



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

經營學碩士 學位論文

인공신경망(ANN) 모델을 활용한 상해 컨테이너
운임지수 예측

A Study on the Forecasting of Shanghai Containerized
Index Using Artificial Neural Network



指導教授 柳 東 瑾

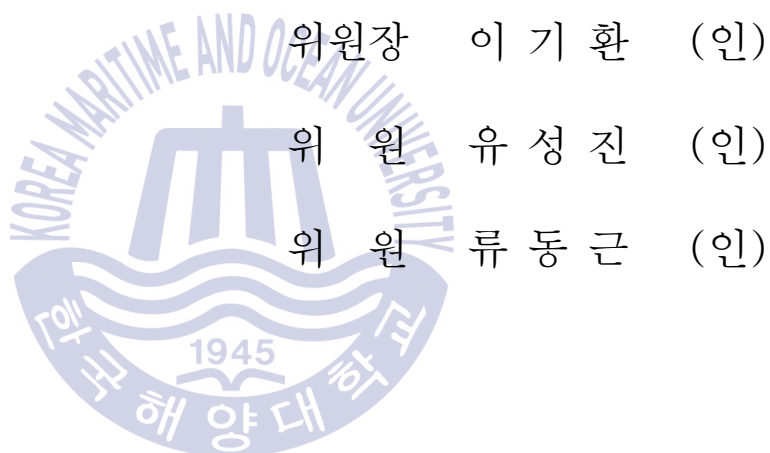
2019년 2월

韓國海洋大學校 大學院

海運經營學科

曹 相 鎬

本 論文을 曹相鎬의 經營學碩士 學位論文으로 認准함



2019년 2월

韓國海洋大學校 大學院
海運經營學科

목 차

Abstract	VI
제1장 서론	1
1.1 연구의 배경 및 목적	1
1.2 연구의 범위 및 방법	2
1.3 연구의 구성	2
제2장 컨테이너 해운 시장 현황 및 운임지수	4
2.1 컨테이너 해운 시장 현황	4
2.1.1 세계 컨테이너선 선복량 추세	4
2.1.2 세계 컨테이너 물동량 추세	7
2.1.3 세계 주요 선사 얼라이언스 개편 및 M&A 추진 현황	10
2.2 컨테이너 해운시장 관련 지수	14
제3장 선행연구 및 예측방법론 고찰	25
3.1 선행연구 고찰	25
3.2 분석방법론	28
3.2.1 ARIMA(Auto-regressive Integrated Moving Average)	28
3.2.2 자기회귀모형	32
3.2.3 인공신경망 모형	34
3.2.4 이중계절형 Holt-Winters 모형	46
제4장 인공신경망을 활용한 SCFI 예측 모형 구축	49
4.1 예측 요인 선정	49
4.1.1 SCFI(Shanghai containerized Freight Index)	49
4.1.2 HRCI(Howe Robinson Containership index)	50
4.1.3 Bunker 380CST Index	51
4.1.4 U.S. Dollar Index	53
4.2 예측 요인 상관 분석	55

4.3 최적 인공신경망 예측 모형 구축	57
4.3.1 최적 인공신경망 예측 모형 구축 과정	57
4.3.2 최적 인공신경망 모형을 통한 예측 수행	64
제5장 결론	67
5.1 연구결과 및 시사점	67
5.2 연구의 한계점 및 향후 연구방향	69
참고문헌	70



List of Tables

Table 1	세계 컨테이너선 선형별 선복량 추이(1997년 ~ 2017년)	4
Table 2	세계 상위 컨테이너 선사 운항선복량 (2018년 4월 기준)	7
Table 3	세계 컨테이너 물동량(1999년 ~ 2017년)	8
Table 4	세계 컨테이너 선복량 및 물동량 증감율(2008년 ~ 2017년)	9
Table 5	주요 컨테이너선사 2018년 2분기 영업실적	13
Table 6	주요 컨테이너선사 2018년 상반기 영업실적	13
Table 7	세계 컨테이션 시장 주요 관련 지수	15
Table 8	WCI(The World Container Index)	18
Table 9	HRCI의 대표 선형 및 가중치	23
Table 10	주요 선행연구 결과요약	25
Table 11	생물학적 신경망과 인공신경망 구성요소 비교	35
Table 12	SCFI 자료 기초통계량분석	49
Table 13	HRCI 자료 기초통계량분석	50
Table 14	Bunker 380CST Index 해당 주요 항만	52
Table 15	Bunker 380CST Index 자료 기초통계량분석	52
Table 16	U.S. Dollar Index 자료 기초통계량분석	54
Table 17	r값 구간에 따른 상관관계	55
Table 18	예측 요인 상관관계 분석 결과	56
Table 19	인공신경망 예측 모형 학습결과	58
Table 20	최적 인공신경망 모형 구성	63

List of Figures

Fig. 1 연구의 구성	3
Fig. 2 세계 컨테이너선 선복량 추세(1997년 ~ 2017년)	6
Fig. 3 Global Alliance 개편	11
Fig. 4 세계 주요선사 M&A 추진현황	12
Fig. 5 SCFI(Shanghai containerized Freight Index) 추이	17
Fig. 6 CCFI(China containerized Freight Index) 추이	17
Fig. 7 Shanghai ↔ Genoa WCI(The World Container Index) 추이	20
Fig. 8 Rotterdam ↔ Shanghai WCI(The World Container Index) 추이	20
Fig. 9 Rotterdam ↔ New York WCI(The World Container Index) 추이	21
Fig. 10 Shanghai ↔ Los Angeles WCI(The World Container Index) 추이	21
Fig. 11 Rotterdam ↔ Los Angeles WCI(The World Container Index) 추이	22
Fig. 12 HRCI(Howe Robinson Containership Index) 추이	23
Fig. 13 Contex(1,100 ~ 2,5500 TEU) All time/Last 6 Months 추이	24
Fig. 14 Contex(2,700 ~ 4,2500 TEU) All time/Last 6 Months 추이	24
Fig. 15 안정적 시계열 및 불안정적 시계열	29
Fig. 16 ARIMA 모형 분석 단계	30
Fig. 17 Durbin-Watson 검정의 구간별 판정	33
Fig. 18 Processing Unit 과 Neuron 비교 1	35
Fig. 19 Processing Unit 과 Neuron 비교 2	36
Fig. 20 인공신경망 계층구조 단순화 모형	37
Fig. 21 경사하강법(Gradient Descent Method)	38
Fig. 22 경사하강법(Gradient Descent Method)	39
Fig. 23 다층 역전파(Back propagation)	40
Fig. 24 ReLU 함수	41
Fig. 25 Single-layer-ANN	42
Fig. 26 데이터 변환과 Classification	44
Fig. 27 Multi-layer-ANN	44
Fig. 28 오차의 역전파	45
Fig. 29 SCFI 추이(2009.10 ~ 2018.9)	50

Fig. 30 HRCI 추이(2009.10 ~ 2018.9)51

Fig. 31 Bunker 380CST Index 추이53

Fig. 32 U.S. Dollar Index 추이54

Fig. 33 예측요인 산점도56

Fig. 34 최적 인공신경망 예측 모형 구축 과정 요약57

Fig. 35 최적 인공신경망 모델과 SCFI 실측값 비교 그래프64

Fig. 36 다단계 예측 모형(Multi-step ahead forecasting)64

Fig. 37 다단계 예측 모형을 통한 SCFI 예측값(2018년 10월 ~ 2019년 1월)65

Fig. 38 SCFI 실측값 및 예측값 비교 그래프(2018년 10월 첫째주 ~ 2018년 12월 둘째주)66



A Study on the Forecasting of Shanghai Containerized Index Using Artificial Neural Network

Cho, Sang Ho

Department of Shipping Management
The Graduate School of Korea Maritime and Ocean University



Abstract

In the global container shipping market, the market imbalance is worsening due to the market imbalance caused by the increase in excess fleet and the reorganization of the alliance, which shows the form of oligopoly. In addition, the container freight rate index SCFI, which is representative of the container shipping industry, has repeatedly fluctuated and reflects the volatility of the freight market. As such, the imbalance and uncertainty of the container shipping market are gradually increasing, and studies on forecasting are needed to prepare for countermeasures against market changes and crises.

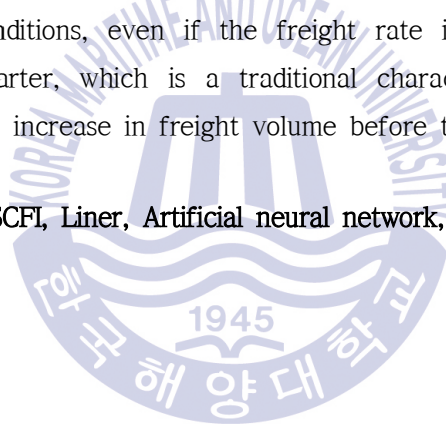
This study predicts SCFI from October, 2018 to January, 2019 through an artificial neural network (ANN) model using the time series data of SCFI which is representative container freight index from October, 2009 to September, 2018. In

addition, the change trends were identified by comparing SCFI predictions and measured values.

For the construction of the artificial neural network prediction model, HRCI, Bunker 380CST Index, U.S. Dollar Index was used. In order to construct an optimal artificial neural network prediction model, trial and error method which adjusts various hyper parameters and experimentally reduces error is used.

Predicted results were a slight uptrend of a maximum of 1,010 points, a minimum of 992 points, but an overall slowdown. In this study, SCFI measured and predicted between the first week of October, 2018 and the second week of December, 2018 were compared. There is a difference in relative value but it shows the same declining trend. The results of the forecast analysis are expected to indicate the deteriorating market conditions, even if the freight rate increase due to the high season in the third quarter, which is a traditional characteristic of the container shipping market, and the increase in freight volume before the Lunar New Year.

Keyword : Container, SCFI, Liner, Artificial neural network, Multilayer perceptron



인공신경망(ANN) 모형을 활용한 상해 컨테이너 운임지수 예측

조 상 호

해운경영학과
한국해양대학교 대학원

초 록

세계 컨테이너 해운시장은 과잉선복량 증대로 인한 시장 불균형 현상 및 과점형태의 양상을 나타내는 얼라이언스 개편 등으로 인해 시장의 불균형은 심화되어가고 있다. 또한 컨테이너 해운 시장을 대표하는 컨테이너 운임지수인 SCFI는 급등락을 반복하며 운임시장의 급격한 변동성을 반영하고 있다. 이처럼 컨테이너 해운시장의 불균형과 불확실성은 점차 증대되고 있으며 시장의 변화와 위기에 대한 대응방안 마련을 위해서는 예측에 관한 연구가 필요하다.

본 연구는 2009년 10월부터 2018년 9월까지 대표적인 컨테이너 운임지수인 SCFI의 시계열자료를 활용하여 인공신경망(ANN) 모델을 통해 2018년 10월부터 2019년 1월까지 SCFI를 예측하였다. 또한 SCFI의 예측치와 실측치 비교를 통해 변화 추세를 파악하였다.

인공신경망 예측모델 구축을 위해 독립변수로 HRCI, Bunker 380CST Index, U.S. Dollar Index를 활용하였다. 최적의 인공신경망 예측모델을 구축하기 위해 다양한 Hyper-Parameters를 조정하며 실험적으로 오차를 줄여나가는 시행착오법을 사용하였다. 인공신경망 예측모델의 성능을 평가하기 위한 기준은 RMSE = 0.136, MSE = 0.018, MAE = 0.116, MAPE = 0.272로 평가 되었다

예측 결과 최대 1,010 포인트 최소 992포인트로 일시적인 소폭상승세를 나타내지만

전반적으로 지속적인 하락세를 나타내는 것으로 예측되었다. 2018년 10월 첫째주부터 2018년 12월 둘째주 간의 SCFI 실측치와 예측치를 비교한 결과 상대적인 값의 차이는 존재하나 동일한 하락 추세를 나타내고 있다. 이러한 예측분석 결과는 컨테이너 해운 시장의 전통적인 특징인 3/4분기 성수기로 인한 운임 상승세 유지 및 춘절 전 물동량 증가를 가만하더라도 악화된 시장상황을 나타내는 것으로 추정된다.

주제어 : 컨테이너, SCFI, 정기선, 인공지능경망, 다층 퍼셉트론



제1장 서론

1.1 연구의 배경 및 목적

세계 컨테이너 해운시장은 2017년 기준 이전 4개 얼라이언스(Alliance)가 3개의 얼라이언스로 통합되었고 상위선사 중심의 적극적인 M&A로 상위 20개 선사가 11개 선사로 통합되어 과점형태의 양상을 보이고 있다. 또한 화물물동량을 고려하지 않는 과잉선복량 증대로 인해 컨테이너 해운시장의 불균형은 점점 심화되어가고 있다.

이러한 컨테이너 해운시장의 문제점으로 인해 시장의 불확실성은 점차 커지고 있으며 정기선 시황을 대표하는 중국 상하이 운임지수(Shanghai containerized Freight Index ; SCFI)는 2016년 3월 18일 400.3으로 최저를 기록하였다. 이 후 등락을 반복하며 운임시장의 변동성은 점차 심화되고 있다.

과거에 그러하였듯 세계 컨테이너 해운시장은 급변하고 있으며 변화를 정확히 파악하지 못하고 이에 대한 적절한 대응방안을 마련하지 못한다면 2008년 리먼 브라더스 사태와 EU 경제위기처럼 악순환을 반복하게 될 것이다. 따라서 이러한 컨테이너 해운시장의 미래 예측에 관한 연구가 시급한 실정이다.

불확실성 높은 시장의 특성 및 정기선 운임의 특수성으로 인해 기존 선행연구는 시계열 자료를 활용한 예측보다는 정기선 운임에 영향을 미치는 변수간의 관계성 분석이 주를 이루고 있다. 또한 일부 예측 연구에서는 수요와 공급측면에 초점을 두어 연구를 수행하였다.

본 연구에서는 다양한 경제요인을 활용하여 컨테이너 해운시장의 향후 시장 변동성을 파악하기 위해 운송거래의 기준이 되는 대표적인 컨테이너 운임지수인 SCFI를 인공신경망 모델을 통해 최적 예측모델 구축함을 목적으로 한다.

1.2 연구의 범위 및 방법

본 연구에서는 2009년 10월부터 상해항운교역소(Shanghai Shipping Exchange)에서 발간한 SCFI의 시계열 자료를 바탕으로 예측 분석을 수행함을 목표로 한다.

이에 활용한 경제요인으로 대표적인 컨테이너 용선료 지수인 HRCI(Howe Robinson Containership index), 주된 선박연료유 종류인 380CST의 평균가격인 Bunker 380CST Index 및 통화가치가 비교적 안정적인 주요 6개국의 통화 대비 미국 달러의 가치를 지수화한 것으로 주요국의 화폐 가치 대비 달러 가치의 움직임을 파악하기 위하여 만들어진 지수인 U.S. Dollar Index를 선정하였다.

연구에 활용된 자료는 2009년 10월 ~ 2018년 9월까지 주별 시계열 자료를 활용하였다.

분석방법론으로는 인공신경망분석(Artificial Neural Network : ANN)을 사용하였고 최적 인공신경망 모델 설계를 위해 다양한 Hyper-Parameters를 조정하여 실험적으로 오차를 줄여나가는 시행착오법을 사용하였다.

1.3 연구의 구성

제1장 서론은 연구의 배경 및 목적을 통해 연구의 필요성을 설명하고, 연구범위와 방법을 설명함으로써 전반적인 연구흐름을 제시하였다.

제2장 컨테이너 해운 시장 현황 및 운임지수에서는 세계 컨테이너 물동량, 세계 컨테이너 선복량 추세 및 세계 컨테이너 선사 얼라이언스 개편 및 M&A 추진 현황을 서술함으로써 컨테이너 해운시장의 전반적인 특징과 변화에 대하여 기술하였다. 또한 컨테이너 해운 운임지수의 도입배경 및 관련지수를 설명하였다.

제3장 선행연구 및 예측방법론 고찰에서는 먼저 관련 선행연구를 검토하였고 시계열 분석에 활용되는 대표적인 분석방법론에 대한 기술을 통해 다양한 분석방법의 장단점 및 특징을 파악하였다.

제4장 인공신경망을 활용한 SCFI 예측 모형 구축에서는 먼저 본 연구에서 활용할 변수 간 상관 분석을 수행하였다. 또한 인공신경망 모형을 활용하여 예측 모형을 설계하였으며 최적 모형을 도출하기 위한 Hyper-Parameters 조절을 통한 모델링을 실시하여 도출된 최적 모형을 바탕으로 SCFI를 예측하였다.

