’과 통계적 유의성이 적은 ‘총인구수’를 제외한 ‘건설업 명목 GDP(비중)’, ‘순상품교역지수’, ‘총건설수주액’, ‘GDP 대비 경상수지’를 최종 변수로 선정했다. 건설업 명목 GDP(비중)을 포함한 모든 추가 변수에서 통계적 유의성을 보였으며 기존 모형에서 고려되었던 ‘총인구수’는 제외됐다.

〈표 4-28〉 시멘트 수요 예측 회귀분석 모형 요약

변수	계수	표준오차	t 검정	F 검정	$\text{adj.}R^2$
상수	-179515154.9	52702935.2	-3.4	15.22***	0.730
GDP 대비 경상수지	668216.4	241525.9	2.7*		
건설업 명목 GDP(비중)	5580249	1465006	3.8***		
순상품 교역지수	748642.6	202514.1	3.6**		
총건설수주액	7470483.5	2016441.6	3.7**		

주: *는 p-value < .05, **는 p-value < .01, ***는 p-value < .001을 의미함

자료: 저자 작성

2000년부터 2021년까지의 실제 시멘트 수요와 기존 모형 예측치 및 개선 모형 예측치를 비교했다. 제안 모형의 수요 예측 정확도를 평가하기 위해 건설수주액 예측의 검증 방법과 동일하게 k -Fold 교차검증법($k = 10$)을 적용했으며 2012년부터 2021년까지의 10년간 실적치와 예측치를 MAPE로 비교했다. 결과는 2018년과 2019년을 제외하고는 모든 예측연도에서 제안된 모형이 높은 정확도를 보였다. 기존 모형의 10년 평균 MAPE는 10.2%이며 제안된 모형은 4.8%로 5.4%p의 정확도가 개선되었다.

〈표 4-29〉 시멘트 수요 예측 결과 요약

단위: %

예측연도	기존 모형	제안된 모형
2012	10.9	0.1
2013	19.0	7.4
2014	5.2	0.4
2015	10.0	3.9
2016	15.1	0.1
2017	13.8	3.8
2018	1.9	9.6
2019	3.9	5.2
2020	13.2	11.7
2021	9.2	6.1
평균	10.2	4.8

자료: 저자 작성

2. 주요 출하항의 단기 물동량 예측 정밀화

기존 모형은 주요 시멘트 출하항인 동해·목호, 삼척 및 옥계항의 전년도 연안 출하 비중으로 물동량을 예측한다. 이러한 방법론은 전년도의 연안 출하 비중이 지속될 것이란 가정으로부터 시작되는데 2021년과 같이 유연탄 및 시멘트 가격 상승, 일부 제조사의 시멘트 운송 시스템 고장 등과

같이 일시적인 물동량 증감에 따라 장기 물동량 예측치가 크게 변화할 수 있다. 따라서 본 장에서는 예측 시점에 따라 변화가 큰 장기 물동량 예측의 불확실성을 감소시키기 위해서 정밀한 단기 예측이 가능한 머신러닝 방법론을 적용한다.

주요 출하항의 물동량 예측을 개선하기 위해서 기존 방식(전년도 비중), ARIMA, XG부스트와 LSTM의 예측 정확도를 비교한다. 기존 방식은 전년도 물동량 비중을 사용하는 방식이다. ARIMA는 대표적인 시계열 예측 모형으로 데이터의 정상성 및 계절성을 고려한 예측이 가능하다. XG부스트는 다수의 약한 의사결정 모형의 예측치를 결합한 앙상블 머신러닝 모형으로 최근 시계열 데이터 분석에서도 좋은 성능을 내는 것으로 밝혀졌다. LSTM은 전통적인 시계열 예측을 위한 머신러닝 모형으로 시계열 예측 문제에 탁월하다고 알려져 있다. 따라서 본 연구에서는 기존 모형과 시계열 예측에는 탁월한 효과를 보이지만 방법론적 성격이 다른 ARIMA, XG부스트와 LSTM을 적용한다.

본 예측 문제의 대상항은 주요 출하항인 동해·묵호항, 삼척항과 옥계항이며 시간적 범위는 1994~2021년이다. 예측력 검증을 위해 1994년~2011년까지를 초기 학습 데이터로 구분하고 2012년~2021년을 검증 데이터로 정의했다. k -Fold 교차검증법($k = 10$)을 적용하여 검증하였으며 단기 물동량 예측력 평가를 위해 예측기간은 1년으로 설정했다. 머신러닝에 사용된 데이터는 월단위 물동량 데이터이며 예측시점 이후의 12개월을 예측하고 롤업 방식으로 연단위 물동량을 계산한다. 물동량 예측의 유형적 범위는 연안 출하 물동량이며 예측 정확도는 MAPE 기준으로 평가한다.

동해·묵호항의 연안 출하 물동량은 LSTM 모형이 7.5%로 가장 좋은 성능을 보였다. 삼척항과 옥계항은 XG부스트가 각각 8.2%, 3.6%로 가장 높은 예측력을 보여줬다. 종합적으로는 XG부스트가 평균 6.5%로 단기 물동량

예측에 가장 좋은 성능을 보여줬다. 기존 방식 또한 7.6%로 준수한 결과를 보여줬으며 LSTM 모형이 8.8%로 가장 낮은 성능을 보였다. LSTM은 삼척항 예측 시 일부 예측시점에서 예측력이 급격하게 떨어지는 경향을 보였는데, 이는 예측 시점 급격한 삼척항 물동량 변화와 모형 최적화에 대한 문제로 판단된다. 따라서 이를 개선하기 위해서는 당시 시멘트 제조사 물동량 변화 요인 및 모형 최적화에 대한 추가적인 연구가 필요하다.

〈표 4-30〉 모형별 연안 출하 물동량 예측 결과

단위: %

대상 항만	기존 방식	ARIMA	XG부스트	LSTM
동해·목호	8.2	9.8	7.6	7.5
삼척	9.3	11.1	8.2	13.7
옥계	5.3	5.1	3.6	5.1
평균	7.6	8.7	6.5	8.8

자료: 저자 작성

3. 지역별 지표 기반 입하 물동량 분배 규칙 개선

기존 모형의 연안 입하 물동량 분배는 시멘트 입하항인 인천, 부산, 광양, 마산 등 14개 항만의 전년도 비중으로 물동량을 예측한다. 이러한 방법론은 전년도의 연안 입하 비중이 유지될 것으로 가정한다.

〈표 4-31〉 국내 주요 시멘트 제조사별 입하항

제조사	출하항	주요 입하항
쌍용C&E	동해·목호	광양, 군산, 목포, 부산, 울산, 인천, 제주, 포항
삼표시멘트	삼척	광양, 군산, 목포, 부산, 인천
한라라파즈	옥계	광양, 마산, 목포, 울산, 인천, 포항

자료: 저자 작성

반면 개선 모형은 지역별 건설수주액과 인구수를 활용하여 연안 물동량 예측치를 도출하고 이동평균법으로 연안 입하 물동량을 보정한다. 입하 물동량 분배 문제는 추가 변수의 활용이 가능하고 2001년~2021년에 꾸준한 물동량이 발생한 10개 항만(인천항, 부산항, 광양항 등)을 대상으로 한다.

기존 모형의 경우 예측 시점에서 전년도 연안 입하 물동량 비중을 사용한 결과이다. 기존 모형을 항만별로 분석한 결과 장항항 62.2%, 광양항 35.2%, 인천항 33.2%, 울산항 31.2% 등 실제 연안 입하량 대비 오차율이 크게 나타났다. 따라서 개선 모형은 예측변수로 활용 가능한 지역별 건설수주액과 지역별 인구수를 사용하여 각 항만의 연안 입하 물동량을 독립적으로 예측했다. ‘건설수주액/인구수’ 모형은 광양항, 인천항, 장항항 및 포항항에서 큰 개선점을 보였다. 또한 ‘건설수주액/인구수’ 모형에 SMA($k = 2$)로 보정했을 시에 부산항을 제외한 모든 입하항에서 기존 모형 대비 MAPE가 개선되었다.

〈표 4-32〉 연안 입항 물동량 예측모형 결과 비교

단위: %

입하항	기존 모형	건설수주액/ 인구수	건설수주액/ 인구수+SMA(2)	차이
광양	35.2	19.1	15.8	19.4
군산	12.1	14.8	10.9	1.2
마산	12.0	15.8	11.3	0.7
목포	17.0	11.1	11.7	5.3
부산	12.0	18.5	13.3	-1.3
울산	31.2	13.5	10.4	20.8
인천	33.2	13.8	12.6	20.6
장항	62.2	14.3	14.4	47.8
제주	25.0	29.2	21.9	3.1
포항	30.9	14.4	10.5	20.4
전체	27.8	16.4	13.3	14.5

자료: 저자 작성

제5절 소결 및 시사점

본 장에서는 앞선 3장에서 제시한 품목별 항만물동량 예측모형 개선 방향에 따라 총 10가지의 고도화 방안을 도출하였다. 각 방안은 실증분석을 통해 구체화되었으며, 교차검증으로 기존 모형 대비 향상된 성능을 평가하였다. 자동차의 경우 ANFIS 모형을 도입하여 국내 자동차 수요량을 예측할 시 기존 모형 대비 2.7%p의 성능 개선이 이뤄졌으며, VAR 모형으로 수입자동차 수요를 예측했을 시 기존 모형 대비 예측 정확도가 5.5%p 개선되는 것으로 나타났다.

석탄은 상관분석을 통해 기존 조강생산능력 외에 추가로 한국은행 정책금리, 실질거래 환율 등 신규 설명변수를 발굴하였으며, 신규 설명변수를 투입한 MLR 모형을 활용하여 제철용 유연탄 수입물동량 예측모형의 성능을 기존 대비 1.3%p 개선하였다. 고철은 국내 고철 수요와 공급을 예측하는 모형을 각각 개발하여 수급분석을 실시하고, 이를 반영한 고철 수입물동량 예측모형 개발을 통해 기존 모형 대비 예측 정확도를 13.9%p 개선하였다.

시멘트의 경우는 상관분석을 통해 도시인구 증가율, 건설업 명목 GDP 등 신규변수 발굴하고 이를 기존 시멘트 수요 예측모형에 투입하여 예측 정확도를 5.4%p 개선하였다. XG부스트 모형 도입 시 주요 출하항의 단기 물동량 예측 정확도가 1.1%p 개선되는 것으로 나타났으며, 지역별 인구수 및 SMA 보정을 반영할 시 연안 입하 물동량 예측모형의 성능이 14.5%p 개선되었다. 다음의 <표 4-33>은 품목별 예측모형 고도화 방안 및 개선 결과를 요약하고 있다.

본 장에서 10가지 고도화 방안에 대해 실증분석을 실시한 결과를 종합하면, 기존 예측은 신규 설명변수를 발굴하여 반영하거나 다양한 시계열 또는 머신

러닝 예측모형들을 도입할 시 예측 정확도가 개선되는 것으로 나타났다. 이는 기존 모형을 즉시 신규 모형으로 대체해야함을 의미하지는 않지만, 항만물동량 데이터가 축적되고 영향 요인들이 변동됨에 따라 지속적으로 예측모형을 점검하고 신규 모형으로의 교체 가능성을 판단할 필요가 있음을 시사한다.

〈표 4-33〉 품목별 예측모형 고도화 방안 및 개선 효과

품목	고도화 방안	개선 효과
자동차	• 국내 자동차 수요량 예측 정확도 개선	• ANFIS 모형 도입 시 기존 모형 대비 예측 정확도 2.7%p 개선
	• 수입자동차 수요 예측모형 개발	• VAR 모형 도입 시 기존 모형 대비 예측 정확도 5.5%p 개선
석탄	• 석탄 수요예측을 위한 신규 변수 발굴	• 상관분석을 통해 중앙은행 정책금리, 실질거래 환율 등 신규 변수 발굴
	• 제철용 유연탄 수입물동량 예측모형 다각화	• MLR 모형 도입으로 기존 모형 대비 예측 정확도 1.3%p 개선
고철	• 국내 고철 수요 예측모형 개발	• 로그 추세모형으로 국내 고철 수요 예측 모형 개발
	• 국산 고철 공급량 예측모형 개발	• 고철 발생 방식에 따라 특화된 국산 고철 공급량 예측모형 개발
	• 수급분석을 반영한 고철 수입물동량 예측모형 개발	• 수급분석 반영 모형 개발을 통해 예측 정확도 4.9%p 개선
시멘트	• 시멘트 수요예측을 위한 신규 변수 발굴	• 도시 인구 증가율, 건설업 명목 GDP 등 신규변수 발굴을 통해 기존 모형 대비 예측 정확도 5.4%p 개선
	• 주요 출하항의 단기 물동량 예측 정밀화	• XG부스트 모형 도입 시 기존 모형 대비 예측 정확도 1.1%p 개선
	• 지역별 지표 기반 입하 물동량 분배 규칙 개선	• 지역별 인구수 반영 및 SMA 보정 시 기존 모형 대비 예측 정확도 14.5%p 개선

자료: 저자 작성

05

결론 및 정책제언

제1절 요약 및 결론

본 연구는 자동차, 석탄, 고철, 시멘트 등 비컨테이너 품목별 항만물동량 예측에 활용되는 기존 모형의 특징 및 한계점을 진단하고, 이를 토대로 예측모형 고도화 방안을 제시하는 것을 목적으로 한다. 궁극적으로는 예측모형 고도화를 통해 예측결과의 대내외적 신뢰도를 제고하고, 중장기 항만개발계획 수립 시 예측력이 높은 물동량 전망치 제공을 위한 여건을 조성하고자 한다. 이를 위해 본 연구에서는 광범위한 국내외 선행연구 분석을 통해 주요 수요예측 방법론을 탐색하고 특성을 분석하였으며, 실증분석을 통해 신규 모형 발굴 및 예측력 검정을 실시하였다. 또한 최종적으로 제안된 고도화 방안은 품목별 전문가 자문과 연구진 브레인스토밍을 통해 수정·보완되었다.

본 연구는 총 다섯 개의 장으로 구성되었으며, 제1장은 본 연구의 배경과 목적, 연구의 내용 및 방법, 선행연구 검토 및 연구의 차별성 등 연구의 개요를 기술하였다. 제2장은 다양한 예측 방법론과 예측모형의 적합도를 평가할 수 있는 지표 및 검증법을 살펴보았다. 수요예측 연구에서 주로 활용되는 시계열 모형을 탐색적 분석기법과 확률적 분석기법으로 구분하여

이동평균 평활법과 ARIMA, SARIMA, VAR 모형 등을 검토하였으며, 회귀 모형으로는 단순회귀 모형과 다중회귀 모형을 검토하였다. 이어서 최근에 활용 빈도가 높은 머신러닝 기반 모형 중 ANN, SVM, LSTM, XG부스트, ANFIS 등을 살펴보았다. 또한 예측모형 개발에 있어서 제안한 모형의 성능을 측정하는 지표로 RMSE, MAPE 등을 검토하였으며, 모형 간 지표의 추정치를 비교 검증하기 위해 k -Fold 교차검증법과 몬테카를로 교차검증법의 활용을 제안하였다.

이어서 제3장에서는 신규 모형 개발에 앞서 기존 품목별 항만물동량 예측모형의 특징 및 한계점을 진단하였다. 품목별 연관 산업의 동향 및 주요 영향요인을 분석하고, 기존 항만물동량 예측에 활용된 변수와 모형을 검토하여 기존 예측모형의 특징과 한계점을 도출하였다. 이를 통해 품목별 항만물동량 예측모형 개선 방향을 제시하였으며, 이는 제4장에서 총 10가지 고도화 방안으로 도출되었다. 실증분석을 통해 개선이 실질적으로 어느 정도의 효과를 가지는지 계량적으로 측정되었으며, 특히 k -Fold 교차검증을 통해 표본 외 검정 시 모형의 성능을 평가하고 기존 모형 대비 신규 모형의 우수성을 증명하였다. 자동차의 경우 국내 자동차 수요량 예측모형과 수입 자동차 수요 예측모형을 개선하였으며, 석탄은 상관분석을 통해 기존 조강 생산능력 외 신규 설명변수 발굴 및 제철용 유연탄 수입물동량 예측모형의 성능 개선이 이뤄졌다. 고철은 국내 고철 수요와 공급을 예측하는 모형을 각각 개발하여 수급분석을 실시하고, 이를 반영한 고철 수입물동량 예측모형을 제시하였다. 시멘트는 신규변수 발굴과 주요 출하항의 단기 물동량 예측 정밀화 및 지역별 지표 기반 입하 물동량 분배 규칙을 개선하였다.

한편 본 연구에서 제시한 10가지 고도화 방안 외에도 기존 모형 개선에는 다양한 방법론들이 적용될 수 있을 것이다. 현실적으로 모든 방법론과 가용한 모든 변수들을 검토하기에는 많은 제약이 있을 수밖에 없다. 또한 예측 정확도가 일부 개선된다고 해서 기존 모형을 신규 모형으로 즉시 교체

해야함을 의미하지도 않는다. 하지만 항만물동량 예측모형의 신뢰도 제고를 위해서는 예측 프로세스를 체계화하고 지속적인 모형의 재검토 및 개선 방안 도출이 중요하다. 이를 위해 상시적 수요예측 고도화 사업 실시, 수요검토 위원회 확대 개편, 단기 물동량 예측모형 개발 및 특정 이벤트 발생 시 물동량 영향 분석 기반 마련 등이 필요할 것으로 판단된다.