제 5장 결론에서는 본 연구에서 도출된 분석결과를 바탕으로 연구결과 및 시사점을 제시하였고 연구를 수행함에 있어 파악된 한계점과 향후 연구방향에 대하여 기술하였다.

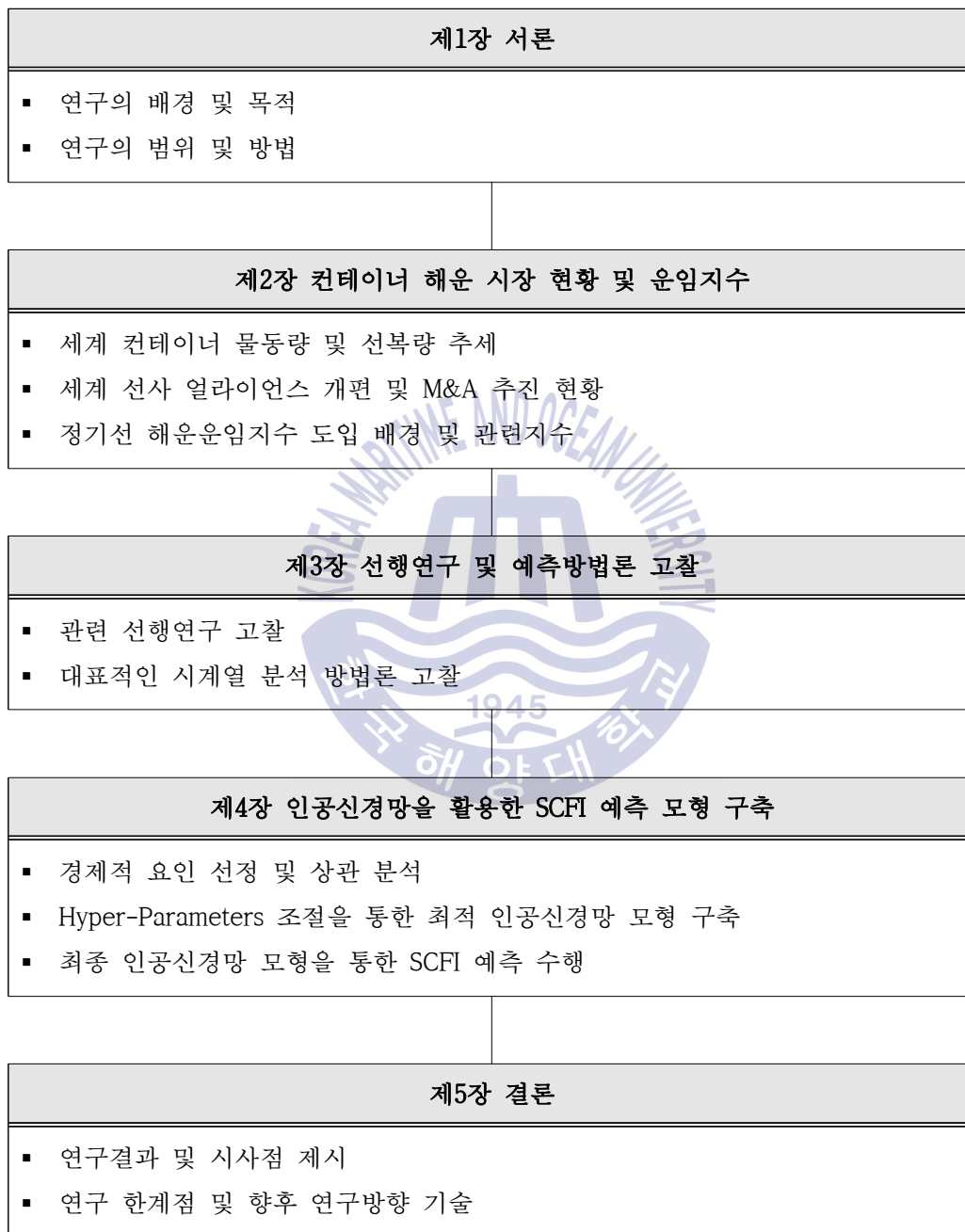


Fig. 1 연구의 구성

제2장 컨테이너 해운 시장 현황 및 운임지수

2.1 컨테이너 해운 시장 현황

2.1.1 세계 컨테이너선 선복량 추세

컨테이너선 선복량은 컨테이너 해운시장에서 공급량을 파악하는 중요한 요인이다. 컨테이너선 선사들은 선대의 대형화를 통해 화물의 유효 적재량을 증가시켜 해운시장에서 우위를 차지하려 한다. 선복량 증가를 통한 규모의 경제 실현을 통해 단위당 원가비용 절감 효과를 얻을 수 있지만 과잉공급으로 인한 시장 불균형을 초래할 수 있다.

2008년 세계 경제위기 당시에도 선복량은 지속적으로 증가하였으며 이후 2016년까지 최소 5.4% 증가율을 나타내고 있다.

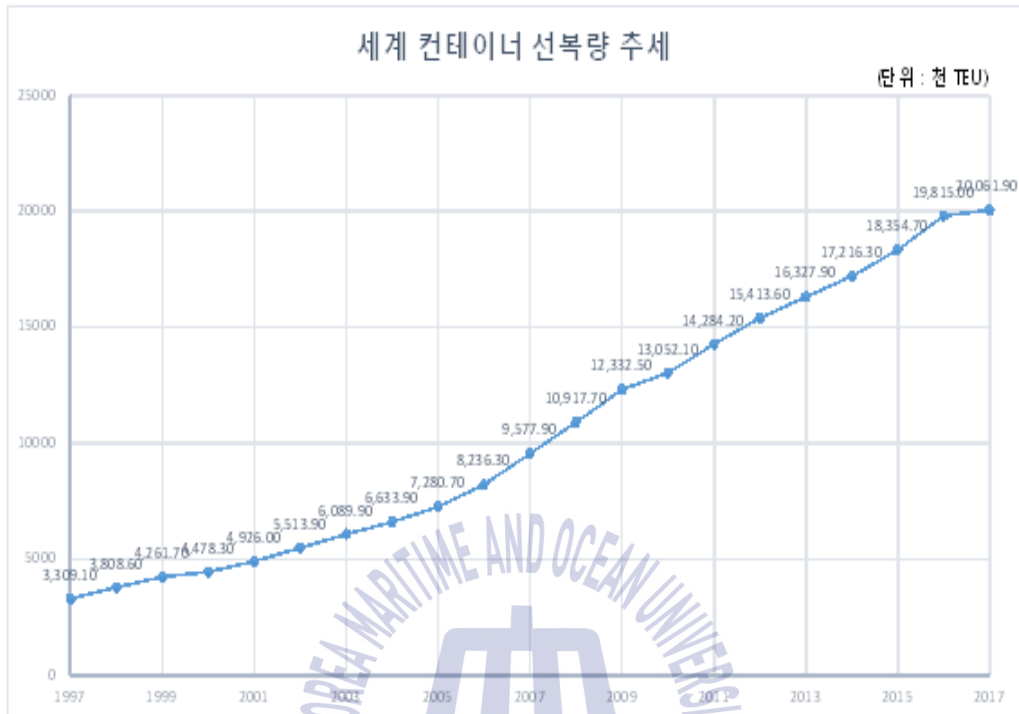
Table 1 세계 컨테이너선 선형별 선복량 추이(1997년 ~ 2017년)

(단위 : 천TEU)

연도	100- 2,999	3,000- 5,999	6,000- 7,999	8,000- 11,999	12,000- 14,999	P-Pmax 15,000+	합 계	증감률 (%)
1997	2,030.0	1,248.5	29.6	-	-	-	3,309.1	13.4
1998	2,268.3	1,470.3	50.8	19.3	-	-	3,808.6	15.1
1999	2,466.5	1,636.2	102.6	56.3	-	-	4,261.7	11.9
2000	2,557.3	1,723.5	102.6	94.8	-	-	4,478.3	5.1
2001	2,685.0	1,953.8	153.9	133.3	-	-	4,926.0	10.0

2002	2,830.0	223.5	317.5	143.0	-	-	5,513.9	11.9
2003	2,946.9	2,498.3	466.3	178.5	-	-	6,089.9	10.4
2004	3,086.3	2,727.9	592.0	227.7	-	-	66,339.	8.9
2005	3,219.9	3,046.9	661.6	352.3	-	-	7,280.7	9.7
2006	3,438.8	3,418.2	738.3	641.1	-	-	8,236.3	13.1
2007	3,742.5	3,761.7	893.0	1,147.3	-	33.4	9,577.9	16.3
2008	4,067.0	4,214.7	1,076.9	1,441.1	-	117.9	10,917.7	14.0
2009	4,356.2	4,651.1	1,262.4	1,913.4	13.8	135.7	12,332.5	13.0
2010	4,298.9	4,941.8	1,401.5	2,192.5	81.7	135.7	13,052.1	5.8
2011	4,329.4	5,298.1	1,573.2	2,527.4	420.4	135.7	14,284.2	9.4
2012	4,339.3	5,457.1	1,675.2	2,922.9	883.3	135.7	15,413.6	7.9
2013	4,188.9	5,573.6	1,729.5	3,244.8	1,439.3	151.8	16,327.9	5.9
2014	4,084.1	5,561.8	1,822.6	3,676.4	1,814.5	256.9	17,216.3	5.4
2015	4,010.8	5,543.0	1,855.5	4,218.5	2,180.0	547.1	18,354.7	6.6
2016	4,014.1	5,532.7	1,876.0	4,922.3	2,448.4	1,021.5	19,815.0	8.0
2017	3,942.3	5,113.3	1,811.9	5,250.9	2,676.6	1,267.0	20,061.9	1.2

자료 : Clarkson Shipping Review Database(2018)



자료 : Clarkson Shipping Review Database(2018) 바탕으로 작성

Fig. 2 세계 컨테이너선 선복량 추세(1997년 ~ 2017년)

이처럼 대부분은 선사들은 선복량 증대를 통해 선대 규모를 증가시켜 컨테이너 해운 시장에서 시장 점유율과 경쟁력을 확보하는 수단으로 활용하고 있다. 하지만 Cullinane et al(2000)은 컨테이너 선사들의 규모의 경제는 짧은 기간 동안 원가경쟁력의 가능성을 제공할 뿐이며 전 세계적으로 지속될 수 있는 경쟁우위는 여러 경쟁업체의 모방 전략으로 인해 실현 불가능하다고 예상하였다.

Table 2 세계 상위 컨테이너 선사 운항선복량 (2018년 4월 기준)

순위	선사	척	TEU	점유율 (%)	수주잔량	
					TEU	척
1	APM-Maersk	763	4,150,731	18.9	146,782	15
2	MSC	513	3,226,407	14.7	332,052	18
3	CMA CGM	498	2,527,286	11.5	277,552	20
4	COSCO	353	1,906,118	8.7	443,799	25
5	Hapag-Lloyd	220	1,578,966	7.2	-	-
6	ONE	235	1,506,507	6.9	125,766	9
7	Evergreen Line	196	1,088,509	5.0	482,330	45
8	OOCL	100	690,376	3.2	-	-
9	Yang Ming	108	646,382	2.9	70,000	5
10	PIL	134	408,511	1.9	72,738	8
11	Zim	87	373,171	1.7	-	-
12	Hyundai M.M	64	341,495	1.6	22,020	2
13	Wan Hai Lines	92	241,648	1.1	-	-
14	X-press Feeders	93	142,961	0.7	-	-
15	KMTC	60	126,848	0.6	16,200	8
16	Antong Holdings	102	126,119	0.6	34,564	23
17	Zhonggu Logi	97	124,310	0.6	29,074	13
18	SITC	76	104,071	0.5	15,666	10
19	IRISL Group	46	96,383	0.4	58,000	4
20	Arkas Line/EMES	42	72,717	0.3	12,400	4
21	TS Lines	34	72,544	0.3	5,616	4
22	Sinotrans	39	61,925	0.3	-	-
23	Sinokor	45	55,990	0.3	-	-
24	Salam Pasific	54	53,712	0.2	700	1
25	SM Line	14	53,542	0.2	-	-
상위 25개사 계		4,056	19,777,229	90.3	2,145,259	214
세계 100개사 합계		5,203	21,130,816	96.4	2,217,037	253

자료 : Alphaliner, 한국선주협회 2018년 해사통계집

2.1.2 세계 컨테이너 물동량 추세

컨테이너 물동량은 컨테이너 해운시장의 수요를 파악하는데 중요한 요인 중 하나이다. 컨테이너 물동량은 1999년 기준 전체 해상 물동량 중 약 9.2%를 차지하였으며 2017년 기준 전체 해상 물동량의 약 15.8%를 차지하는 것으로 나타났다.

Table 3 세계 컨테이너 물동량(1999년 ~ 2017년)

연도	컨테이너 (백만톤)	해상 물동량 합계 (백만톤)	비중(%)
1999	544	5,909	9.2
2000	609	6,347	9.6
2001	641	6,391	10.0
2002	693	6,572	10.5
2003	802	6,943	11.6
2004	911	7,424	12.3
2005	1,002	7,786	12.9
2006	1,092	8,136	13.4
2007	1,215	8,479	14.3
2008	1,272	8,684	14.6
2009	1,134	8,343	13.6
2010	1,291	9,120	14.2
2011	1,411	9,517	14.8
2012	1,458	9,917	14.7
2013	1,532	10,257	14.9
2014	1,622	10,607	15.3
2015	1,660	10,823	15.3
2016	1,734	11,150	15.6
2017	1,829	11,587	15.8

주 : 해상물동량 합계 = 철광석, 석탄, 곡물, 마이너벌크, 기타 건화물, 원유, 석유제품, 가스 포함
 자료 : Clarkson Shipping Review Database, Spring 2018 및 한국선주협회 2018년 해사통계집을 바탕으로 제작성

컨테이너 물동량과 컨테이너선 선복량은 수요와 공급요인으로 해운시장에 중요한 영향을 미치는 요인이다. 물동량의 경우 2008년 세계 경제위기 이후 2009년 약 12% 전년 대비 감소를 나타냈으며 이후 회복세를 거쳐 2012년 ~ 2017년까지 약 2% ~ 5% 정도 전년대비 상승하였다.

반면 선복량의 경우 지속적인 증가량을 나타내고 있다. 2009년 기준 컨테이너 물동량이 약 12% 전년대비 감소하였지만 선복량의 경우 약 11% 증가 추세를 나타내었다. 이러한 현상은 선복량 과잉현상이며 컨테이너 해운시장의 수급불균형을 나타내고 있다.

Table 4 세계 컨테이너 선복량 및 물동량 증감율(2008년 ~ 2017년)

연도	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
컨테이너 선복량 (백만TEU)	10.9	12.3	13.1	14.3	15.4	16.3	17.2	18.4	19.8	20.1
선복량 증감율	12%	11%	6%	9%	7%	6%	5%	6%	7%	1%
컨테이너 물동량 (백만톤)	1,272	1,134	1,291	1,411	1,458	1,532	1,622	1,660	1,734	1,829
물동량 증감율	4%	-12%	12%	9%	3%	5%	6%	2%	4%	5%

이처럼 지속적인 선복량 과잉공급과 수급조절이 어려운 원인으로는 첫째, 경쟁사와의 경쟁에서 우위를 확보하기 위한 경쟁적 대형선 발주, 둘째, 시장점유율 확대 및 대형선 증가에 따른 규모의 경제효과를 기대 하였으나 슬롯 비용절감과 더불어 운임감소를 동시 유발, 셋째, 금융시장의 낮은 금리로 인한 과잉공급의 확대 등을 들 수 있다. (박동주, 2017)

2.1.3 세계 주요 선사 얼라이언스 개편 및 M&A 추진 현황

보통 기업의 전략적 제휴의 목표는 일반적으로 규모의 경제, 위험분산, 상대 경쟁기업간의 경쟁우위 확보 등의 목적을 가지고 있다. 이와 마찬가지로 정기선사간 전략적 제휴는 해운동맹이나 컨소시엄 등의 다양한 형태로 물동량 집화 및 서비스항로 확보 등의 목적으로 발전해 왔다.(정원섭, 2001)

컨테이너 정기선 시장은 고정자산인 선박 및 컨테이너 등의 확보를 위하여 막대한 자산이 소요되는 반면, 해운시황에 따른 용선료 및 선가의 변화폭이 크기 때문에 다른 산업에 비하여 재무레버리지와 영업레버리지가 모두 높은 산업이라 할 수 있는데,(송채현 외1, 2007) 1984 미국 신해운법의 제정으로 해운동맹의 시장 지배력이 대폭적으로 약화되고 동맹과 비동맹선가간의 차별이 사라지면서 컨테이너 정기선사간의 경쟁은 더욱 치열해졌다.

경쟁심화의 원인으로서는 막대한 자산투자에도 불구하고 수요측면에서의 경기 침체와 공급측면에서 원가절감을 위하여 선박의 대형화 및 고속화와 각국의 해운업 지원 정책 등에 따라 공급이 수요를 초과하는 선복과잉 현상의 지속으로 운임수준이나 시장 환경이 더욱 열악해졌기 때문이다.

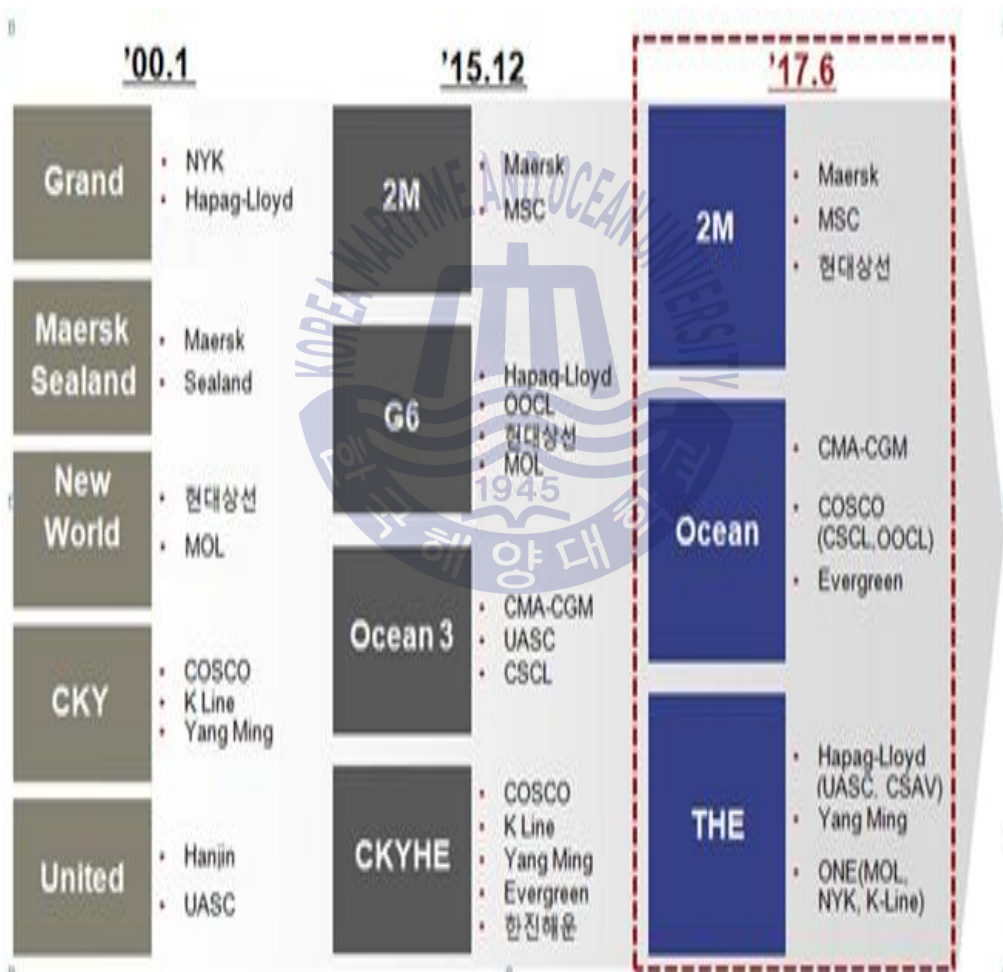
이러한 경쟁의 심화는 정기선사간 서비스질의 차이가 줄어들고, 시장이 거대화되고 다극화됨으로서 선사 단독으로 서비스영역을 확대하여 범세계적인 서비스망을 구축하는데 한계가 있기 때문에 독립적인 마케팅 및 영업활동을 허용하면서, 선박이나 컨테이너 기타 항만/하역장비 등을 공동으로 사용하여 서비스질의 향상을 도모하는 동시에 비용절감을 실현할 수 있는 컨테이너 정기선사간 전략적 제휴의 필요성이 증대되었다.(구근완, 2001)

따라서 선사의 독립적인 항로운항보다는 둘 또는 그 이상의 선사가 모여 전략적 제휴를 통하여 투자위험을 회피하고 규모의 경제를 통한 비용절감으로 치열한 경쟁시작에서 생존하는 방안을 추구하게 되었고 생존전략으로서 컨테이너 정기선사간 전략적 제휴가 보편화되었다.

즉, 컨테이너 정기선사간 전략적 제휴를 추진하게 된 근본 동기는 서비스영역, 영업망, 물류체제 등 선사 고유의 경쟁우위를 보다 광범위한 시장에서 보다 효율적으로 시행하거나 타 선사들의 우위요소들을 이용 또는 획득하려는 것으로 볼 수가 있으며 앞서 언급한 바와 같이 레버리지가 높은 산업특성에 따른 막대한 자산투자와 불확실한

시장 환경에서 경쟁에 따른 손실을 최소화하고 수익 증대를 위해서 추진되었다고 하겠다.(정원섭, 2001)

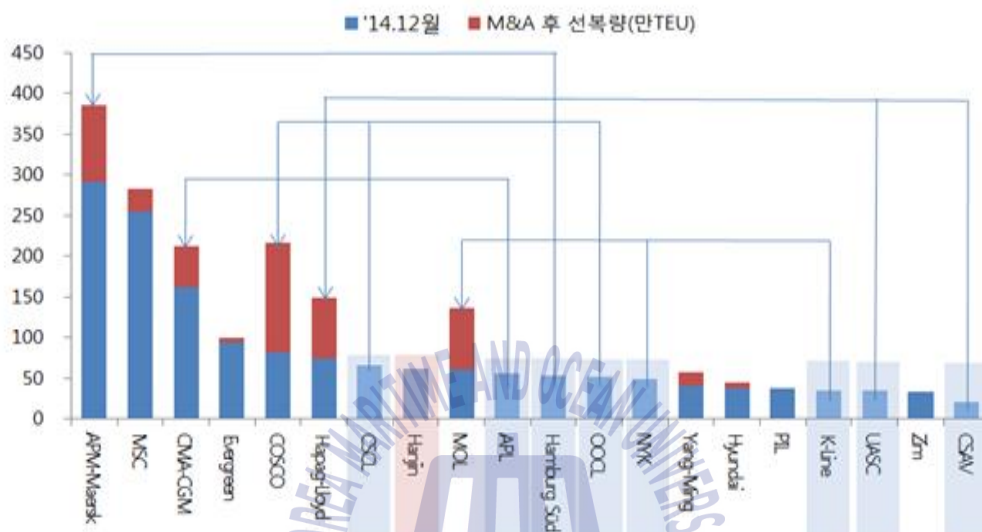
이처럼 선사간 경쟁우위 선점과 수익 증대를 위해 선사들은 얼라이언스를 체결하여 선복량 확보와 경쟁력 향상을 도모하였다. 2000년 Grand, Maersk Sealand, New World, CKY, United 총 5개의 얼라이언스는 2015년 2M, G6, Ocean 3, CKYHE 4개의 얼라이언스로 개편되었으며 2017년 2M, Ocean, THE 총 3개의 얼라이언스로 개편되었다. 기존의 5개였던 얼라이언스가 3개로 통합되며 차츰 과점형태가 강화되고 있다.



자료 : 해양수산부(2018), 해운재건 5개년 계획을 바탕으로 재구성

Fig. 3 Global Alliance 개편

또한 상위선사 중심의 적극적인 M&A로 상위 20개 선사가 11개 선사로 통합되었으며 이는 선복량을 최소 100만 TEU이상 확보하여 규모의 경제를 통한 원가절감 실현을 위한 조치로 파악된다.



자료 : 해양수산부(2018), 해운제건 5개년 계획을 바탕으로 제작됨

Fig. 4 세계 주요선사 M&A 추진현황

세계 1위 선사인 머스크라인은 2018년 2분기 매출액은 69억5200만달러(약 7조8100억원)를 기록, 전년 대비 25.5% 증가하였다. 반면 이자·세금·감가상각 전 이익(EBITDA)은 8억7600만달러에서 6억7400만달러(약 7600억원)로 23% 감소하였다. 2018년 상반기 EBITDA 역시 15.1% 감소한 11억5500만달러(약 1조3000억원)를 기록하였다.

현대상선은 2018년 2분기 영업이익이 -2094억원으로 전년 동기 -1237억원 대비 적자폭이 확대됐다. 순이익 역시 -2419억원으로 2017년도 같은 기간 -1865억원 대비 적자가 지속됐다. 매출액 역시 1조1854억원으로 전년 1조2324억원과 비교해 3.8% 감소했다. 2018년 상반기 매출액 역시 2조2590억원으로 전년 대비 9.1% 감소하였으며 영업이익은 -3856억원으로 적자가 지속된 반면, 순이익은 전년 대비 약 4600억원 개선된 -4164억원을 기록했다.

Table 5 주요 컨테이너선사 2018년 2분기 영업실적

선사	매출액	전년비 증감	영업이익	전년비 증감	순이익	전년비 증감
Maersk Line	69억 5,200만달러	25.5%	6억 7,400만달러	-23%	2,600만달러	흑자 전환
CMA CGM	57억 달러	7.4%	6,700만달러	-85.8%	3,300만달러	-85%
Hapage-Lloyd	28억 1,000만유로	15%	3,500만유로	-57.9%	-6,700만유로	적자 전환
Evergreen	382억 대만달러	2%	-20억 대만달러	적자 전환	-18억 대만달러	적자 전환
Yang-Ming	336억 대만달러	1.1%	-35억 대만달러	적자 전환	-38억 대만달러	적자 확대
현대상선	1조 1,854억원	-3.8%	-2,094억원	적자 확대	-2,419억원	적자 축소
OOCL	14억 6,200만달러	4%	-	-	-	-
Zim Line	8억 320만달러	7.7%	-1,380만달러	적자 전환	-3,320만달러	적자 전환
ONE	20억 6,600만달러	-	-	-	-1억 2,000만달러	-

주 : Maersk Line의 영업이익은 EBITDA(이자, 세금, 감가상각 전 이익), 순이익은 그룹전체 실적
 자료 : <http://www.ksg.co.kr>를 참고하여 제작성

Table 6 주요 컨테이너선사 2018년 상반기 영업실적

선사	매출액	전년비 증감	영업이익	전년비 증감	순이익	전년비 증감
Maersk Line	137억 6,200만달러	31.2%	11억 5,500만달러	-15.1%	27억 8,800만달러	흑자 전환
CMA CGM	111억 1,400만달러	11.9%	1억 5,500만달러	-79%	-3,400만 달러	적자 전환
Hapage-Lloyd	54억 2,500만달러	20%	8900만유로	-2.1%	-1억 유로	적자 확대
COSCO	432억 6470만 위안	1%	4억 2,770만위안	-74.1%	-1억 5,850만 위안	적자 전환
Evergreen	751억 대만달러	5%	-20억 대만달러	적자 전환	-18억 대만달러	적자 전환
Yang-Ming	646억 대만달러	1.8%	-56억 대만달러	적자 확대	-58억 대만달러	적자 확대
현대상선	2조 2590억원	-9.1%	-3856억원	적자 확대	-4,164억원	적자 축소
OOCL	28억 4,100만달러	9.6%	-	-	-	-
Zim Line	15억 5,460만 달러	11%	-2090만 달러	적자 전환	-6,730만달러	적자 확대

주 : Maersk Line의 영업이익은 EBITDA(이자, 세금, 감가상각 전 이익), 순이익은 그룹전체 실적
 자료 : <http://www.ksg.co.kr>를 참고하여 제작성

2.2 컨테이너 해운시장 관련 지수

2.2.1 컨테이너 해운시장 운임지수 도입 배경

현재 컨테이너 운임지수로 발표되고 가장 많이 활용되는 지수는 중국의 상해항운교역소(Shanghai Shipping Exchange)에서 발간한 SCFI(Shanghai containerized Freight Index)와 CCFI(China containerized Freight Index)가 있다.

하지만 세계 해운시장에서 주된 운임지수는 건화물 시장 중심이었다. 과거 건화물선 시장의 운임지수가 개발된 것은 해상운임선물거래(Freight Forward Agreement)가 도입된 것이 배경이 되었다. 해운에 선물거래가 등장한 것은 해운의 특징에서 비롯된다. 1970년 ~ 2006년까지 세계 해상물동량은 약 3배 증가하였다. 이와 같은 양적 팽창으로 해운시장에서 운임 리스크를 관리하는 방법으로 선박의 매도, 공동운항, 기간용선, 장기용선(또는 대선)을 하거나 장기계약(COA 등)을 하는 것외에는 특별한 방법이 없었다. (김태일, 정봉민, 2012)

또한 해운 리스크 관리를 위한 파생상품(Derivatives)의 도입이 어려웠던 것은 해운의 기초자산(Underlying asset)이 일반 상품과 달리 물리적 실체가 없을 뿐만 아니라 저장할 수도 없는 서비스라는 점으로 인해 파생상품화하기 어려운 내적 문제점이 존재하였다. (Alizadeh&Nomikos, 2009)

이러한 상황에서 1982년 미국 캔자스 상품거래소에서 Value-Line 지수를 대상으로 하는 상품을 시작으로 시카고 상품거래소에서 S&P500 지수 선물이 상장되어 거래가 시작되었다. 이 같은 주가지수선물거래의 도입에 따라 해운에서도 선물거래가 가능하게 되는 계기가 되었다. 처음으로 생겨난 해상운임선물거래는 1985년 5월 1일에 도입된 BIFFEX이다. 이것은 유로넷(Euronet LIFFE)의 일부인 런던거래소(London Commodity Exchange)에서 거래되었다. 그리고 이 거래에서 기초자산(Underlying asset)은 BFI(Baltic Freight Index)였다. (김태일, 정봉민, 2012)

최근 들어 컨테이너 시장에도 이와 같은 선물거래가 도입되고, 지수 개발이 활발해지고 있다. 이는 컨테이너선 시장도 운임리스크를 관리할 필요성이 대두되고 있고, 더 나아가 금융기법이 발달하면서 컨테이너 운임지수도 파생상품화 되고 있기 때문인 것으로 판단된다(김태일, 정봉민, 2012)

컨테이너선 관련 지수로는 크게 운임지수, 용선료지수, 기타지수로 나누어 볼 수 있다. 첫째, 운임지수는 상해항운교역소가 발간하고 있는 SCFI(Shanghai containerized Freight Index)와 CCFI(China containerized Freight Index)이 대표적이며 비교적 최근 발표된 WCI(The World Container Index)와 Drewry가 발표한 Drewry Freight rate Benchmark 가 있다.

둘째, 용선료 지수는 그 종류가 매우 다양하며 운임지수가 선물거래의 기초자산으로 활용되는데 비해서 용선료 지수는 시장의 수준을 파악하는데 주로 활용되고 있는 것으로 파악된다.(김태일, 정봉민, 2012) 용선료 지수로는 HRIC(Howe Robinson Containership Index), 함부르크 선박브로커협회(Hamburg shipbroker's Association)가 발간하는 Cotex(Containership time charter assessment Index), 그리고 영국선박 브로커 회사인 Braemar Seascope Ltd.)가 발간하는 Braemar BOXi Index 등이 있다.

셋째, 기타 지수로는 태평양항로 운임안정화 협정(Transpacific Stabilization Agreement)이 발간한 수입지수인 TSA's revenue index가 있다.

Table 7 세계 컨테이너선 시장 주요 관련 지수

구분	지수	발간일	발간 주체
운임 지수	<ul style="list-style-type: none"> ▪ SCFI(Shanghai containerized Freight Index) ▪ CCFI(China containerized Freight Index) 	2009.10.	상해항운교역소 (중국)
	<ul style="list-style-type: none"> ▪ WCI(The World Container Index) 	2011.07.	Drewry와 Cleartrade Exchange 합작
	<ul style="list-style-type: none"> ▪ Drewry Freight rate Benchmark 	-	Drewry
용선료 지수	<ul style="list-style-type: none"> ▪ HRIC(Howe Robinson Containership Index) 	1977.01.	Howe Robinson
	<ul style="list-style-type: none"> ▪ Cotex(Containership time Charter assessment Index) 	2009.02	Hamburg shipbroker's Association
	<ul style="list-style-type: none"> ▪ Braemar BOXi Index 	-	
기타 지수	<ul style="list-style-type: none"> ▪ TSA's revenue Index 	2008.06	Braemar Seascope Ltd.