제2절 정책제언

1. 상시적 항만물동량 예측모형 고도화 체계 구축

정부는 항만의 개발 및 운영계획 등 주요 항만 관련 정책 수립의 근거로 항만물동량에 대한 중장기적 예측을 활용하고 있다. ‘전국 무역항 기본계획’이 10년마다 수립되고 있으며, 5년마다 수정계획을 고시하고 있다. 이에 항만물동량 예측의 목표연도는 예측 시점부터 30년까지로 초장기 예측이 요구된다. 이에 비해 항만물동량이 체계적으로 집계되기 시작한 시점은 25년이 채 되지 않았다. 예측 기간에 비해 상대적으로 짧은 시계열 자료를 바탕으로 추정된 미래 시점의 물동량 전망은 근본적으로 추정치의 변동성이 클 수밖에 없다. 특히 최근 대내외 물류 환경이 급변하면서 향후 수년 내의 물동량 전망도 예측이 정확하기를 기대하기 어렵다.

방법론적인 측면에서도 여러 후보 모형 중에서 최적의 모형을 선별하는 과정은 일률적이지 않은 경우가 많다. 모형 선택에 앞서 모형의 우수성 평가 기준을 확정해야 하며, 여기에는 정량적 평가와 더불어 정성적 평가 기준이 포함되기도 한다. 또한 모든 기준에서 다른 모형 대비 상대적 우위를

차지하는 모형이 존재하지 않을 수도 있으며, 일부 조건이 변경되거나 평가 기준이 변경되면 현시점에서 가장 우수한 모형으로 평가받던 모형이 상대적 열위에 놓이기도 한다. 기존에 없던 새로운 데이터가 수집되거나 기존 데이터의 차기 시점 데이터가 업데이트되면 기존 모형에 대한 평가가 바뀔 수 있고, 신규 예측 방법론 도입을 통해 비교 집단이 바뀌면서 최적의 모형이 변경될 수도 있다.

이에 본 연구에서는 항만 수요예측 고도화 사업을 상시적·전문적으로 추진할 수 있는 체계 구축을 제안한다. 고도화 사업의 내용적 범위는 기존 예측모형의 검증 및 신규 모형 개발을 포함하며, 이를 위해 방대한 양의 데이터를 효율적으로 수집·분석할 수 있는 시스템과 통계분석 툴 및 전문 인력 확보가 뒤따라야 할 것으로 판단된다. 또한 개발한 예측모형에 대한 체계적인 평가 시스템 마련을 위한 연구가 필요하다.

2. 예측모형 확장을 통한 데이터 기반 정책 개발 지원

정부는 최근 데이터 3법 개정 등 관련 제도 개선을 통해 개인과 기업이 정보를 활용할 수 있는 폭을 넓혔으며, 각 부처에서는 공공데이터를 선도적으로 개방하여 데이터의 공유 및 활용을 활성화하는 정책을 추진 중이다. 이처럼 정부와 민간을 통해 과거에는 불가능했던 데이터의 확보가 가능해지면서 데이터의 중요성이 더욱 강조되고 있으며, 이에 따라 데이터 기반 정책 개발 연구의 수요도 급격히 증가하고 있다.

또한 본 연구를 진행하면서 수행한 산·학·연·정 전문가 자문을 통해 살펴본 바에 따르면 항만 분야에서도 항만물동량 예측 수요가 증가하고 있음을 확인할 수 있었다. 항만 분야 대형 신규 공공투자사업에 대한 예비타당성 조사, 민간 제안사업의 적격성 검토 등에서 항만물동량 예측 자료가

수시로 요구되고 있으며, 코로나19에 따른 항만 봉쇄, 러시아·우크라이나 전쟁 등 물류 공급망 충격 발생 시 항만물동량 영향에 대한 관심이 높아지고 있다.

이처럼 증가하는 항만물동량 예측 관련 연구 수요에 대응하여 정책담당자, 항만당국, 부두 운영사 등이 다양한 목적으로 활용할 수 있는 기초자료를 생산하기 위해서는 기존 중장기 물동량 예측모형 고도화뿐만 아니라 단기적 관점에서의 물동량 예측 및 외부 충격에 따른 영향 추정 모형 개발이 필요하다. 월별 혹은 분기별 자료를 이용한 단기 물동량 예측에는 본 연구에서 소개한 다양한 머신러닝 기반의 예측모형의 활용도가 높을 것으로 예상되며, 공급망 충격에 따른 항만물동량 영향 예측에는 구조방정식 모형, 시스템다이내믹스 등과 같이 시나리오 기반 분석이 가능한 모형이 적합할 것으로 판단된다.

3. 정부-연구기관 간 데이터 공유 체계 구축

한국해양수산개발원 항만수요예측센터는 지난 2010년부터 ‘항만개발 시기 및 규모 산정에 대한 조사·연구기관’으로 선정되어 지속적으로 국내외 항만의 물동량을 모니터링하고 지속적으로 국내 항만물동량 예측 관련 연구 사업을 수행하고 있다. 사업의 일환으로 매년 『품목별 항만물동량 예측보고서』를 발간하고 있으며, 사업의 결과물은 정부 정책의 근거자료로 활용되고 있다.

항만수요예측센터의 품목별 항만물동량 예측시스템의 근간이 되는 물동량 데이터는 해양수산부의 Port-MIS로부터 수집된다. Port-MIS는 전국 무역항을 대상으로 선박의 입출항, 항만시설, 화물반출입 등 항만운영과 관련된 정보를 처리하는 시스템이다. 현재는 Port-MIS의 방대한 양의 데이터 중

에서 월별로 집계된 항만별 물동량만이 항만수요예측센터에 제공되어 물동량 예측에 활용되고 있다. 이에 Port-MIS와 항만수요예측센터 간 데이터 공유 체계 구축을 통해 고빈도(high-frequency) 대용량 데이터의 활용도를 제고할 필요가 있다. Open-API 등의 방식으로 공유된 물동량 정보를 항만수요예측센터에서 지속적으로 모니터링하고 시의 적절하게 분석결과를 정부에 제공함으로써 공공데이터의 활용도를 높이는 한편 정책 연구의 품질 향상도 자연스레 뒤따를 것으로 기대된다.

참고문헌

국내 문헌

- 김경태 외(2014), 벌크시멘트 철도물류 프로세스 개선 방안, 한국물류학회, 23(1), pp. 147-169.
- 김동윤 외(2021), MLP 를 이용한 공컨테이너 수요예측, 한국빅데이터학회지, 6(2), pp. 85-98.
- 김두환 · 이강배(2020), LSTM 을 활용한 부산항 컨테이너 물동량 예측, 한국항만경제학회지, 36(2), pp. 53-62.
- 김정은 외(2018), 유류화물 항만물동량 예측모형 개발 연구, 한국항만경제학회지, 34(1), pp. 81-98.
- 민경창 · 하헌구(2014), SARIMA 모형을 이용한 우리나라 컨테이너 물동량 예측, 대한교통학회지, 32(6), pp. 65-82.
- 박상수 외(2017), 한국 자동차산업의 수요구조 및 정책효과 분석 -사회후생 및 국민경제 파급효과 분석을 중심으로-, 산업연구원.
- 산업통상자원부(2020), 제9차 전력수급기본계획(2020~2034).
- 서홍용(2014), 우리나라 연안해운 물류체계의 공간적 네트워크 분석 : 시멘트를 중심으로, 지리학논총, 59/60, pp. 141-169.
- 신창훈 외(2008), 하이브리드 ARIMA-신경망 모델을 통한 컨테이너물동량 예측에 관한 연구, 한국항해항만학회지, 32(2), pp. 81-88.
- 신창훈 · 정수현(2011), 항만물동량 예측력 제고를 위한 ARIMA 및 인공신경망모형들의 비교 연구, 한국항해항만학회지, 35(1), pp. 83-91.

-
- 에너지경제연구원(2021a), 에너지전환 정책의 성과 및 향후 추진방향 연구.
- _____ (2021b), 제9차 전력수급기본계획의 주요 내용과 향후 과제, 에너지포커스 2021 봄호(제18권 제1호 통권79호), pp. 23-27.
- _____ (2022), IEA의 최근 세계 석탄 수급 분석과 중기 전망, 세계 에너지 시장 인사이트, 제22-3호, pp. 3-17.
- 유승훈(2021), 제9차 전력수급기본계획의 주요 내용과 향후 과제, 에너지포커스 2021 봄호(제18권 제1호 통권79호), 에너지경제연구원, pp. 23-27.
- 이민규 외(2014), 산업연관분석을 활용한 수출입 항만물동량 유발효과의 계산, 해운물류연구, 30(3), pp. 617-635.
- 이재길(2017), R프로그램에 기반한 시계열 자료 분석, 황소걸음 아카데미.
- 인천항만공사(2010), 인천남항 석탄부두 전용사용료 산정 연구.
- 장원익(2010), 광물자원의 해상물동량 전망에 관한 연구: 철광석 및 석탄을 중심으로, 자원환경경제연구, 19(2), pp. 341-360.
- 전찬영 · 송주미(2006), 데이터 마이닝 기법을 이용한 항만물동량 예측 활용방안 연구, 한국해양수산개발원.
- 전찬영 · 송주미(2007), 인공신경망모형의 항만물동량 예측 적용에 관한 연구, 해운물류연구, (53), pp. 65-82.
- 조상호 외(2020), 단계적 회귀분석과 인공신경망 모형을 이용한 광양항 석탄 · 철광석 물동량 예측력 비교 분석, 한국항해항만학회지, 44(3), pp. 187-194.
- 조찬혁(2014), 동해항 시멘트 물동량의 추정에 관한 연구, 한국물류학회, 18(1), pp. 33-53.
- 하준수 외(2021), Datamining 기법을 활용한 단기 항만 물동량 예측, 한국항만경제학회지, 37(3), pp. 01-17.
- 하준수 외(2021), 시계열 분석 기반 신뢰구간 추정을 활용한 항만물동량 이상감지 방안, 한국항만경제학회지, 37(1), pp. 179-196.
- 한국광해광업공단(2022), 월간자원 인사이트(제22-05호), pp. 3-5.
- 한국시멘트협회(2021a), 2021 한국의 시멘트산업 통계.
- _____ (2021b), 2021년 시멘트지(제226호).
-

- 한국자동차산업협회(2021a), 2021년 자동차산업 평가 및 2022년 전망, pp. 2-23.
- _____ (2021b), 2020년 기준 세계 자동차 통계(종합편), pp. 1-329.
- _____ (2022a), 2021년 세계 자동차 생산 현황과 시사점, pp. 1-7.
- _____ (2022b), 2021년 해외 주요 자동차시장 판매 및 정책동향, pp. 1-15.
- _____ (2022c), 2021년 주요국 전기동력차 보급현황 분석, pp. 1-25.
- 한국해양수산개발원(2022), 2021년 품목별 항만물동량 예측보고서.

국외 문헌

- Andrikopoulos, A., Merika, A., Merikas, A. & Tsionas, M.(2020), The dynamics of fleet size and shipping profitability: The role of steel-scrap prices, *Maritime Policy & Management*, 47(8), pp. 985-1009.
- Benalcazar, P., Krawczyk, M. & Kamiński, J.(2017), Forecasting global coal consumption: An artificial neural network approach, *Gospodarka Surowcami Mineralnymi*, 33(4), pp. 29-44.
- Boser, B. E., Guyon, I. M., & Vapnik, V. N.(1992, July). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pp. 144-152.
- Chan, H. L. & Lee, S. K.(1997), Modelling and forecasting the demand for coal in China, *Energy Economics*, 19(3), pp. 271-287.
- Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785-794.
- Ding, M. J., Zhang, S. Z., Zhong, H. D., Wu, Y. H., & Zhang, L. B. (2019). A prediction model of the sum of container based on combined BP neural network and SVM. *Journal of Information Processing Systems*, 15(2), pp. 305-319.
- Dragan, D., Keshavarzsaleh, A., Intihar, M., Popović, V. & Kramberger, T.(2021), Throughput forecasting of different types of cargo in the adriatic seaport Koper, *Maritime Policy & Management*, 48(1), pp. 19-45.

-
- Dworak, S. & Fellner, J.(2021), Steel scrap generation in the EU-28 since 1946–Sources and composition, *Resources, Conservation and Recycling*, 173, 105692.
- Farhan, J. & Ong, G. P.(2018), Forecasting seasonal container throughput at international ports using SARIMA models, *Maritime Economics & Logistics*, 20(1), pp. 131–148.
- Fışkın, C. S. & Cerit, A. G.(2019), Forecasting domestic shipping demand of cement: Comparison of SARIMAX, ANN and hybrid SARIMAX-ANN, 2019 4th International Conference on Computer Science and Engineering, pp. 68–72.
- Fonseca-Delgado, R. & Gómez-Gil, P.(2013), An assessment of ten-fold and Monte Carlo cross validations for time series forecasting, 2013 10th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), pp. 215–220, doi: 10.1109/ICEEE.2013.6676075.
- Fradinata, E., Suthummanon, S. & Suntiamorntut, W.(2015), Forecasting determinant of cement demand in Indonesia with artificial neural network, *Journal of Asian Scientific Research*, 5(7), pp. 373–384.
- Fradinata, E., Suthummanon, S., Sirivongpaisal, N. & Suntiamorntut, W.(2014), ANN, ARIMA and MA timeseries model for forecasting in cement manufacturing industry: Case study at Lafarge cement Indonesia-Aceh, 2014 International Conference of Advanced Informatics: Concept, Theory and Application, pp. 39–44.
- Gao, Y., Luo, M. & Zou, G.(2016), Forecasting with model selection or model averaging: A case study for monthly container port throughput, *Transportmetrica A: Transport Science*, 12(4), pp. 366–384.
- Gosasang, V., Chandraprakaikul, W., & Kiattisin, S. (2011), A comparison of traditional and neural networks forecasting techniques for container throughput at Bangkok port, *The Asian Journal of Shipping and Logistics*, 27(3), pp. 463–482.
- Gosasang, V., Yip, T. L., & Chandraprakaikul, W. (2018), Long-term container throughput forecast and equipment planning: the case of Bangkok Port. *Maritime Business Review*.
- Hatayama, H., Daigo, I., Matsuno, Y. & Adachi, Y.(2010), Outlook of the world steel cycle based on the stock and flow dynamics, *Environmental Science & Technology*, 44(16), pp. 6457–6463.

- Hebb, D. O. (1949). The first stage of perception: growth of the assembly. *The Organization of Behavior*, 4, pp. 60-78.
- Hochreiter, S., & Schmidhuber, J. (1997), Long short-term memory, *Neural computation*, 9(8), pp. 1735-1780.
- Islam, R., Ghani, A. B. A., Kusuma, B. & Hong, E. T. Y. (2016), An analysis of factors that affecting the number of car sales in Malaysia, *International Review of Management and Marketing*, 6(4), pp. 872-882.
- Jiang, L., Wang, J., Jiang, H. & Feng, X. (2019), Prediction model of port throughput based on game theory and multimedia Bayesian regression, *Multimedia Tools and Applications*, 78(4), pp. 4397-4416.
- Lam, W. H., Ng, P. L., Seabrooke, W. & Hui, E. C. (2004), Forecasts and reliability analysis of port cargo throughput in Hong Kong, *Journal of Urban Planning and Development*, 130(3), pp. 133-144.
- Lapedes, A., & Farber, R. (1987), Nonlinear signal processing using neural networks: Prediction and system modelling (No. LA-UR-87-2662; CONF-8706130-4).
- Lee, H. & Sohn, I. (2015), Global scrap trading outlook analysis for steel sustainability, *Journal of Sustainable Metallurgy*, 1(1), pp. 39-52.
- Li, M. W., Geng, J., Hong, W. C., & Chen, Z. Y. (2017), A novel approach based on the Gauss-vSVR with a new hybrid evolutionary algorithm and input vector decision method for port throughput forecasting, *Neural Computing and Applications*, 28(1), pp. 621-640.
- Li, Y. & Li, Z. (2019), Forecasting of coal demand in China based on support vector machine optimized by the improved gravitational search algorithm, *Energies*, 12(12), 2249.
- Liu, P., Chen, S., Yang, H., Hung, C. & Tsai M. (2008), Application of artificial neural network and SARIMA in portland cement supply chain to forecast demand, 2008 Fourth International Conference on Natural Computation, pp. 97-101.
- Niu, M., Hu, Y., Sun, S. & Liu, Y. (2018), A novel hybrid decomposition-ensemble model based on VMD and HGWO for container throughput forecasting, *Applied Mathematical Modelling*, 57, pp. 163-178.

-
- McCulloch, W. S., & Pitts, W. (1943), A logical calculus of the ideas immanent in nervous activity, *The bulletin of mathematical biophysics*, 5(4), pp. 115-133.
- Mitchell, T. M., & Mitchell, T. M. (1997). *Machine learning* (Vol. 1, No. 9). New York: McGraw-hill.
- Oda, J., Akimoto, K. & Tomoda, T.(2013), Long-term global availability of steel scrap, *Resources, Conservation and Recycling*, 81, pp. 81-91.
- Pauliuk, S., Milford, R. L., Müller, D. B. & Allwood, J. M.(2013), The steel scrap age, *Environmental Science & Technology*, 47(7), pp. 3448-3454.
- Pae, D. S., Choi, I. H., Kang, T. K., & Lim, M. T. (2018), Vehicle detection framework for challenging lighting driving environment based on feature fusion method using adaptive neuro-fuzzy inference system, *International Journal of Advanced Robotic Systems*, 15(2), 1729881418770545.
- Rosenblatt, F. (1958), The perceptron: a probabilistic model for information storage and organization in the brain, *Psychological review*, 65(6), 386.
- Ruiz-Aguilar, J. J., Moscoso-Lopez, J. A., Urda, D., Gonzalez-Enrique, J., & Turias, I. (2020), A clustering-based hybrid support vector regression model to predict container volume at seaport sanitary facilities, *Applied Sciences*, 10(23), 8326.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986), Learning representations by back-propagating errors. *nature*, 323(6088), pp. 533-536.
- Summaries, M. C. (2022), *Mineral Commodity Summaries: USGS Numbered Series 2022*, US Geological Survey: Reston, VA, pp. 45.
- Vahabi, A., Hosseininia, S. S. & Alborzi, M.(2016), A sales forecasting model in automotive industry using adaptive neuro-fuzzy inference system (Anfis) and genetic algorithm (GA), *International Journal of Advanced Computer Science and Applications*, 7(11), pp. 24-30.
- van Dorsser, C., Wolters, M., & van Wee, B. (2012), A very long term forecast of the port throughput in the Le Havre–Hamburg range up to 2100, *European Journal of Transport and Infrastructure Research*, 12(1).