자료 : 김태일, 정봉민, 2012. 컨테이너선 시장 가격지수 동향과 전망. 계간 해양수산 제5권을 바탕으로 제작성

2.2.2 SCFI(Shanghai containerized Freight Index) 및 CCFI(China containerized Freight Index)

SCFI는 2009년 10월부터 발간한 것으로 상해항을 기항하는 선사 15개사 및 17개 중국 포워더로부터 데이터를 수집하여 산출한다. 대표적인 선사로는 머스크라인(Maersk line), CMA CGM, 코스코(Cosco Container Lines), NYK 및 OOCL 등이 있다. 최초 2005년 12월 7일에 발간을 시작했으나 내용을 업데이트하여 새로운 SCFI를 2009년 10월 (=1,000 기준)부터 발표하고 있다.

운임은 20피트 및 40피트 컨테이너 기준이며, 매주 종합운임지수를 산출한다. 이 지수는 해상운임과 부대운임을 포함하며, 수출운임 기준이다. 따라서 상해발 주요 항로가 포함된다. 대표적인 항로로는 유럽, 지중해, 미국 서안, 미국 동안, 페르시아만, 호주·뉴질랜드, 서아프리카, 남아프리카, 남아메리카, 서일본, 동일본, 동남아시아, 한국, 대만 및 홍콩 등 15개 항로이다. 운임의 내용은 선형과 관계없이 산출된 운임이며, 유류할증료 등 각종 부대 운임은 포함되나 THC, 보안할증료(security surcharge), 내륙운송료(inland on-carriage surcharge) 등은 제외된다.

한편, CCFI는 중국 정부(교통부)가 지원하여 1998년 4월 13일(=1,000기준) 상해항운교역소에서 발간을 시작했다. 운임지수를 산출하는 주요 항로는 11개 항로인데, SCFI와 달리 일본항로가 서일본과 동일본 항로로 나뉘지 않고 통합되어 있으며, 페르시아만, 대만 및 홍콩 항로가 없다. 그리고 이 운임지수도 SCFI와 마찬가지로 중국발 수출 운임이나 주요 항만이 전 중국에 걸쳐 있다는 특징이 있다. 주요 기점은 다롄(Dalian), 텐진(Tianjin), 칭다오(Qingdao), 상하이(Shanghai), 난징(Nanjing), Ningbo), 샤먼(Xiamen), 푸저우(Fuzhou), 선전(Shenzhen) 및 광저우(Guangzhou) 등 10개 항만이다. SCFI와 또 다른 것은 SCFI가 선물거래의 기초자산으로 활용되는 데 반해, CCFI는 그렇지 않다는 점이다.(김태일, 정봉민, 2012)



자료 : 상해해운교역소(<http://www.sse.net.cn/home>)

Fig. 5 SCFI(Shanghai containerized Freight Index) 추이



자료 : 상해해운교역소(<http://www.sse.net.cn/home>)

Fig. 6 CCFI(China containerized Frieght Index) 추이

2.2.3 WCI(The World Container Index)

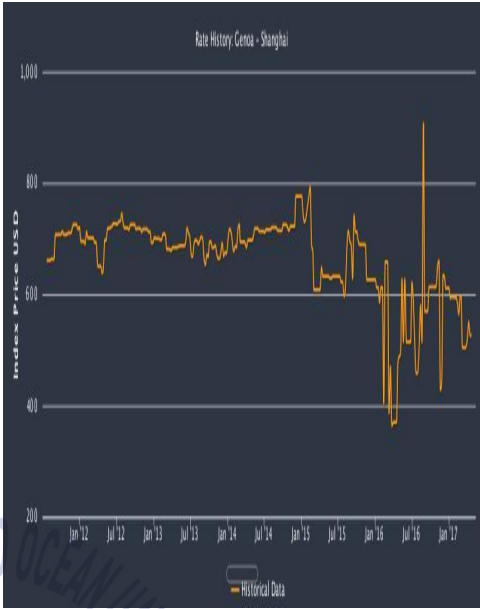
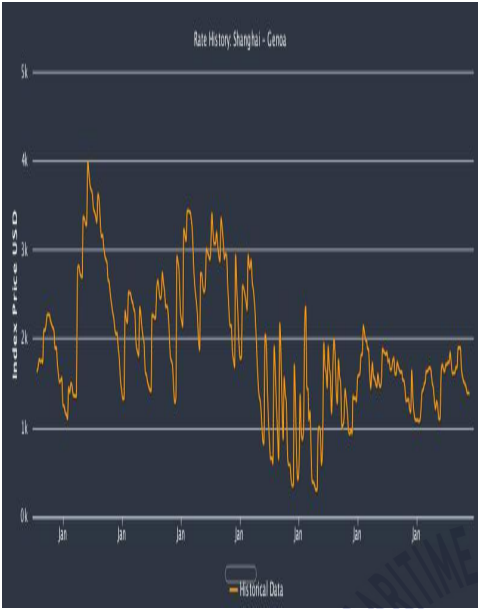
2011년 5월 5일 발표를 시작한 WCI(World Container Index)는 컨테이너선 동서항로(아시아, 북미, 유럽 간)에 대한 운임지수이다. 11개 주요 항로에 대한 운임지수로 종합 지수는 매주 발표한다. 11개 항로는 2010년 기준 총 물동량이 약 3,700만 TEU에 달하는 주요 항로로 볼 수 있다. 대표항로는 상해/로테르담, 로테르담/상해, 상해/제노아, 제노아/상해, 상해/LA, LA/상해, 상해/뉴욕, 뉴욕/로테르담, 로테르담/뉴욕, LA/로테르담, 로테르담/뉴욕 등이다.

WCI는 Drewry와 Cleartrade Exchange가 50:50으로 출자하여 설립한 ‘WCI Marketing Services’ 에서 발간한다. 이 지수는 컨테이너선 운임파생거래에 활용되며, 이러한 측면에서 SCFI와 마찬가지로 기능을 한다고 볼 수 있다. 운임의 내용은 SCFI와 다소 다른 측면이 있다. SCFI는 THC가 포함되지 않으나 WCI는 포함 여부가 항로마다 다르다. 또한, 유류할증료 등 각종 부대운임을 포함하고, 내륙운송비가 포함되지 않는 것은 SCFI와 기본적으로 같다. 그러나 부대운임의 경우 SCFI가 미국 항만보안 문서료(Documentation fee)와 미국의 사전화물정보 통보료(Automated Manifest Fee)를 제외하고 있는데, 이러한 점은 WCI와 차이가 있다고도 볼 수 있다. 또한 SCFI는 미국 서안 및 동안 항로의 경우 달러/FEU 기준이고 나머지 항로는 달러/TEU이나 WCI는 모든 항로가 달러/FEU 기준이다.(김태일, 정봉민, 2012)

Table 8 WCI(The World Container Index)

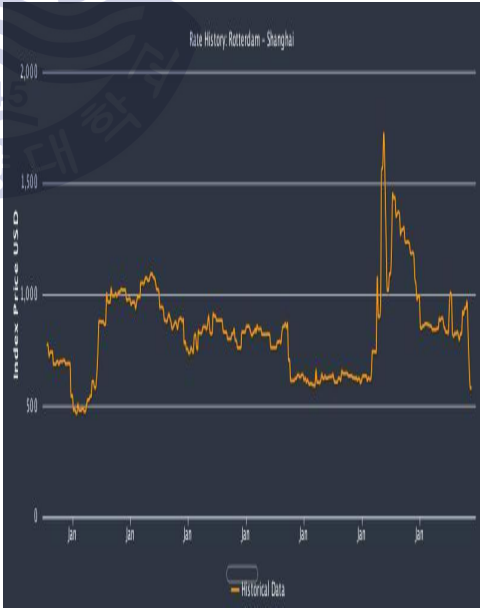
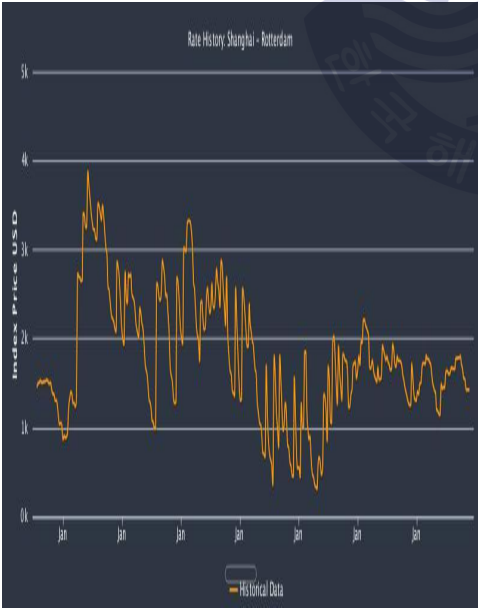
		04-OCT-18	11-OCT-18	18-OCT-18	25-OCT-18	Weekly Trend
Composite Index		1668.47	1635.87	1666.15	1666.1	-0.06
Shanghai-Rotterdam	USD/FEU	1433	1424	1427	1427	0
Rotterdam-Shanghai	USD/FEU	808	653	578	578	0

Shanghai-Genoa	USD/ FEU	1456	1391	1392	1378	-14
Genoa-Shanghai	USD/ FEU	0	0	0	0	
Shanghai-LosAngeles	USD/ FEU	2382	2375	2590	2621	31
LosAngeles-Shanghai	USD/ FEU	505	505	518	505	-13
Shanghai-NewYork	USD/ FEU	3569	3534	3386	3348	-38
LosAngeles-Rotterdam	USD/ FEU	0	0	0	0	
Rotterdam-LosAngeles	USD/ FEU	0	0	0	0	
NewYork-Rotterdam	USD/ FEU	562	561	567	558	-9
Rotterdam-NewYork	USD/ FEU	2009	2037	1970	1956	-14



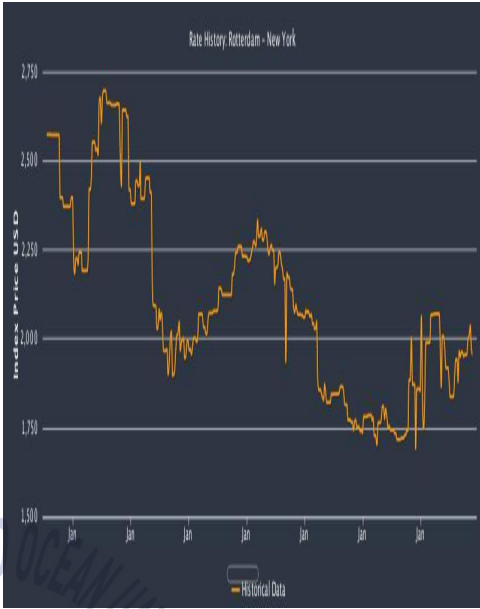
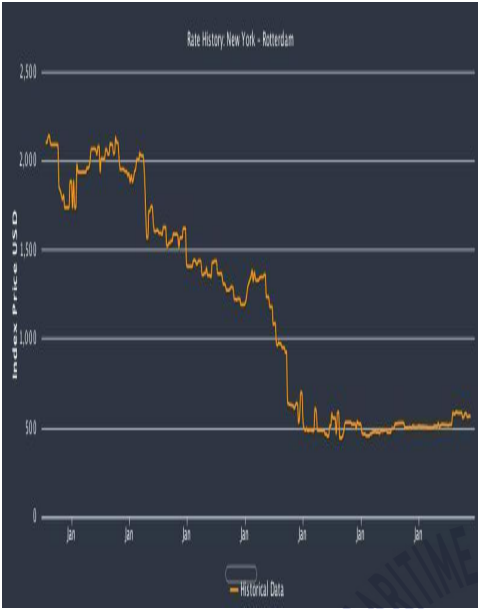
자료 : <http://www.worldcontainerindex.com>

Fig. 7 Shanghai ↔ Genoa WCI(The World Container Index) 추이



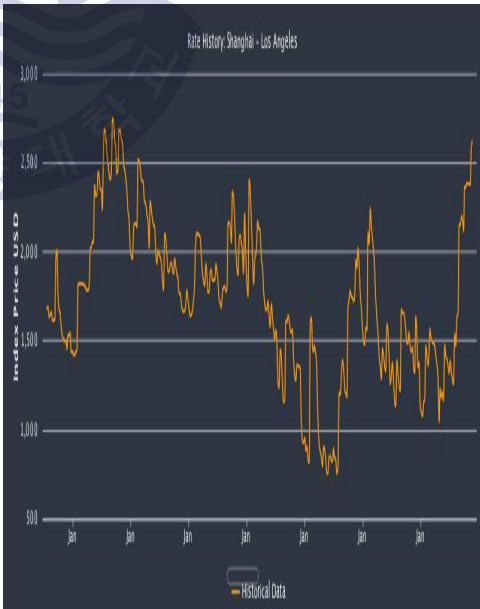
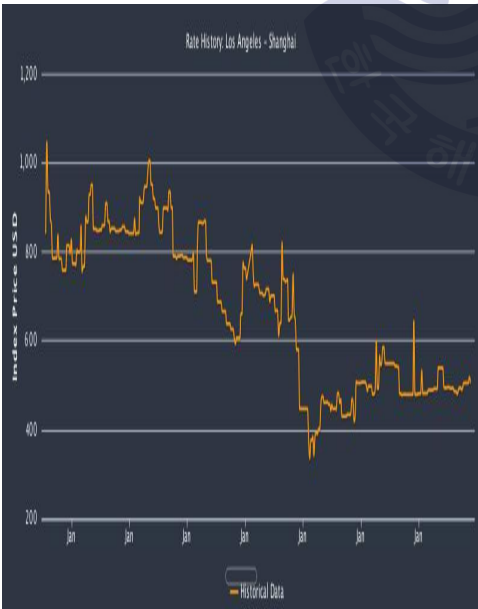
자료 : <http://www.worldcontainerindex.com>

Fig. 8 Rotterdam ↔ Shanghai WCI(The World Container Index) 추이



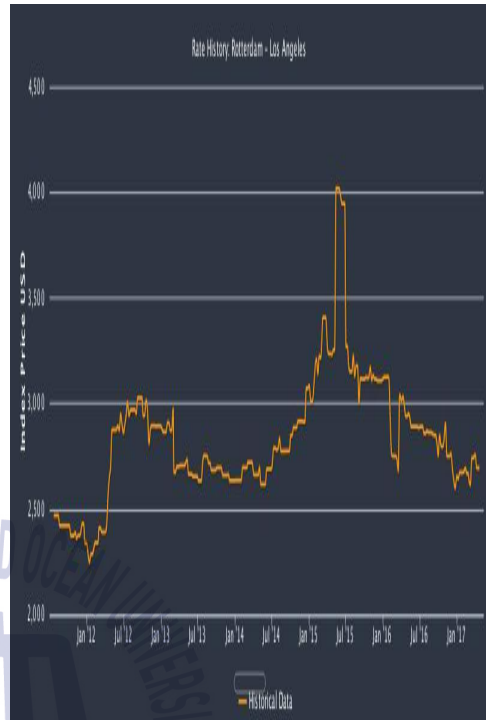
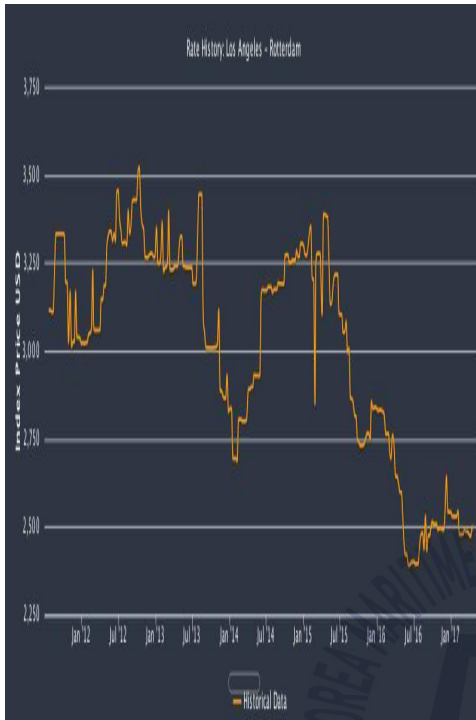
자료 : <http://www.worldcontainerindex.com>

Fig. 9 Rotterdam ↔ New York WCI(The World Container Index) 추이



자료 : <http://www.worldcontainerindex.com>

Fig. 10 Shanghai ↔ Los Angeles WCI(The World Container Index) 추이



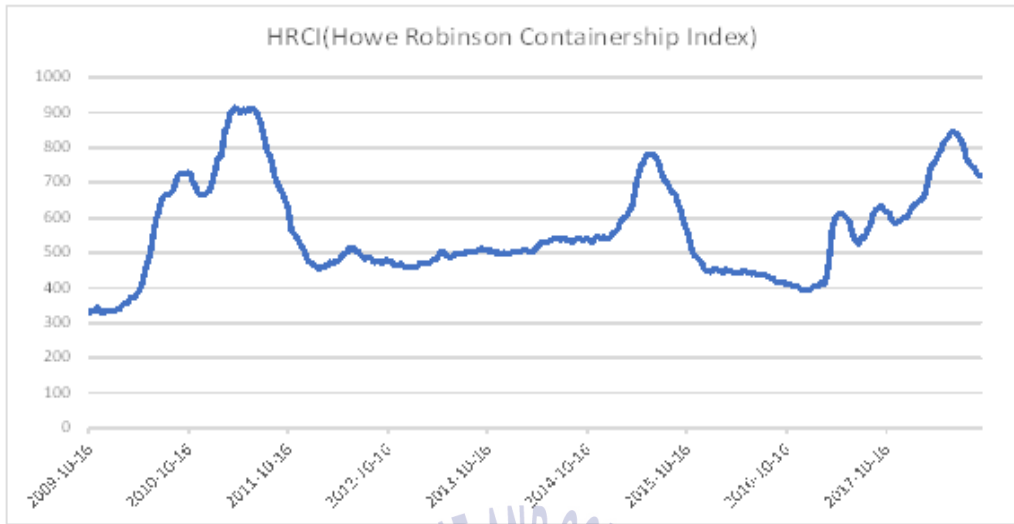
자료 : <http://www.worldcontainerindex.com>

Fig. 11 Rotterdam ↔ Los Angeles WCI(The World Container Index) 추이

2.2.4 HRCI(Howe Robinson Containership index)

HRCI는 영국 런던 브로커인 Howe Robinson & Co. Ltd.가 발간하는 컨테이너선 용선료 지수이다. 510TEU~4,300TEU의 14개 주요 대표 선형으로 이루어져 있으며, 1997년 1월=1,000을 기준으로 한다.

1997~1999년 아시아 금융위기, 러시아 및 브라질 경제위기에 따른 세계 경제침체로 지수가 하락했으며, 1999년 2월 HRCI는 기준 지수보다 40% 하락한 595.3포인트를 기록했다. 한편, 2000년 9월 20일 지수는 953.1 포인트를 회복했으나 2001~2002년 HRCI는 다시 침체시기를 겪고, 2002년 1월 16일 451.4 포인트를 기록했다. 그리고 중국 효과가 나타나기 시작한 2003년 이후 상승세를 보이다가 2008년 금융위기 이후 최근까지 낮은 수준을 유지하고 있다.(김태일, 정봉민, 2012)



자료 : Lloy' s list를 바탕으로 재구성

Fig. 12 HRCI(Howe Robinson Containership Index)추이

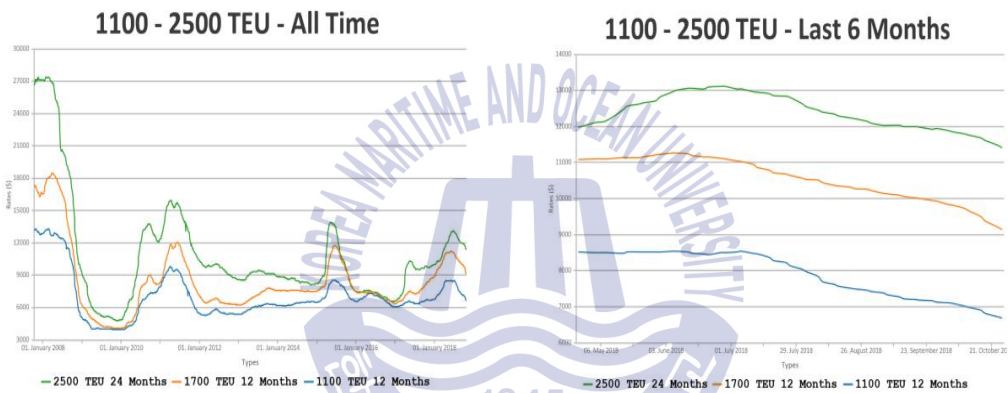
Table 9 HRCI의 대표 선형 및 가중치

구분	TEU	DWT	기어장착 여부	속력(노트)	가중치(%)
Class 1	510	6500	No	15.0	2.5
Class 2	520	5500	Yes	15.5	2.5
Class 3	650	9500	Yes	15.0	5.0
Class 4	1000	12000	Yes	17.5	2.5
Class 5	1000	13750	No	19.0	7.5
Class 6	1000	18500	Yes	19.0	7.5
Class 7	1200	17000	No	19.0	10.0
Class 8	1600	25000	No	18.0	12.5
Class 9	1700	24000	Yes	19.5	12.5
Class 10	2080	30000	No	21.5	7.5
Class 11	2500	34000	Yes	22.0	10.0
Class 12	2800	35000	No	22.0	10.0
Class 13	3500	45000	No	22.5	5.0
Class 14	4300	55000	No	24.0	5.0

자료 : 김태일, 정봉민, 2012. 컨테이너선 시장 가격지수 동향과 전망. 계간 해양수산 제5권을 바탕으로 재작성

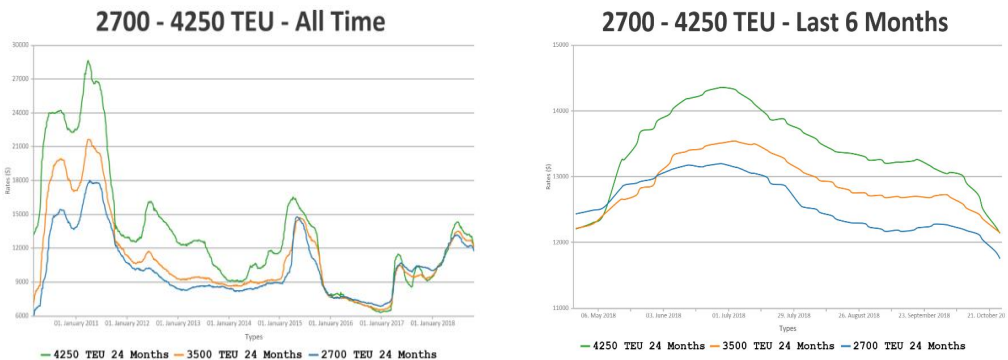
2.2.5 Contex(Container ship time charter assessment Index)

Contex는 독일 함부르크 선박브로커협회(Hamburg Shipbrokers Association: VHSS)가 발간하고 있는 용선료 지수이다. 이 지수는 2007년 10월 11일 1,000을 기준으로 시작되었다. 이 지수는 초기 3개 선형에 대한 용선료 수준을 나타내는 지수였는데, 2010년 5월 11일부터 2,700TEU, 3,500TEU, 4,250TEU 등 3개 선형이 추가되어 모두 6개 선형으로 구성된 지수가 발표되고 있다. 이러한 지수의 산정에는 60여 명의 선박브로커가 참여하고 있는 것으로 알려졌다. 브로커들의 국적은 독일, 덴마크, 영국, 프랑스 등 다양하다.(김태일, 정봉민, 2012)



자료 : Hamburg Shipbrokers Association

Fig. 13 Contex(1,100 ~ 2,550 TEU) All time/Last 6 Months 추이



자료 : Hamburg Shipbrokers Association

Fig. 14 Contex(2,700 ~ 4,250 TEU) All time/Last 6 Months 추이

제3장 선행연구 및 예측방법론 고찰

3.1 선행연구 고찰

정기선 운임지수 및 운임에 영향을 미치는 요인에 관련한 기존 연구들의 경우 부정기선 부문에 비해 상대적으로 미시적인 요인들에 중점을 두고 있으며 예측에 관한 연구 또한 상대적으로 적다고 할 수 있다. 이는 정기선과 부컨테이너 해운시장의 운임측면의 차이점에서 비롯된 것으로 파악된다. 이에 기존 선행연구 검토를 통해 정기선 운임지수 및 운임에 영향을 미치는 요인을 검토한 결과 다음과 같다.

Table 10 주요 선행연구 결과요약

구분	분석방법론	선정변수
박호건 외 1(2002)	구조방정식을 통한 경로분석	수요-공급요인(물동량, 선복량 변화) 선사적요인(선사 영향력 및 경쟁정도) 해운원가요인(자본비, 선원비, 연료비) 기타요인(해운정책, 항만, 항로)
이현호(2013)	다중회귀 분석	국가별 GDP, 환율, 유가, 생산자 출하지수, 제조업 가동률, 선박 수요 가용성, 선박 공급 가용성, 적재율
강효원 외 2인(2014)	회귀분석	운항 항로간 거리, 컨테이너 시장 상황, 경제 발전 정도
박동주(2017)	단위근 검정 벡터자기회귀모형	HRCI, Time Charter Index, Bunker 380 CST, Orderbook, CNTRSHIP Fleet
김석환(2017)	회귀분석	Bunker 380 CST, Clarkson Index, Time charter Index

박호건 외 1(2002)은 정기선 운임에 미치는 네 가지의 주요 요인으로 연구모형을 설정하고, 설문조사를 통하여 이에 대한 구조방정식에 의한 경로 분석을 하였다. 네 가지의 주요 요인은 첫째, 수요-공급요인(물동량과 선박량 변화), 둘째, 선사적 요인(운임 결정에 대한 선사의 영향력과 선사들간의 경쟁정도), 셋째, 운임결정에 해운원가요인(선박 자본비, 선원비, 연료비), 넷째, 회부적 기타요인(정부의 해운정책, 항만, 항로적 요인)이다. 네 가지 요인은 모두 해상운임 결정에 타당한 것으로 평가 되었다.

이현호(2013)는 컨테이너 해상운임 예측모형 연구에서 한국을 포함한 극동지역 중심의 국제 해운시장에 있어서 정기선 운임에 영향을 미치는 요인들 즉 국가별 GDP, 환율, 유가, 생산자 출하지수, 제조업 가동률, 선박 수요 가용성, 선박 공급 가용성, 적재율 등을 분석하여 독립변수로 찾아내고 다중회귀 분석을 통하여 가장 높은 결정계수 값을 찾아 적정운임을 예측할 수 있는 모델링을 구축하였다. 이에 대해 적정성 유무 판단결과, 조정된 결정계수로 예측된 모형이 CCFI의 트렌드를 재현 한다고 볼 수 있으며, 분산분석을 통한 F값도 유의한 F인 값보다 크므로, 귀무가설(H0)인 독립변수들이 시장운임(CCFI)과 연관성이 없음을 기각하여 유의하다고 판단 할 수 있는 것이다. 특히나 유럽항 모델링의 결정계수 0.91이 북미항보다 높은 것으로 도출되어 보다 최적화된 모델링인 것으로 나타났다.

Gouveral et al.(2012) 는 컨테이너 해상운임에 영향을 미치는 운항 항로간 거리, 컨테이너 시장 상황, 경제 발전 정도 세 가지 요인들과 해상운임 사이의 관계에 대하여 실증 분석을 하였다. 이 중, 해상운임과 운항 항로간 거리 사이의 상관관계에 대한 분석 결과로는 연구에 사용된 2007년, 2008년, 2009년의 운임 자료 중 2007년의 자료만이 운항 항로의 거리와 해상운임간의 정(+)의 상관관계를 보였으며, 2008년, 2009년의 자료는 상관관계가 없는 것으로 분석 되었다.

강효원 외 2인(2014)은 HR 지수, Time Charter 지수 등 정기선 해상운임에 영향을 미치는 요소들에 대한 실증분석을 수행하였다. 이 연구를 통해 HR 지수와 Time Charter 지수의 높은 유사성이 있다는 점, 중고선가는 HR 지수, Time Charter 지수와 부(-)관계를 보인다는 점, 신조선계약과 HR 지수, Time Charter 지수와는 정(+)의 관계가 있는 점을 보여 주었다.

박동주(2017)는 SCFI 운임지수, HRCI, Time Charter Index, Bunker 380CST, Orderbook, CNTRSHIP Fleet의 월간 시계열 자료를 단위근 검정과 벡터자기회귀모형, 충격반응분석, 예측오차 분산분해를 통해 실증 분석을 하였다. 실증 분석 결과 종속변

수 SCFI 종합지수에 영향을 미치는 공급 변수의 비중은 Orderbook이 2.7573%, Time charter Index가 1.5598%, CNTRSHIP Fleet이 1.1603%, HRCI가 0.8356%, Bunker가 0.0594%로 나타났다. SCFI 종합지수의 예측오차 분산을 설명하는데 가장 큰 비중을 차지하는 변수는 Orderbook 변수로 나타났다.

김석환(2017)은 정기선 컨테이너 선사들의 해상운임 결정에 영향을 미치는 요인들에 대한 문헌 연구와 조사를 바탕으로 변수를 선정하고 각 변수들과 본 연구를 위하여 수집된 실제 컨테이너 해상운임 사이의 상관관계에 대한 분석을 하였다. 연구 결과 실제 해상운임 자료들과 Bunker 비용과 용선비용 지수인 Clarkson index와 Time Charter index와의 상관관계를 설명해주는 요소들을 발견하지 못하였다. 최근 컨테이너 선사들의 운임 결정에 미치는 요인으로는 Bunker 비용과 용선비용이 고려되고 차지하는 비중이 상대적으로 크지 않았다는 것으로 나타났다.



3.2 분석방법론

3.2.1 ARIMA(Auto-regressive Integrated Moving Average)

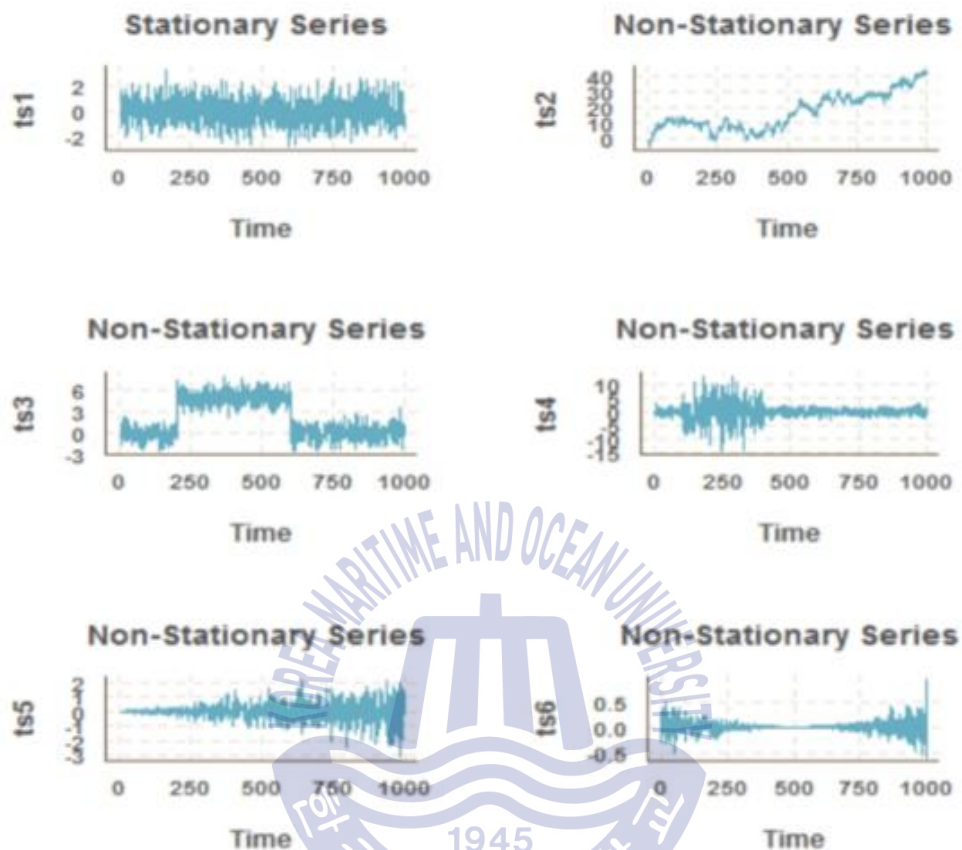
시계열 분석에서 가장 많이 활용되고 있는 모형인 자기회귀이동평균(ARIMA : Auto-regressive Integrated Moving Average) 모형은 Box and Jenkins(1976)에 의해 이론적으로 체계화 되었으며 선형시스템에 기반을 둔 통계학적 분석 기법으로 과거의 많은 연구들을 통해 그 우수성이 입증 되었다. 이는 해당 시계열의 과거 경향에 기초한 확률과정에 대하여 특정한 모형을 설정한 뒤 미래 경향을 설명하고 예측하는 기법으로 모든 형태의 단변량(univariate) 시계열 분석에 적용가능 하다.