- Xuan, Y. & Yue, Q.(2016), Forecast of steel demand and the availability of depreciated steel scrap in China, Resources, Conservation and Recycling, 109, pp. 1-12.
- Zhang, C., Huang, L. & Zhao, Z.(2013), Research on combination forecast of port cargo throughput based on time series and causality analysis, Journal of Industrial Engineering and Management, 6(1), pp. 124-134.
- Zhang, Y., Zhong, M., Geng, N. & Jiang, Y.(2017), Forecasting electric vehicles sales with univariate and multivariate time series models: The case of China, PloS One, 12(5), e0176729.

인터넷 자료

- 그린피스, <https://www.greenpeace.org/korea/press/12618/boom-and-bust-2020/> (검색일: 2022.6.1.)
- 뉴스시스, https://mobile.newsis.com/view.html?ar_id=NISX20211114_0001650117 (검색일: 2022.6.1.)
- 에너지정책소통센터, <https://e-policy.or.kr/policy/03.php>(검색일: 2022.6.5.)
- 한국무역통계진흥원, HS CODE 내비게이션, https://www.bandtrass.or.kr/hsnavi.do?page=F&site=#table_content(검색일: 2022.6.5.)
- 한국철강협회, STEELDATA, https://sd.kosa.or.kr/actions/Front.action?pageId=C012200&_flush_ (검색일: 2022.9.12.)
- 한국표준협회, <https://standard.go.kr/KSCI/standardIntro/getStandardSearchView.do?menuId=919&topMenuId=502&upperMenuId=503&ksNo=KSL5201&tmpRksNo=KSL5201&reformNo=24>(검색일: 2022.5.1.)
- 해양수산부 Port-MIS, <https://new.portmis.go.kr/portmis/websquare/websquare.jsp?w2xPath=/portmis/w2/main/index.xml&page=/portmis/w2/cm/sys/UI-PM-MT-001-021.xml&menuId=0045&menuCd=M4735&menuNm=BB%E7%C0%CC%C6%AE%B8%CA>(검색일: 2022.9.12.)

IHS Markit, <https://connect.ihsmarkit.com/data-browser/buildQuery/Economics>
(검색일: 2022.9.12.)

KESIS 국가에너지통계종합정보시스템, <http://www.kesis.net/main/main.jsp>
(검색일: 2022.6.5.)

Olah, C. (2015), Understanding lstm networks, <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>(검색일: 2022.6.15)

Singh, R. (2020), Mathematics behind Random forest and XGBoost, <https://medium.com/analytics-vidhya/mathematics-behind-random-forest-and-xgboost-ea8596657275>(검색일: 2022.7.20)

부 록

선행연구 현황

■ 국내연구 사례

NO	연구 목적	연구 방법	주요 연구 내용
1	- 과제명: 데이터 마이닝 기법을 이용한 항만물동량 예측 활용 방안 연구 - 연구자(년도): 전찬영 · 송주미 (2006) - 연구목적: 인공신경망모형을 활용한 단기 품목별 항만물동량 예측	- 국내외 문헌자료 분석 - 인공신경망(Artificial Neural Network) 모형 - 몬테카를로 시뮬레이션	- 데이터마이닝 기법 고찰 - 주요 물동량 추세 변화 분석 및 설명변수 설정 - 품목별 예측의 적합도 비교분석 - 몬테카를로 시뮬레이션 분석
2	- 과제명: Datamining 기법을 활용한 단기 항만 물동량 예측 - 연구자(년도): 하준수 외 (2021) - 연구목적: 데이터마이닝 기법을 활용한 단기 물동량 예측 모형 제시	- 국내외 선행연구 분석 - CART(Classification and Regression Tree) 모형 - ARIMA 모형	- ARIMA 및 CART 모형을 활용한 two stage 단기 물동량 예측 모형 개발 - 실증분석 및 예측 정확도 분석
3	- 과제명: 산업연관분석을 활용한 수출입 항만물동량 유발효과 분석 - 연구자(년도): 이민규 외 (2014) - 연구목적: 산업연관분석 방법을 이용하여 국내 항만의 수출입 항만물동량 유발효과를 분석	산업연관분석	- 산업연관표와 SPIDC 화물품목 매칭을 통해 28개 산업부문 재분류 - 수출입 항만물동량 유발효과 분석

NO	연구 목적	연구 방법	주요 연구 내용
4	- 과제명: 인공신경망모형의 항만 물동량 예측 적용에 관한 연구 - 연구자(년도): 전찬영 · 송주미 (2007) - 연구목적: 인공신경망모형을 활용한 단기 품목별 항만물동량 예측	- 국내외 문헌자료 분석 - 인공신경망 모형 - 몬테카를로 시뮬레이션	- 데이터마이닝기법 고찰 - 주요 물동량 추세 변화 분석 및 설명변수 설정 - 품목별 예측의 적합도 비교분석 - 몬테카를로 시뮬레이션 분석
5	- 과제명: SARIMA 모형을 이용한 우리나라 컨테이너 물동량 예측 - 연구자(년도): 민경창 · 하현구 (2014) - 연구목적: 분기별 항만컨테이너 물동량 자료를 활용하여 단기 수요를 예측	SARIMA(Seasonal Autoregressive Integrated Moving Average)모형	- 분기별 항만 컨테이너 물동량 예측 - ARIMA모형과 SARIMA모형 간 예측결과 비교
6	- 과제명: 시계열 분석 기반 신뢰 구간 추정을 활용한 항만물동량 이상감지 방안 - 연구자(년도): 하준수(2021) - 연구목적: 수요예측치의 신뢰 구간 추정을 활용한 항만 물동량 이상감지 방안 제안	- 항만수요예측 관련 선행연구 검토 - ARIMA 모형	- 일별, 부두별 물동량 자료를 활용한 단기 물동량 예측 - 예측치를 토대로 부두 장치물 관리 방안 제시
7	- 과제명: 항만물동량 예측력 제고를 위한 ARIMA 및 인공신경망모형들의 비교 연구 - 연구자(년도): 신창훈(2011) - 연구목적: ARIMA모형과 인공신경망모형을 결합한 하이브리드모형을 활용하여 컨테이너 물동량 예측	- 국내외 문헌자료 분석 - 하이브리드모형(ARIMA&인공신경망) - 실증분석	ARIMA모형, 인공신경망모형, 하이브리드모형 간 컨테이너 물동량 예측치 비교 분석
8	- 과제명: LSTM을 활용한 부산항 컨테이너 물동량 예측 - 연구자: 김두환 · 이강배 (2020) - 연구목적: 부산항 컨테이너물동량 예측을 위해 SARIMA와 LSTM 모형의 예측 정확도 비교	- SARIMA과 LSTM 모형 적용 및 비교 - RMSE와 MAPE를 활용해 예측 정확도 검증	- SARIMA와 LSTM을 활용한 부산항 컨테이너 물동량 예측 - 부산항의 14년간 월별 물동량 데이터 활용, 단기 (24개월) 물동량 예측 - SARIMA와 LSTM의 물동량 예측 결과를 비교 분석, LSTM의 예측력 검증