ARIMA(Auto-regressive Integrated Moving Average) 모형은 시계열 데이터 기반 분석 기법으로 과거지식이나 경험을 바탕으로 한 행동에 따라 경제가 움직이고 있음을 기초로 한다. 시계열 분석은 주가 전망이나 수요 예측 등에 사용되며 매우 다양한 기법으로 발전하였으나, 계량경제모델이 점점 대형화되면서 오히려 경제의 본질적인 부분을 제대로 설명하기 어려워졌다.

이러한 상황을 극복하기 위해 소수의 시계열 데이터만을 이용한 단순 모델로 경제 분석 및 예측을 시도하는 움직임이 최근 활발히 진행되고 있다. ARIMA 모형은 과거의 관측 값과 오차를 사용해서 현재의 시계열 값을 설명하는 ARMA(Auto-regressive Moving Average) 모델을 일반화 한 것으로, 분기/반기/연간 단위로 다음 지표를 예측한 다거나 주간/월간 단위로 지표를 리뷰하며 트렌드에 이상치가 없는지를 모니터링 하는데 사용되는 분석 기법이다.

ARMA 모델이 안정적 시계열(Stationary Series)에만 적용 가능한 것에 비해, 분석 대상이 다소 불안정적인 시계열(Non Stationary Series)의 특징을 보여도 적용이 가능하다.

안정적인 시계열이란 시간의 추이와 관계없이 평균 및 분산이 불변하거나 시점 간의 공분산이 기준시점과 무관한 형태의 시계열이다. 시계열이 안정적이지 않을 때는 로그를 이용하거나 차분을 통해 시계열을 안정적으로 변환한 뒤에 분석을 진행한다. 다음 Fig. 15는 안정적인 시계열과 불안정적인 시계열의 몇 가지 예시이다.



자료 : <https://stats.stackexchange.com/questions/182764/stationarity-tests-in-r-checking-mean-variance-and-covariance>

Fig. 15 안정적 시계열 및 불안정적 시계열

ARIMA모형은 크게 1단계 모형의 식별, 2단계 모수 추정, 3단계 모델 적합, 4단계 예측으로 총 4단계로 구분이 되는데, 1단계 모형의 식별에서 시계열의 정상성 여부를 확인하게 되며, 보통은 그래프를 먼저 그려보고 추세선이 존재하는지, 상수항이 존재하는지(즉 시계열 변수의 평균값이 0이 아닌지) 확인한다.

이후 ADF 검정법으로 단위근을 검정하거나, 혹은 자기상관함수나 편자기상관함수가 줄어드는 양태를 보고 단위근 여부를 판단하거나, 혹은 그냥 그래프가 추세, 순환, 계절, 불규칙 변동을 보이는지 눈으로 확인하고 정상성 여부를 확인하게 된다.

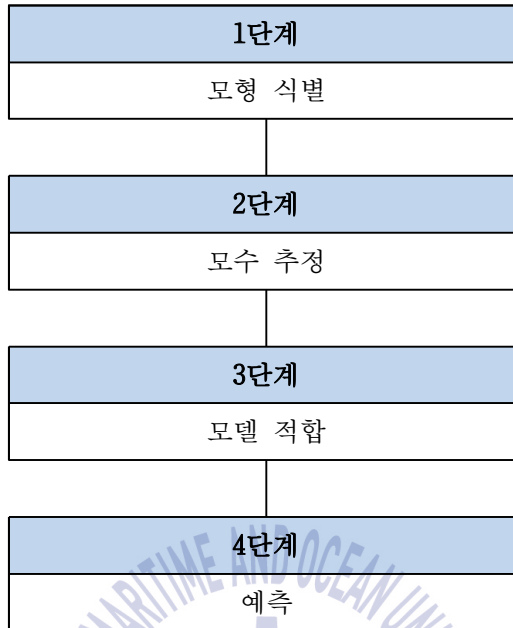


Fig. 16 ARIMA 모형 분석 단계

정상성 확인이 중요한 이유는, 일반적으로 선형확률과정은 임의의 표본공간으로부터 시계열 값들이 선출되면서 실현된 것이라고 보기 때문이다. 좀 더 엄밀히 말하면 우리가 구한 시계열 자료는 시간에 따른 확률분포에서 실현된 값들의 나열이다. 그런데 시간이 흐름에 따라 확률분포가 크게 변동한다면 그 실현값들의 평균이나 분산 등 모멘트가 의미가 없기 때문에 적어도 이 모멘트들이 우리가 다루고자하는 확률과정을 설명하기에 문제가 없도록 하기 위하여 필요한 조건이 정상성 조건이다.

마찬가지로 미래의 수치도 이런 방법으로 추측이 가능한데, 문제는 현실사회에선 이상적인 그래프는 존재하지도 않고, 어디까지나 모수치는 없고(가정은 하지만) 과거값인 표본들만 존재하는 상황이라는 데 있다. 즉, 유한개의 관찰된 시계열 자료로 추정을 하려면 필연적으로 그래프를 단순화 시키는 가정이 필요하다.

그래프를 정상화 시키는 가장 효율적인 방법은 시계열 변수들을 백색화(Whitening)시키는 것이다. 백색잡음으로 구성된 항들은 평균은 일정하고, 분산은 일정하며, 변수들간 공분산과 자기상관은 시점에 의존하지 않고(즉 특정 시점에 갈수록 커지거나 작아지지 않음) 어디까지나 각 변수들의 시점의 차이인 “시차“에만 의존한다는 정상성 조건을 만족하게 된다.

변수들을 “백색잡음“으로 만드는 가장 효율적인 방법은 바로 차분(differencing)을 이용하는 방법인데, 바로 현재 변수에서 바로 전 차수의 변수를 빼줌으로서 가능하다. 단위근이 존재한다는 것은 $Z_t = aZ_{t-1} + e_t$ (단, e_t 는 백색잡음)에서 a 가 1이라는 뜻인데, a 가 1이 되어버리면 $Z_t = Z_{t-1} + e_t$ 가 되어 분산이 일정 값으로 수렴하지 않고 무한히 발산하게 된다. 즉, $Z_t = Z_{t-1} + e_t$ 에서 Z_{t-1} 을 좌변으로 이항하여 $Z_t - Z_{t-1} = e_t$ 가 되어버린다면 이 변수는 백색잡음(e_t)가 되면서 자연스럽게 정상성을 만족하게 된다.

이렇게 시계열 정상화가 끝나면 다음은 ARIMA(p,q,r)의 차수를 결정해야 하는데, 차수는 AR(p) 모형과 MA(q)모형이 각각 몇 차시까지 상관성을 갖느냐를 나타내는 수치이다. 가령, ARIMA(2,1,2)라면, 이 시계열은 1차 차분이 완료된($q=1$) 그리고 자기상관모형인 AR항이 2차시까지 ($Y_t = aY_{t-1} + bY_{t-2} + e_t$) 가며, MA항이 2차시($Y_t = e_t - c_{t-1} - d_{t-2}$)까지 존재하는 모형임을 나타낸다. 식으로 표현하면 다음과 같다.

$$Y_t = aY_t + bY_{t-2} - c_{t-1} - d_{t-2} + e_t$$

적정 차수를 추정하는 방법은 첫째, 자기상관함수(ACF)와 부분자기상관함수(PACF)가 줄어드는 양태를 보고 적정 차수를 결정하는 방법과 둘째, AIC, BIC등 각종 모형의 적합도를 나타내는 수치를 확인하고 적합도 수치가 가장 적절한 모형을 찾는 방법 두 가지가 존재한다.

ACF와 PACF로 차수를 결정하는 방법은 방법론에 익숙해지면 직관적이거나 그 양태가 그래프에서 뚜렷히 나타나는 경우가 생각보다 적고, 또 MA(q)항을 추정하는 과정에서 그 오류가 누적적으로 커지기 때문에[9] 0이 아닌데도(귀무가설이 아닌데도) 0으로 간주하는 1종오류를 범할 확률이 매우 높아진다. 따라서 보통은 ACF와 PCAF가 줄어드는 양태 + 적합 후 나타나는 적합도 수치의 결과를 보고 타당한 모형을 지속적으로 적용해야 한다.

모수를 추정하는 방법은 대개 최대우도 추정법(MLE), 최소자승추정법(OLS)이 쓰인다. 적률추정법은 가장 간단하긴 하나 그 예측치가 효율적이지 않아서 자주 쓰이지 않는다. 좀 더 정확히 말하면 이동평균 component가 들어가면 OLS를 쓸 수 없다. AR process를 Wold theorem과 iteration을 통해 MA process로 바꾸고 MLE를 쓰고 이를 다시 ARIMA로 바꾸는 것이다. 그런데 계산량이 많아서 특정한 점에서 OLS를 하고 이 점의 근방에 이 점보다 더 큰 Likelihood를 주는 점은 없는지 확인하는 수치적 방법을 더 많이 활용한다.

3.2.2 자기회귀모형

회귀분석(Regression Analysis)이란 어떤 현상이 변수들의 인과관계에 의하여 나타날 때 그 관계를 수학적으로 설명하기 위하여 사용되는 통계적 방법을 말한다. 회귀분석에서 사용되는 변수는 종속변수와 독립변수가 있다. 종속변수(dependent variable)는 서로 관계를 가지고 있는 변수들 중에서 다른 변수에 의해 영향을 받는 변수를 말하며, 독립변수(independent variable)는 종속변수에 영향을 주는 변수를 말한다. 회귀분석은 이러한 종속변수와 독립변수들 간의 함수관계를 규명하기 위하여 어떤 수학적 모형을 설정하고, 주어진 자료로부터 모형을 추정하거나 검정하는 방법이다.

회귀분석에서 얻어진 모형의 추정식을 회귀방정식(Regression equation)이라 하고, 이 식에 대한 통계적 유의성검정을 실시하거나 이 식을 이용하여 종속변수 값에 대한 통계적 예측을 실시하게 된다.

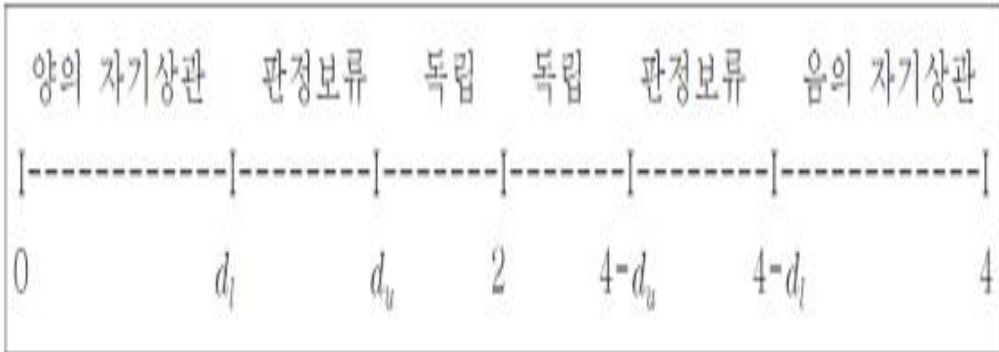
회귀분석에서 가장 간단한 모형은 단순선형회귀분석(simple linear regression analysis)이다. 여기서 ‘단순’은 독립변수가 한 개임을 의미하며, ‘선형’은 회귀식이 독립변수의 1차식으로 표현됨을 의미한다. 이 경우 회귀식은 다음과 같이 나타내어질 수 있다.

$$Z_t = \beta_0 + \beta_1 x_t + \varepsilon_t, \varepsilon_t \sim N(0, \sigma^2) \text{이며 서로 독립, } t = 1, 2, \dots, n$$

여기서 ε_t 는 t번째 측정치의 오차(error)이고, β_0 와 β_1 은 미지의 회귀계수(regression coefficient)로서, 이것은 최소제곱법(method of least squares)로 추정한다. 오차항에 대해서는 위의 식에서 보는 바와 같이 정규성(normality), 불편성(unbiasedness), 등분산성(equal variance), 독립성(independence)을 일반적으로 가정하고 있다.

만약 주어진 자료들이 시계열적 특성을 띠고 있어 오차항들 사이에 자기상관이 존재한다면 회귀분석에서 가정하는 독립성을 위배하게 되므로 추정된 회귀식의 수정이 필요하다. 오차항들 사이에 자기상관이 존재하는 것은 Durbin-Watson 통계량(d)를 이용하여 검정할 수 있다.

Durbin-Watson은 검정통계량 d의 임계값으로 상한값 d_u 와 하한값 d_l 을 사용하여 Fig. 17과 같은 판정을 제안했다. 이 때, 임계값 d_l , d_u 는 자유도와 독립변수의 수에 의하여 결정된다. Durbin-Watson의 d-통계량의 판정기준은 독립변수가 1개이고 데이터 수가 100개 이상일 경우 유의수준 0.05에서 $d_l=1.65$ 이며, d_u 는 1.69이다.



자료 : 하정아, 2006. 자기회귀분석과 시계열분석을 이용한 교통량 예측자료 보정 알고리즘

Fig. 17 Durbin-Watson 검정의 구간별 판정

검정결과 오차항들간에 자기상관이 존재한다면 오차항을 시계열모형을 따르는 확률변수로 생각할 수 있고, 이 확률변수를 ε_t (t=시차)라고 한다면 위의 회귀식을 다음과 같이 변환하여 생각할수 있으며 이와같은 모형을 자기회귀모형이라 한다.

$$Z_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$

$$\text{여기서, } \varepsilon_t = \rho\varepsilon_{t-1} + \delta_t, \delta_t \sim N(0, \sigma_\delta^2), \text{Cov}(\delta_t, \delta_{t+i}) = 0$$

단순선형회귀분석을 하기 전에 선행되어야 할 분석방법으로 상관분석을 들 수 있다. 상관분석은 두 양적 변수간의 선형관계를 분석하는 것이다. 일반적으로 두 변수간의 직접적인 선형관계를 재는 상관성의 척도로 Pearson의 단순상관계수(Pearson's simple correlation coefficient)를 이용하며, 단순상관계수를 이용하여 두 변수간의 상관관계 유무를 검정한다.

두 양적 변수간의 상관계수가 높다면 하나의 변수는 다른 변수의 1차식으로 표현될 수 있을 것이다. 표현되는 방정식은 두 변수간의 단순선형회귀분석의 회귀방정식이며, 상관분석결과 상관계수가 높아 두 변수 간에 선형관계가 있다고 분석이 되면, 단순선형회귀분석을 이용하여 두 변수간의 관계식을 설정하여 예측모형에 이용하게 된다.

3.2.3 인공신경망 모형

인공신경망은 패턴인식(pattern recognition), 제어(control), 화상처리(image processing) 등의 여러 분야에서 기존의 방법론으로써 해결하기 어려웠던 부분들을 해결해주거나 성과를 제고시키는 새로운 기법으로 각광을 받고 있다. 인공신경망의 출현은 상당히 긴 역사를 가지고 있지만 최근 이에 대한 연구가 폭발적으로 증가했고 이로 인해 다양한 인공신경망 모델들이 등장하였다. 최근에는 인공신경망의 임의의 함수에 대한 근사(approximation)능력과 일반화(generalization) 능력에 기초하여 예측(forecasting)과 추정(estimation)등의 분야에서도 많이 응용되고 있다.(Toll, 1996)

인공신경망은 인식이나 의사결정 등에 있어 생물학적 신경세포의 특성을 모델링 하여 프로그래밍 한 것으로서 단순한 프로그램에 의해서 순차적으로 수행하는 컴퓨터를 보완하기 위한 것이다. 인공신경망은 뉴런이라고 불리는 인간의 신경세포와 유사한 처리소자들의 병렬분산처리에 의하여 학습 및 예측을 수행하며 대표적인 특성은 다음과 같다.(이철욱, 1993, 나상민, 1996, 손창우, 2004)

첫째, 각 신경세포는 다른 신경세포들과 완전히 독립된 기능을 갖는다. 즉, 각 신경세포의 출력은 자신과 연결된 연결강도를 통하여 직접 전달 되는 정보에만 의존할 뿐 다른 정보들과 무관하다. 이와 같은 특징으로 인하여 병렬처리가 가능하므로 연산속도가 매우 빠르다.

둘째, 신경망은 무수히 많은 연결강도를 가지고 있다. 따라서 정보의 분산표현 및 처리가 가능하다. 또, 일부의 정보로부터 전체를 얻을 수 있는 연산기억 특성을 갖는다. 또, 중복성이 크므로 Fault - Tolerant 할 뿐만 아니라 일부의 정보로부터 전체를 얻을 수 있는 연산기억 특성을 갖는다.

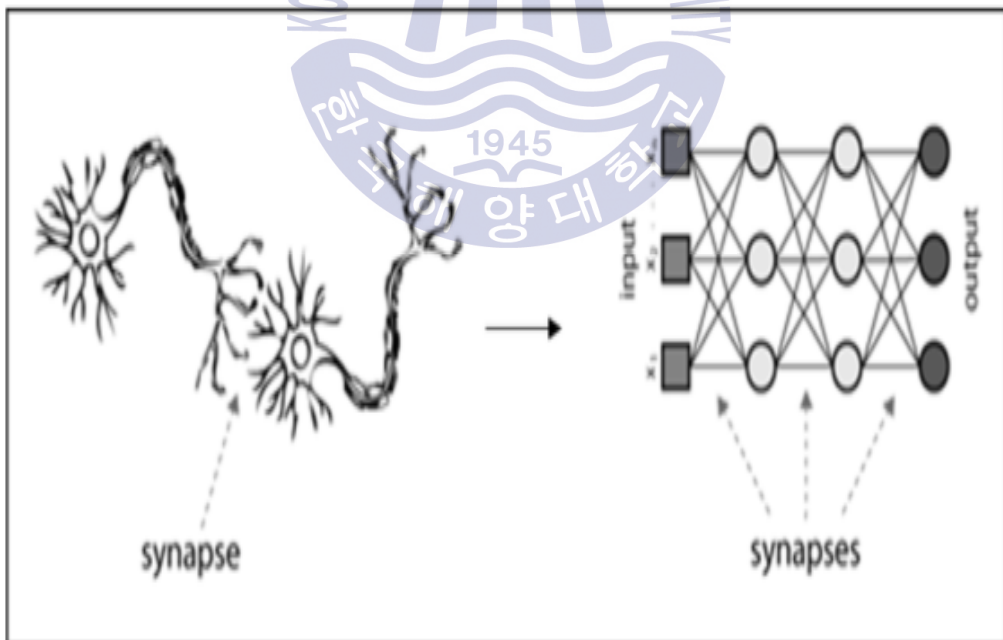
셋째, 지속적인 학습을 통해 연결강도를 조정함으로써 새로운 정보를 추가하거나 변경할 수 있는 적응특성을 가지고 있다.

생물학적 신경조직이 단순한 신경세포들의 대단위 병렬연결로 이루어져 있듯이, 인공신경망 또한 이러한 매커니즘을 구현하기 위해 구성되었다.(석태용, 2007) 인공신경망은 생물학적 신경망을 모방하여 인공신경망을 모델링한 내용 살펴보면 처리 단위(Processing unit) 측면에서는 생물적인 뉴런(neurons)이 노드(nodes)로, 연결성(Connections)은 시냅스(Synapse)가 가중치(weights)로 모델링 되어있다.

Table 11 생물학적 신경망과 인공신경망 구성요소 비교

생물학적 신경망	인공신경망
세포체	노드(Node)
수상돌기(Dendrite)	입력(Input)
축삭(Axon)	출력(Output)
시냅스	가중치(Weight)

자료 : 문병선, 2016. 각광받는 딥러닝의 근원 인공신경망(Artificial Neural Network)를 참조하여 제작성

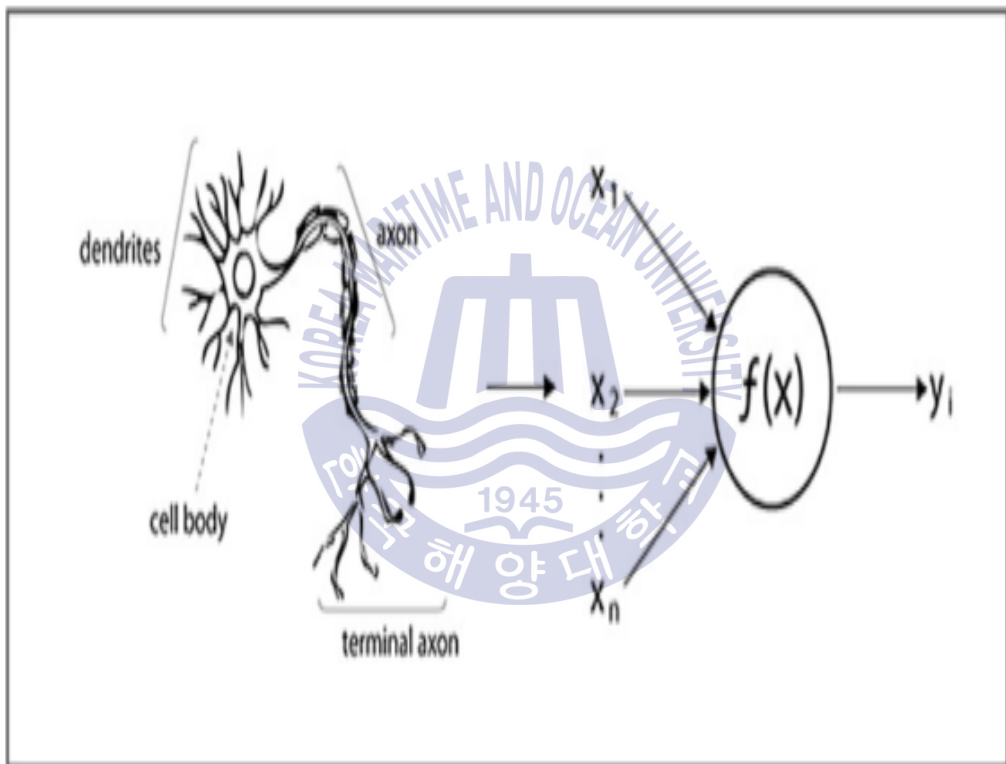


자료 : 문병선, 2016. 각광받는 딥러닝의 근원 인공신경망(Artificial Neural Network)

Fig. 18 Processing Unit 과 Neuron 비교 1

신경세포(뉴런)의 입력은 다수이고 출력은 하나이며, 여러 신경세포로부터 전달되어 온 신호들은 합산되어 출력된다. 합산된 값이 설정값 이상이면 출력 신호가 생기고 이 하이면 출력 신호가 없다. 이를 모형화 한 것은 Fig. 19와 같다.

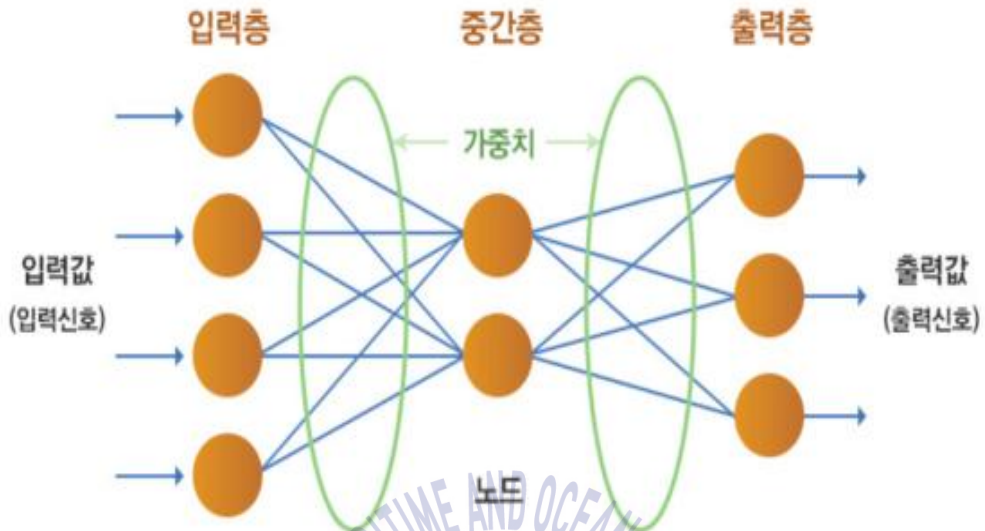
인간의 생물학적 신경세포가 하나가 아닌 다수가 연결되어 의미 있는 작업을 하듯, 인공신경망의 경우도 개별 뉴런들을 서로 시냅스를 통해 서로 연결시켜서 복수개의 계층(layer)이 서로 연결되어 각 층간의 연결 강도는 가중치로 수정(update) 가능하다.



자료 : 문병선, 2016. 각광받는 딥러닝의 근원 인공신경망(Artificial Neural Network)

Fig. 19 Processing Unit 과 Neuron 비교 2

인공신경망의 계층구조(Layer Structure)는 생물학적인 뇌(인간의 뇌)를 기반으로 모델링 하였고 인간 뇌의 적응성을 활용하여 적응 학습 능력과 병렬구조를 구현하고자 하였으며 인공신경망의 모델링을 구조를 매우 단순화하면 Fig. 20과 같다.



자료 : 문병선, 2016. 각광받는 딥러닝의 근원 인공신경망(Artificial Neural Network)

Fig. 20 인공신경망 계층구조 단순화 모형

각 노드들은 가중치가 있는 링크들로 연결되어 있고, 전체 모델은 가중치를 반복적으로 조정하면서 학습한다. 가중치는 장기 기억을 위한 기본 수단으로서 각 노드들의 중요도를 표현하며 단순화 시켜서 이야기하자면 인공신경망은 이들 가중치를 초기하고 훈련시킬 데이터 셋으로 가중치를 갱신하여 조정하여 전체 모델을 훈련시키는 것이다.

훈련이 완료된 후에 새로운 입력값이 들어오면 적절한 출력값을 추론해 내게 된다. 인공신경망의 학습원리는 경험의 일반화로부터 지능이 형성되는 과정이라고 보면 되고 bottom-up 방식으로 이루어지게 된다. Fig. 20에서 중간층이 2개 이상(즉 5~10개)일 경우를 층이 깊어진다고 보고 심층신경망(Deep Neural Netowkr)이라 하며, 이러한 심층 신경망을 통해서 이루어진 학습과 추론 모델을 딥러닝 모델이라 칭한다.

$$Y = f(X) = \sigma(\sum_i w_i g_i(X))$$

f : 비선형 함수

X : 연속형 그리고/ 혹은 이산형 변수들의 벡터, *Input nodes*

Y : 연속형 그리고/ 혹은 이산형 변수들의 벡터, *Output nodes*

g : X 의 함수, $g(x)$ 는 *hidden nodes*

σ : *Activarion Function*, 활성화 함수

w_i : 가중치

위 수식에서 활성화 함수란 임계치를 적용해 의미 없는 데이터를 사정에 필터링하고 미분을 편하게 계산하기 위해서 사용하는 함수이다. 신경망에서 중요한 것은 바로 가중치를 조절하는 것이며 가중치 조절을 어떻게 수행하느냐에 따라 신경망 모델의 정확도가 달라지게 된다.

Feedforward은 입력층, 은닉층, 출력층까지 순서대로 계산하는 방법이고, Backpropagation은 레이블된 학습 데이터를 가지고 여러개의 은닉층을 가지는 Feedforward 신경망을 학습시킬 때 사용되는 대표적인 지도학습 알고리즘이다. 즉 Forward Phase + Backward Phase이며 입력층에서 은닉층, 출력층까지 순서대로 거쳐, 다시 출력층에서 은닉층, 입력층으로 돌아가면서 가중치를 수정하고 위 과정을 계속 반복하며 가장 좋은 결과를 도출하는 방식이다.

Forward phase란 입력층에서 출력층 순서대로 활성화하면서 각 뉴런의 가중치와 활성화 함수를 적용하는 것이며, Backward Phase란 Forward Phases에서 생성된 결과 신호를 Train 데이터의 실제 목표 값과 비교하고, 망의 출력 신호와 실제 값의 차이가 나면 경사하강법(Gradient Descent)를 이용하여 뉴런간의 가중치를 변경하는 것이다.

경사하강법이란 최대 오차를 줄이는 방향으로 가중치를 변경하기 위한 방법으로 수식에서 w_i 를 수정해 나가는 방법론이다. 이러한 과정을 거치면서 신경망의 결과값이 달라지게 된다.



자료 : <https://ko.wikipedia.org/>

Fig. 21 경사하강법(Gradient Descent Method)

Fig. 21에서 x 는 가중치, $f(x)$ 는 그에 따른 오차이다. 즉 가중치를 조절함으로써 오차가 최소가 되도록 값을 조절해야하는 데 이때 사용하는 방법이 경사하강법이다. 경사하강법을 수행하면서 학습률(Learning rate)이라는 개념이 사용되는데 학습률이 너무 높게 설정하면 최적점(Convergence)에 도달하기 전에 높은 학습률로 인해 최적점을 무시하고 다시 높아지는 오차를 발생시킬수 있다.

은닉 계층(Hidden Layer)에 대한 활성화 함수는 네트워크에 비선형성을 적용하는데 필요하다. 활성화 함수가 적용되고, 결과는 네트워크 내의 다음 뉴런으로 전달된다. 대부분의 비선형 함수들이 사용되는데, 주로 시그모이드(Sigmoid) 함수가 주로 사용된다. 시그모이드 함수는 다음과 같은 수식으로 정의할 수 있다.

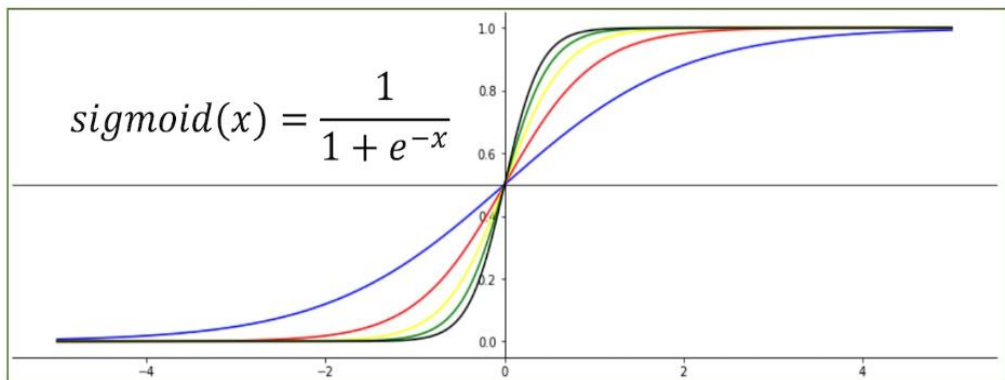
$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

e 는 자연상수이며 대략 2.71828이며 시그모이드 함수를 활성화 함수로 사용하는 뉴런을 시그모이드 뉴런이라 칭한다. 먼저 변수 z 에 가중치 합을 입력하면 다음 단계로 시그모이드 함수를 통과 시킨다.

$$z = b + \sum_i w_i x_i$$

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

입력값 z 에 대한 함수로 $\sigma(z)$ 의 값을 그래프로 나타내면 Fig. 22와 같다.

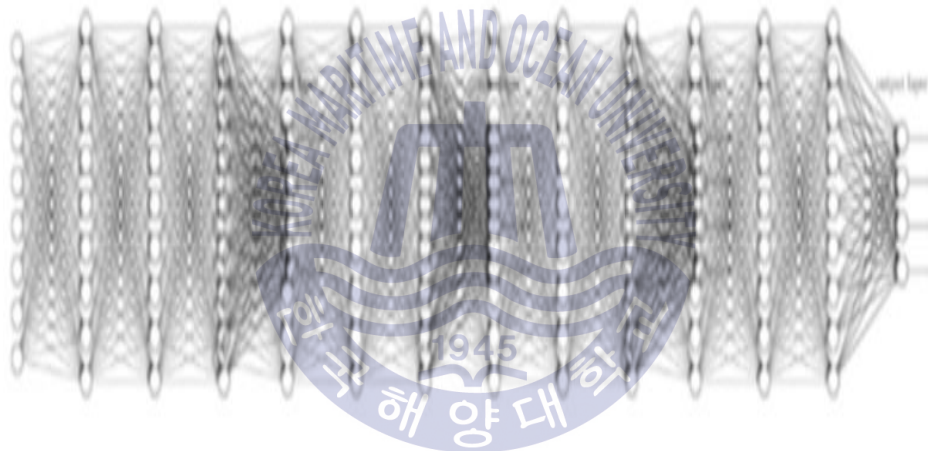


자료 : <https://ko.wikipedia.org/>

Fig. 22 경사하강법(Gradient Descent Method)

무한대 출력을 0에서 1사이로 응축시키므로 $\sigma(z)$ 를 압축 함수의 하나로 볼 수 있다. $z = 0$ 인 가운데 $\sigma(0) = 1/(1+e^0) = 1/2$ 이다. z 가 큰 음수일 경우 분모 e^{-z} 가 커져서 $\sigma(z)$ 가 0에 수렴한다. 반대로, z 가 아주 큰 양수이면 e^{-z} 가 0에 가까워지므로, $\sigma(z)$ 가 1에 수렴한다.