NO	연구 목적	연구 방법	주요 연구 내용
9	- 과제명: 한국 자동차산업의 수요구조 및 정책효과 분석 - 연구자(년도): 박상수 외 (2017) - 연구목적: 한국 승용차 시장의 수요와 공급 함수 추정 모형 설계 및 개별소비세 인하에 따른 가격인하 효과 추정	- 국내외 문헌자료 분석 - 확률적 계수 로짓모형 분석 - 정책효과 시뮬레이션 분석	- 한국 자동차산업의 현황 및 정책 동향 조사 - 한국 자동차산업의 수요 및 공급 함수 추정 - 승용차 개별소비세 인하의 사회후생효과 및 국민경제 파급 효과 분석
10	- 과제명: 광물자원이 해상물동량 전망에 관한 연구: 철광석 및 석탄을 중심으로 - 연구자(년도): 장원익(2010) - 연구목적: 조선경기 예측을 위해 건화물(철광석, 석탄)의 물동량을 전망	- 국내외 선행연구 검토 - 선형 회귀분석	- 철광석은 조강생산과 세계 GDP를 독립변수로 설정 - 석탄은 조강생산과 세계 발전량을 독립변수로 설정 - 선형 회귀분석을 통해 석탄과 철광석의 단기/중기 해상물동량 전망
11	- 과제명: 단계적 회귀분석과 인공신경망 모형을 이용한 광양항 석탄·철광석 물동량 예측력 비교 분석 - 연구자(년도): 조상호 외 (2020) - 연구목적: 항만의 주요정책 및 운영계획 수립을 위해 예측모형간 예측력 비교 분석	- 선행연구 검토 - 시계열 분석 - 단계적 회귀분석 - 인공신경망(Artificial Neural Network) 모형	- 석탄 및 철광석 물동량에 영향을 주는 요인 선정 - 단계적 회귀분석과 인공신경망모형간 물동량 예측력 비교
12	- 과제명: 유류화물 항만물동량 예측모형 개발 연구 - 연구자(년도): 김정은(2018) - 연구목적: 유류화물의 항만물동량 예측방식 비교 분석	- 국내외 선행연구 검토 - 실증분석(ARIMA, 윈터스 기법, 단순계절모형, TSMR 모형)	- ARIMA모형, 윈터스 기법모형, 단순계절모형, TSMR모형의 예측방식 비교 - 모형별 적합도, 예측력, 정확성 검증
13	- 과제명: 우리나라 연안해운 물류체계의 공간적 네트워크 분석-시멘트를 중심으로- - 연구자: 서홍용(2014) - 연구목적: 연안해운의 공간적 특성과 시멘트의 품목 특성을 고려한 공간적 네트워크 분석 연구	- 국내외 연안물류 체계에 대한 선행연구 검토 - 다이에딕 요인분석 - 네트워크 분석	- 전국 30개 무역항의 시멘트 물동량을 대상으로 다이에딕 요인분석과 사회 네트워크 분석 실시 - 항만별 반입·반출 중심성을 정의하고 물품 종류별 출발지-행선지의 유사성을 기반으로 군집화 실시

NO	연구 목적	연구 방법	주요 연구 내용
14	- 과제명: 벌크시멘트 철도물류 프로세스 개선 방안 - 연구자: 김경태 외(2014) - 연구목적: 시멘트 물류비를 줄이기 위해 물류 프로세스를 진단, 벌크시멘트 수송용 컨테이너 도입 검토	- 벌크시멘트 철도, 육로 및 해상 운송 프로세스 진단 - AS-IS/TO-BE 분석을 통한 물류 프로세스 변화 분석	철도운송용 컨테이너 도입에 따른 시멘트 물류 프로세스 변화와 효과 분석
15	- 과제명: 동해항 시멘트 물동량의 추정에 관한 연구 - 연구자: 조찬혁(2008) - 연구목적: 동해항의 시멘트 반출량 예측을 위해 Holt-Winter 지수평활방식 및 ARIMA 예측력 비교	- Holt-Winter 지수평활모형 적용 - ARIMA(Box-Jenkins 방식) 모형 적용	- 국내 최대 시멘트 반출항인 동해항을 대상으로 시멘트 물동량 추정 - 동해항의 12년간 월별 물동량 데이터를 IMF 이전, IMF 기간, IMF 이후로 구분하여 예측 - Holt-Winter 지수평활모형과 ARIMA 방식의 예측능력 비교

■ 국외연구 사례

NO	연구 목적	연구 방법	주요 연구 내용
1	- 과제명: Throughput forecasting of different types of cargo in the adriatic seaport Koper - 연구자(년도): Dragan et al. (2021) - 연구목적: 품목별 항만물동량 예측모형 개발	- 선행연구 분석 - Dynamic Factor Analysis-ARIMAX - Principal component analysis - Multiple linear regression - Monte Carlo simulation	54개의 거시경제변수를 PCA, MLR, MC를 활용해 압축하고, DFA를 활용해 동적 요인을 추출, 다양한 모수 값을 대입한 ARIMAX 모형을 통해 품목별 물동량 예측 2001~2017년 분기 데이터를 이용해 각 모형을 통계적 비교 검증
2	- 과제명: Forecasting seasonal container throughput at international ports using SARIMA models - 연구자(년도): Farhan&Ong (2018) - 연구목적: 계절성을 반영한 컨테이너 항만물동량 월별 단기 예측모형 개발	- 선행연구 분석 - SARIMA - DF-GLS test	- 세계 주요 20개 항만에 대해 계절성을 반영한 월별 물동량 예측모형 개발 1999~2007년 월별 데이터를 이용해 각 모형을 통계적 비교 검증

NO	연구 목적	연구 방법	주요 연구 내용
3	• 과제명: Forecasting with model selection or model averaging: a case study for monthly container port throughput • 연구자(년도): Gao et al. (2016) • 연구목적: 다양한 예측모형의 선택 방법 및 평균화 기법을 비교하여 가장 우수한 모형 선택 기준 파악	• Structural change VAR • Model selection methods • Model averaging methods	• 홍콩항과 센젠항의 1999~2014년 월별 컨테이너 물동량 데이터를 바탕으로 SC-VAR (m) 모형을 통해 예측 • 6개의 모형 선택 기법과 6개의 모형 평균화 기법을 적용하여 가장 우수한 기법 확인
4	• 과제명: Prediction model of port throughput based on game theory and multimedia Bayesian regression • 연구자(년도): Jiang et al. (2019) • 연구목적: 게임이론을 통해 배후단지가 겹치는 항만의 물동량 예측모형을 개발하고 실증 분석을 통해 검증함	• 선행연구 분석 • Game theory • Multimedia theory • Bayesian regression • Simulation analysis	• 확률예측이론과 도로 임피던스 함수를 도입해 베이지안 회귀 모형을 개발 • 중국 연안항만 1997~2017년 품목별 물동량을 이용해 실증
5	• 과제명: Forecasts and reliability analysis of port cargo throughput in Hong Kong • 연구자(년도): Lam et al. (2004) • 연구목적: 품목별 항만물동량 예측 모형 개발	• 선행연구 분석 • Neural networks model • Regression analysis • Monte Carlo method	• 1983~2000년 홍콩항의 37개 종류의 물동량 데이터를 통해 2002~2011년 물동량을 예측 • 수출입 물동량(금액), 인구, 전기 사용량, 건설 · 건축비, GDP 등의 설명변수를 사용하여 NN 적용
6	• 과제명: A novel hybrid decomposition-ensemble model based on VMD and HGWO for container throughput forecasting • 연구자(년도): Niu et al. (2018) • 연구목적: 컨테이너 물동량 예측 방법론 안정성 및 정확도 향상	• 선행연구 분석 • Variational mode decomposition • Support vector regression • Hybridizing grey wolf optimization • Hybrid decomposition-ensemble model • ARIMA	• VMD 알고리즘을 적용해 컨테이너 물동량 시계열 데이터를 요소별로 분류한 후, 주기가 긴 시계열 데이터는 ARIMA, 주기가 빠른 시계열 데이터는 SVR를 통해 예측 • 최종 결과는 앙상블 모형을 적용하여 도출

NO	연구 목적	연구 방법	주요 연구 내용
7	- 과제명: Research on combination forecast of port cargo throughput based on time series and causality analysis - 연구자(년도): Zhang et al. (2013) - 연구목적: 화물 물동량 시계열 예측 방법론 개발	- 선행연구 분석 - Combined model - Logistic growth curve model - Grey forecast model	Grey 모형과 logistic 모형을 혼합하여 예측모형 개발
8	- 과제명: Forecasting electric vehicles sales with univariate and multivariate time series models - 연구자(년도): Zhang et al. (2017) - 연구목적: 단변량 및 다변량 시계열 모형을 활용한 전기차 판매량 전망	- 선행연구 분석 - SSA(Singular Spectrum Analysis) 모형 분석 - VAR(Vector Auto Regressive) 모형 분석	중국 전기차 판매량에 대해 소비자 물가지수, 소비자 신뢰지수, 공산품 가격 지수, 소매 연료 가격, 자동차 가격, Baidu 검색 지수 등을 독립변수로 활용한 시계열 모형을 통한 수요 예측
9	- 과제명: An analysis of factors that affecting the number of car sales in Malaysia - 연구자(년도): Islam et al. (2016) - 연구목적: 거시경제 변수를 활용한 회귀모형을 통해 자동차 수요 추정	- 선행연구 분석 - 거시경제 변수 선정 - 회귀분석 - ANOVA를 활용한 가설 검정	- 말레이시아 자동차 판매량에 영향을 주는 거시경제 변수를 활용하여 회귀분석을 통해 모형 설계 - 자동차 판매량과 개별 거시변수 간 상관관계를 ANOVA로 검정
10	- 과제명: A sales forecasting model in automotive industry using adaptive neuro-fuzzy inference system and genetic algorithm - 연구자(년도): Vahabi et al. (2016) - 연구목적: ANFIS 모형을 활용한 연간 자동차 판매량 예측	- 선행연구 분석 - ANFIS 모형 분석 - 모형 적합도 검정	인구, 실업률, 환율, 생산지수, 제조업 판매지수, 전체 소비, 소비자 물가, 가스 가격, GDP, 물가상승률, 대출 기준금리, 추가 지수, 1인당 자본, 자동차 수출 세금 등 방대한 거시경제 데이터를 활용하여 ANFIS 모형으로 판매량 예측

NO	연구 목적	연구 방법	주요 연구 내용
11	- 과제명: Modelling and forecasting the demand for coal in china - 연구자(년도): Chan & Lee (1997) - 연구목적: 중국의 주요 에너지 원인 석탄의 수요 분석 및 예측	- 선행연구 분석 - Engle-Granger 모형 분석	- 공적분 분석과 ECM을 사용하여 중국의 석탄에 대한 수요를 모델링 및 예측 - Engle-Granger's ECM과 Hendry's ECM, General-to-specific 접근 등 세 모델의 성능 비교를 위해 사후 예측 비교 분석
12	- 과제명: Forecasting global coal consumption: An artificial neural network approach - 연구자(년도): Pablo Benalcazar et al.(2017) - 연구목적: 인공신경망 분석방법론을 이용하여 글로벌 석탄 소비를 예측함	- 선행연구 분석 - 인공신경망 분석 - 다층 퍼셉트론 신경망(MLP) 모형	1970년부터 2016년까지 IMF의 세계 경제전망 데이터를 이용, 다층 퍼셉트론 신경망 모델을 적용하여 세계 석탄 소비량을 예측 - 예측의 성과 개선을 위해 하이브리드 ANN 기반 기법 및 사회학적 변수 추가 적용을 제안
13	- 과제명: Forecasting of Coal Demand in China Based on Support Vector Machine Optimized by the Improved Gravitational Search Algorithm - 연구자(년도): Yanbin Li and Zhen Li(2019) - 연구목적: 회색 관계 분석을 활용한 중국의 석탄수요 예측	- 선행연구 분석 - 회색 관계 분석(시스템 요인간 영향도 분석 방법론) - 서포트 벡터 머신(SVM) 모델 분석	- 석탄수요에 영향을 미치는 요인 및 지표체계를 구축 - 회색관계 분석법을 이용한 주요 입력변수를 선정 및 종력 탐색 알고리즘 최적화를 통해 IGSA-SVM 모델을 구축, '25년까지 중국의 석탄 수요를 예측함
14	- 과제명: Long-term global scrap availability of steel scrap - 연구자(년도): Oda et al. (2013) - 연구목적: 2050년 고철 가용성 파악	- 선행연구 분석 - Material flow analysis (MFA) - Life-time model	고철 발생원 중 폐기 철에 따른 재활용률을 분석하고, 여러 시나리오에 따른 장래 철 생산량을 가정하여 장래 고철 가용성을 파악
15	- 과제명: The Steel Scrap Age - 연구자(년도): Pauliuk et al. (2013) - 연구목적: 장래 철 수요량 및 고철 공급량 예측	- 선행연구 분석 - Dynamic stock model - Demand-driven capacity model	장래 철 재고량과 고철 발생 시점, 재고량 포화 지점을 고려하여 장래 철 수요량 예측

NO	연구 목적	연구 방법	주요 연구 내용
16	• 과제명: Steel scrap generation in the EU-28 since 1946 – Sources and composition • 연구자(년도): Dworak & Fellner (2021) • 연구목적: 1946~2017년 EU-28 고철 세부품목별 발생량 조사	• 선행연구 분석 • MFA • Material pinch analysis	• 1946~2017년 EU 내 발생한 고철 발생량 및 세부 품목 변화를 분석 • EU 고철은 철 생산과정에서 발생한 신규 고철에 비해 폐기 철에 의해 발생한 고철의 비중이 확대 • 2008년 기준 세계 고철 비중 구조는 약 30년 전 EU 고철 비중과 유사함
17	• 과제명: Global Scrap Trading Outlook Analysis for Steel Sustainability • 연구자(년도): Lee & Sohn (2015) • 연구목적: 국가 간 연도별 고철 무역 구조 파악 및 고철 수요 예측	• 선행연구 분석 • 네트워크 분석 • 시계열 분석(ARIMA)	• 국가 간 고철 수출입 물동량 네트워크 구조 분석 • 시계열 기법을 통해 철 생산량 및 고철 수요량 예측
18	• 과제명: Outlook of the World Steel Cycle Based on the Stock and Flow dynamics • 연구자(년도): Hatayama et al. (2010) • 연구목적: 철 재고량 분석을 통한 세계 철 수요량 예측	• 선행연구 분석 • Dynamic MFA	• 42개국의 철 사용량 및 재고량을 분석 • 2005년까지의 철 재고량과 경제 성장률을 통해 국가별 장래 철 재고량을 예측 • 재고량을 활용해 2050년까지 장래 철 수요 예측
19	• 과제명: The dynamics of fleet size and shipping profitability: the role of steel-scrap prices • 연구자(년도): Andrikopoulos et al.(2020) • 연구목적: 국제 스크랩 가격을 활용한 선박 스크랩 가격 예측	• 선행연구 분석 • VEC 모형	• 선박 스크랩 가격 예측에 활용할 수 있는 변수로 국제 스크랩 등을 제시 • 국제 스크랩 가격과 연관된 거시경제변수를 제시하여 본 연구 내 참고 가능
20	• 과제명: Forecast of steel demand and the availability of depreciated steel scrap in China • 연구자(년도): Xuan & Yue (2016) • 연구목적: 중국 내 철 수요 예측 및 고철 가용성 파악	• 선행연구 분석 • IPAT 모형	• IPAT 모형을 통해 철 생산량, 단위 철 생산당 소요되는 고철 양, 고철 수요량을 예측

NO	연구 목적	연구 방법	주요 연구 내용
21	- 과제명: Forecasting Domestic Shipping Demand of Cement: Comparison of SARIMAX, ANN and Hybrid SARIMAX-ANN - 연구자: Fiskin & Cerit (2019) - 연구목적: 시멘트 항만수요 예측을 위한 SARIMA, ANN 및 하이브리드 SARIMA&ANN의 예측력 비교	- Shapiro-Wilk 및 Jarque-Bera 데이터 정규성 검정 - Augmented Dickey-Fuller 및 Philips Perron 일관성 분석 - SARIMAX 예측모형 적용 - ANN 예측모형 적용 - Hybrid SARIMAX &ANN 예측모형 적용	- 터키의 시멘트 연간 물동량 예측을 위한 SARIMAX, ANN 및 하이브리드 SARIMAX &ANN 비교 13년간 월별 물동량을 활용하여 4분기(19개월) 물동량 예측 - 하이브리드 SARIMA&ANN 모형이 가장 좋은 예측력을 보임
22	- 과제명: Forecasting Determinant of Cement Demand in Indonesia with Artificial Neural Network - 연구자: Fradinata et al. (2015) - 연구목적: 시멘트 항만수요 예측을 위한 외생변수 분석 및 ANN 기반 수요예측 모형 제안	- GDP, 인구 등의 외생변수 분석 - Durbin Watson 테스트를 통해 변수별 상관관계 분석 - ANN 기반 예측모형 적용 - Mean Squared Error(MSE)를 통해 모형 예측력 평가	- 인도네시아의 시멘트 물동량 예측을 위해 ANN 적용, 관련 변수별 예측력 비교 8년간 월별 데이터를 활용하여 4분기(7개월) 물동량 예측 - 기본적인 ANN을 적용하며 GDP, 인구, 고객수, 가격, 판매량, 판촉량, 품질 등 변수 고려
23	- 과제명: ANN, ARIMA and MA Timeseries Model for Forecasting in Cement Manufacturing Industry - 연구자: Fradinata et al. (2015) - 연구목적: 인도네시아 Arch 지역의 시멘트 판매량을 ANN, ARIMA, MA로 예측한 사례 연구	- 포졸란, 포틀랜드, 벨크 시멘트로 구분하여 예측 - ANN, ARIMA, 예측모형 적용 및 결과 비교 - MSE와 RMSE(Root MSE)를 적용해 모형 예측력 비교	- 인도네시아 Aceh 지방의 시멘트 판매량과 판매 수요 예측을 위한 시계열 예측 모델 개발 22개월 데이터를 활용하여 4분기(2개월) 시멘트 수요예측 - ANN, ARIMA와 MA의 비교 연구 및 ANN 예측력 검증
24	- 과제명: Application of Artificial Neural Network and SARIMA in Portland Cement Supply Chain to Forecast Demand - 연구자: Liu et al.(2008) - 연구목적: 시멘트 수요 예측을 위해 SARIMA와 ANN의 예측력 비교	- 시멘트 수요 데이터를 월별, 분기별, 연별로 분석 - SARIMA와 ANN 모형 적용 및 예측력 비교 - Mean of Absolute Percentage Errors(MAPE)를 적용해 예측력 비교	- 대만의 시멘트 수요 발생을 예측하기 위한 모델로 SARIMA, ANN을 적용 20년의 시멘트 수요 데이터를 활용하여 4분기(14개월) 예측 - SARIMA와 ANN의 예측력 비교 - 분기별 데이터 학습으로 예측력 상승

기본연구보고서 발간목록

Ⅰ 2022년

01	선박투자 가치평가 및 위험관리 모형 연구	박성화
02	항만 컨테이너 반출입 예약시스템 가격결정 모형 연구	서정용
03	연안재해 대응을 위한 그린인프라 구축방안 연구	정지호
04	탄소중립이 해양수산업에 미치는 경제적 영향분석 연구	권장한
05	수산물 안전성 관리체계 개선방안 연구	조현주
06	준해양사고 통보제도 개선방안 연구	박상원
07	MZ세대 소비트렌드를 반영한 해양관광 추진방안 연구	최일선
08	연근해 어종별 어획쿼터제도 도입방안 연구	심성현
09	행위자기반 해양공간계획 시뮬레이션 개발 연구	조성진
10	항만분야 탄소중립 관리체계 개선방안 연구	안승현
11	탄소중립 시대 양식산업 대응전략 연구 - 육상 어류양식을 중심으로 -	마창모
12	수산물식품산업의 탄소중립 대응방안 연구 - 가공업을 중심으로 -	김지연
13	해양환경보건정책 도입방안 연구	박수진
14	포스트 코로나 시대의 컨테이너 해운산업 대응방안 연구 - 디지털 플랫폼을 중심으로 -	최건우
15	항만분야 공공갈등 관리방안 연구	김세원
16	글로벌 가치사슬 변화와 국제물류 분야 대응방안 연구 - 전기자동차 산업을 중심으로 -	권보배
17	항만개발제도 개선방안 연구	이수영
18	클러스터 혁신을 통한 선박관리업 발전방안 연구	허성례
19	수입수산물 예방적 관리체계 도입방안 연구 - IUU 수산물을 중심으로 -	안지은
20	양식수산물 유통 빅데이터 구축방안 연구	이기영
21	글로벌 혁신성장을 위한 신남방·신북방 정책 연구 - 신남방 지역 ICT산업 가치사슬 변화에 따른 물류 공급사슬 대응방안	최나영환

Ⅱ 2021년

01	시민참여형 해안돌봄 정책 도입방안 연구	정치호
02	해양 플라스틱 쓰레기가 선박 운항에 미치는 영향분석 연구	김보람
03	여성어업인 노동정책 방향 연구	홍혜수
04	순환경제 시스템을 활용한 어업폐기물의 자원화 방안 연구	고동훈
05	1conomy 시대, HMR 시장 확대에 따른 수산부문 대응전략 수립 연구	마창모
06	AIS 데이터 기반 해상교통 안전 평가모델 개발 연구	황선일
07	항만 에너지 관리시스템 도입을 위한 로드맵 구축 연구 - 부산항 신항 전기에너지 소비를 중심으로 -	김근섭
08	해양분야 리빙랩 활성화 방안 연구	좌미라
09	빅데이터의 연계·활용을 통한 선박의 배출량 산정체계 고도화 방안 연구	안용성
10	해양공간계획 집행체계 고도화 방안 연구	최희정
11	수산식품산업 푸드테크 환경분석 및 적용전략 연구	이상건
12	AIS데이터기반 해상물동량 추정 연구	황수진
13	항만부문 정책의사결정 지원 빅데이터 플랫폼 구축 방안 연구	이기열
14	항만 컨테이너 물류 프로세스 디지털화 모형 연구	서정용
15	해양수산 분야 국제산업연관표 구축기반 연구	정수빈
16	해양수산 과학기술 정책평가모형 연구 - 해양수산 R&D를 중심으로 -	전형모
17	글로벌 경제위기와 해운산업 대응체계 연구	박성화
18	항만산업 경기진단체계 구축 연구 - 경기동행지수 개발을 중심으로 -	김성아
19	전국 무역항 부두 분류 체계 개선방안 연구	이수영
20	IMO 규제기반 해사산업의 글로벌 지속발전방안 연구 - 新해사산업의 육성 및 지원을 위한 법제화 연구(5차년도) -	박한선
21	수소에너지 거점 구축을 위한 항만의 대응방안 연구	신수용

수시연구보고서 발간목록

Ⅰ 2022년

01	항만의 탄소중립 이행·관리 표준안 연구	안용성
02	매립지 소유권 분리 제도화의 이해관계 분석과 관리방안	윤성순
03	글로벌 공급망 리스크별 영향분석 및 대응방안 연구 - 에너지·곡물을 중심으로	조지성
04	대북제재 강화와 코로나 팬데믹 이후 북한 해양수산 이슈와 대응 방안	윤인주
05	남극환경보호의정서 제6부속서의 국내 이행을 위한 법제 정비방안 연구	박예나
06	해양바이오산업 활성화를 위한 해양수산생명자원 법령 정비 방안 연구	좌미라
07	우리나라 수산종자 관리체계 개선방안 연구	조현주
08	2050 신해양강국 미래비전 수립 연구	김민수

Ⅰ 2021년

01	해양환경산업 육성 방안 연구	한기원
02	선원의 인권 및 근로여건 향상을 위한 선원근로감독관 제도 개선 방안 연구	허성례
03	포스트 코로나 시대를 대비하는 지역 해양축제 활성화 연구	최일선
04	회복탄력성 개념을 적용한 항만 위기관리 시스템 구축 연구	김성기
05	중대재해처벌법 시행에 따른 항만에서의 대응방안 연구	최상균
06	크루즈산업 COVID-19 방역체계 구축방안 연구	황진희
07	한국형 선주사의 최적 운영방안 연구	고병욱
08	어촌형 생활서비스 전달체계 개선 연구	이호림
09	원양산업의 ESG 도입 기초 연구	윤미경
10	수상레저활동 관리체계 개선방안 연구	홍장원
11	연안지역 자연성 회복을 위한 정책 방향_간척지·담수호 중심으로	정지호
12	스마트항만 구축에 따른 탄소저감 효과 연구 - 동력전환을 중심으로 -	김가현

일반연구보고서 발간목록

Ⅰ 2022년

01	건화물선 시장 비용분석을 통한 해운시장 위험관리지표체계 구축	류희영
02	해상운임 예측모형 고도화 연구(Ⅰ)	황수진
03	IMO 시장기반조치 도입이 국내 해운기업에 미치는 영향	김한나
04	연안지역발전지수(CoDI) 개발 연구	황재희
05	CGE 기반 국제통상환경 변화의 해양수산부문 파급효과 분석	임병호
06	글로벌 수산부문 이머징 이슈 분석 모델 개발 연구	한기욱
07	양식 수산물 중장기 수급전망모형 구축 연구 - 굴을 중심으로 -	김철현
08	Network DEA를 이용한 물류기업 경쟁력 비교 분석 연구	황선일
09	해양수산 사업체 성과 및 효율성 분석 연구	김주현
10	베이지안 방법을 이용한 양식 멧게 단수 추정방안 연구	천성훈
11	수입수산물과 국산 간의 대체관계 분석 연구 - 활·산선냉장품을 중심으로 -	박혜진
12	국내 컨테이너 해운기업의 디지털 전환 활성화 방안 연구	전서연
13	비컨테이너 항만물동량 예측모형 고도화 방안 연구(Ⅰ) - 자동차, 석탄, 고철, 시멘트를 중심으로 -	최석우
14	동중국해 해양경계획정 방안 기초 연구	박영길
15	해양수산 분야 미래 리스크 발굴 및 파급효과 분석 연구(Ⅱ)	김찬호
16	항만개발사업의 정책효과 적용방안 연구	이종필
17	국가 해양전략 기본구상 연구	정현욱
18	선박대기시간 예측모형 개발을 위한 방법론 연구	조아현
19	안전항만 구축을 위한 비용 산정 및 지원체계 마련을 위한 연구	최상균

2021년

01	해양수산 정책영향평가를 위한 기초 연구	김주현
02	해양교육의 사회·경제적 가치 평가 연구	이슬기
03	해양수산업 경기진단체계 기초연구 - 해운업을 중심으로 -	권장한
04	선박 기술진보를 고려한 탄소 배출량 추정 연구	최건우
05	서포트벡터머신(SVM) 기법을 활용한 해운시황 예측 연구	김병주
06	크루즈 여객 수요 전망 모형 구축 연구	안승현
07	디지털 공급사슬 물류정보통합 구축전략 연구(Ⅲ)	서정용
08	수출용 전복의 유통경로 분석 연구	이정필
09	인공지능기반 해상운임 예측 연구(3차년도)	황수진
10	인공신경망모형을 이용한 양식수산물 단수 전망에 관한 연구 - 김 양식을 중심으로 -	천성훈
11	해양수산분야 미래 리스크 발굴 및 파급효과 분석 연구	박광서

일반연구 2022-13

**비컨테이너 항만물동량 예측모형
고도화 방안 연구(Ⅰ)**
- 자동차, 석탄, 고철, 시멘트를 중심으로

인쇄 2022년 12월 29일

발행 2022년 12월 31일

발행인 김 종 덕

발행처 한국해양수산개발원

주 소 49111 부산시 영도구 해양로 301번길 26(동삼동)

연락처 051-797-4800 (FAX 051-797-4810)

등록 1984년 8월 6일 제313-1984-1호

조판·인쇄 공감 Tel: 051-903-9909

판매 및 보급: 정부간행물판매센터 Tel: 02-394-0337

정가 6,000원