이처럼 시그모이드 뉴런은 수십 년간 대부분의 신경망의 기본 활성화 함수로 활용되었지만 최근 활용빈도가 감소하는 추세이다. 시그모이드 함수는 그래디언트 소실 (Vanishing gradient) 문제를 가지고 있어서 많은 층으로 이루어진 신경망을 훈련시키기 어렵기 때문이다.



자료 : <https://ko.wikipedia.org/>

Fig. 23 다층 역전파(Back propagation)

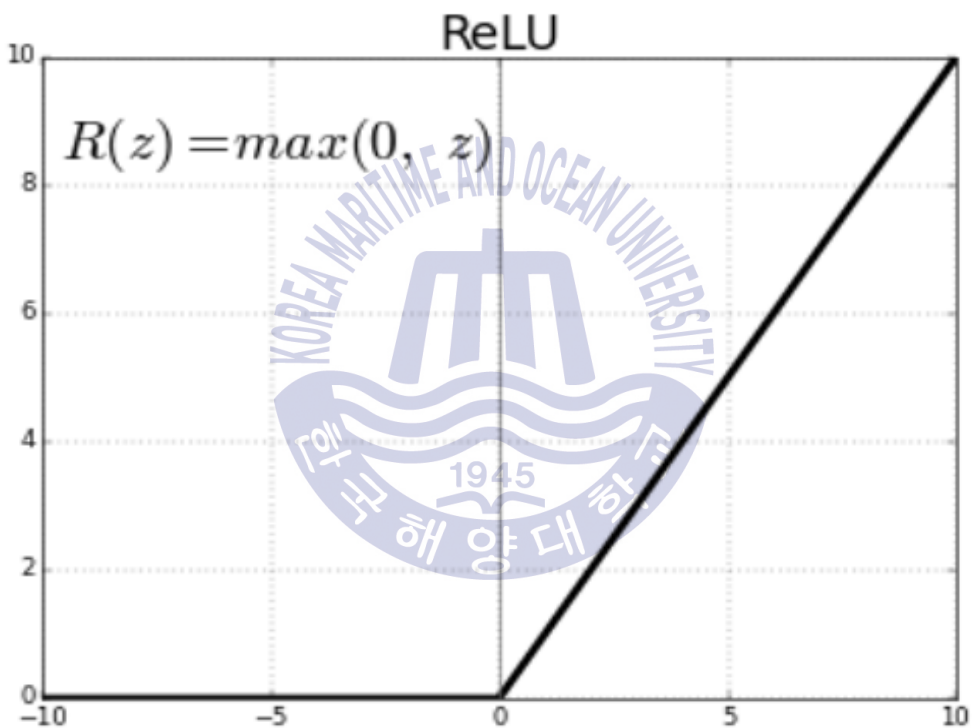
Fig. 23는 역전파(Back propagation)을 나타낸 그림이다. 이처럼 Layer가 많을 경우 미분 결과를 최초 Layer까지 전달하는 것은 불가능하다. 보통 2~3개의 Layer의 경우 정상적으로 계산되지만 그 이상 Layer를 가지는 경우 문제가 발생한다. 보통 역전파에서 결과를 전달할 때 활성화 함수인 시그모이드 함수를 활용한다. 반면 시그모이드 함수는 전달된 값을 0과 1사이로 심하게 변형한다.

이처럼 일정 비율로 줄어들기 때문에 왜곡은 아니지만 값이 현저하게 작아지는 현상을 나타낸다. 3개의 Layer를 거치면서 계속해서 1/10으로 줄어들면, 현재 사용 중인 값

은 처음 값을 1/1000이 된다. 이처럼 작아진 값을 원래 영향력을 그대로 복원한다는 것은 불가능하다.

따라서 최근 연구에서는 다른 종류의 활성화 함수인 ReLU를 사용하는 경우가 많다. ReLU함수는 Rectified Linear Unit의 약자로 기존 Linear 함수인 Sigmoid를 개선했다는 뜻이다. ReLU함수는 다음 수식과 같이 정의된다.

$$R(z) = \max(0, z)$$

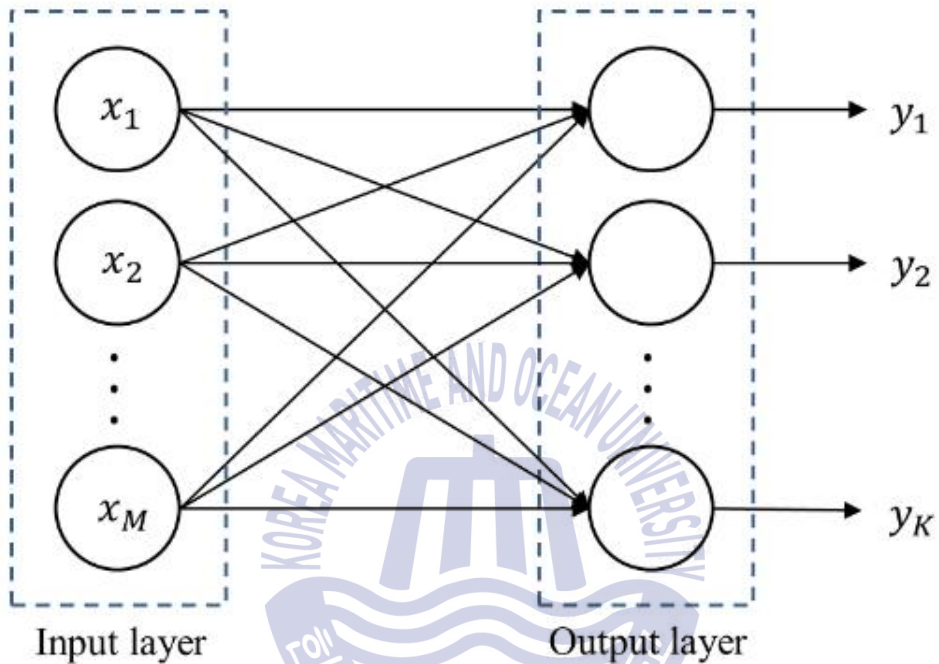


자료 : <https://ko.wikipedia.org/>

Fig. 24 ReLU 함수

ReLU함수는 모든 양수의 값은 통과시키지만 음수 값은 0으로 변환시킨다. 이처럼 ReLU 함수를 활용함으로 양 극단값이 포화되지 않고 최대값 연산이 1개로 매우 효율적인 계산이 가능하며 수렴속도가 시그모이드 함수에 비해 약 6배가 가량 빠르다는 장점을 가지고 있다.

Single-layer-ANN은 뉴런으로 이루어진 layer가 하나 뿐인 것을 말한다. Fig. 25는 Single-layer-ANN의 구조를 보여주며 각각의 원은 하나의 뉴런, 화살표는 뉴런이 연결되어 있음을 나타낸다. 이때 각각의 연결에는 weight가 부여된다.



자료 : <https://ko.wikipedia.org/>

Fig. 25 Single-layer-ANN

Fig. 25과 같이 2개의 layer가 존재하는데 이러한 ANN을 Single-layer라 하는 이유는 Input layer의 경우 단순히 주어진 x 를 그대로 전달하는 역할만 수행하기 때문이다. 따라서 Fig. 25에서 실제 뉴런으로 이루어진 layer는 Output layer 하나 뿐이다.

일반적으로 ANN에서는 하나의 뉴런은 binary class를 분류하도록 설계된다. 즉 뉴런의 Output y 는 $\{0,1\}$ 의 값을 갖거나 $[0,1]$ 사이의 값을 갖도록 설정된다. 따라서 K개의 class가 존재하는 classification 문제를 풀기 위해서는 ANN이 Output layer에 K개의 뉴런을 포함하도록 설계하여야 한다.

Single-layer ANN을 트레이닝하기 위한 Sum of squared error형태의 cost function C는 다음 수식과 같다.

$$C = \frac{1}{2} \sum_{n=1}^N \| t_n - y \|_2^2 = \frac{1}{2} \sum_{n=1}^N \sum_{i=1}^K (t_{\ni} - y_i)^2$$

위 수식에서 t_{n_i} 는 n 번째 데이터에 대한 label vector t_n 의 i 번째 요소이며, y_i 는 Output layer를 구성하는 i 번째 뉴런의 Output이다. 뉴런의 각 파라미터에 대한 update rule을 유도때 와 마찬가지로 Cost function C 를 i 번째 뉴런의 j 번째 weight w_{ij} 에 대해 편미분 하면 다음 수식과 같다. 이때 활성화 함수는 시그모이드 함수이다.

$$\frac{\partial C}{\partial w_{ij}} = \frac{\partial C}{\partial y_i} \frac{\partial y_i}{\partial s_i} \frac{\partial s_i}{\partial w_{ij}} = - \sum_{n=1}^N (t_{\ni} - y_i) y_i (1 - y_i) x_j$$

따라서, i 번째 뉴런의 j 번째 weight w_{ij} 에 대한 Update rule은 아래 수식과 같다.

$$w_{ij} = w_{ij} - \eta \frac{\partial C}{\partial w_{ij}} = w_{ij} + \sum_{n=1}^N (t_{\ni} - y_i) y_i (1 - y_i) x_j$$

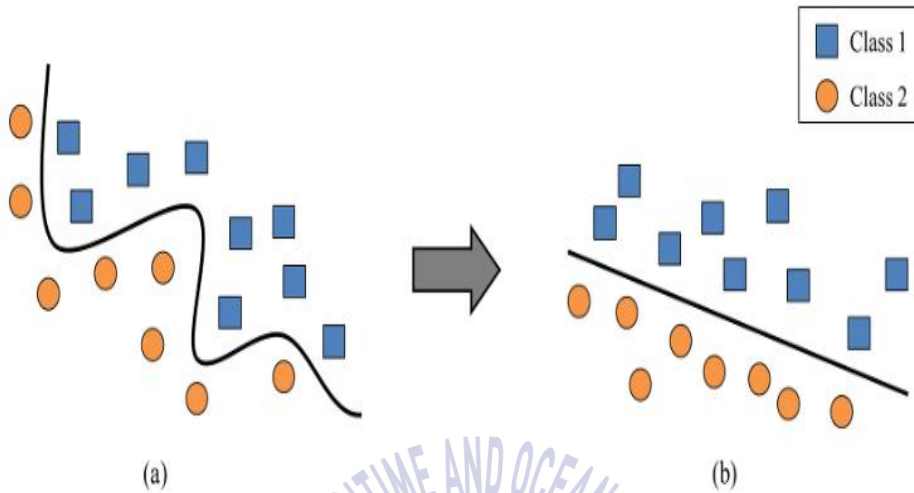
또한, i 번째 뉴런의 bias b_i 에 대한 Update rule은 아래 수식과 같다.

$$b_i = b_i - \eta \frac{\partial C}{\partial b_i} = b_i + \sum_{n=1}^N (t_{\ni} - y_i) y_i (1 - y_i)$$

Single-layer ANN은 실질적으로 서로 독립적인 뉴런을 여러개 나열한 것이므로 Single-layer ANN의 각각 뉴런에 대한 Update rule은 정의된 Update rule과 크게 다르지 않다.

Multi-layer-ANN은 다소 복잡한 function이 필요한 경우, 단순한 Single-layer ANN으로 모델을 구축하는 것이 불가능한 경우 많이 활용된다. Multi-layer ANN의 가장 중요한 특성은 주어진 데이터를 그대로 이용하여 Classifier를 형성하는 것이 아니라, 은닉층 구조를 추가함으로써 Classifier를 형성함과 동시에 주어진 데이터에 최적화된 변환 연산을 트레이닝한다는 점이다.

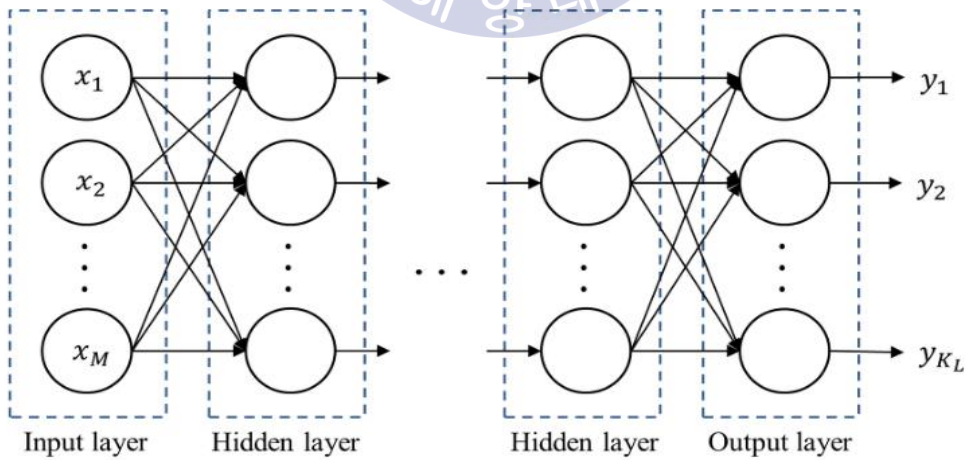
또한 은닉층의 뉴런 수에 따라서는 주어진 데이터의 차원을 축소하거나 증가시키는 기능도 수행할 수 있다. 이처럼 Multi-layer-ANN을 활용하면 복잡한 데이터 분포를 단순한 형태의 Linear function만으로도 데이터를 완벽하게 Classification 할 수 있다.



자료 : <https://ko.wikipedia.org/>

Fig. 26 데이터 변환과 Classification

Multi-layer-ANN은 뉴런으로 이루어진 layer를 2개 이상 포함하는 ANN을 말한다. Fig. 27과 같이 Multi-layer-ANN은 Output layer를 제외한 모든 layer를 hidden layer라고 한다.



자료 : <https://ko.wikipedia.org/>

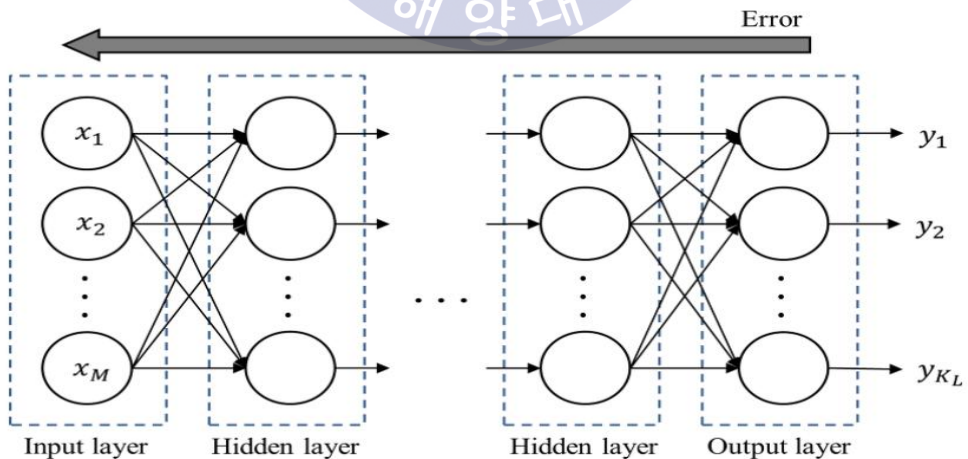
Fig. 27 Multi-layer-ANN

Multi-layer-ANN은 Output layer와 동일하게 Cost function을 정의하여 이를 최소화하는 방식으로 트레이닝 할 수 있다. 따라서 활성화 함수는 시그모이드 함수를, Cost function으로 Sum of squared error를 사용할 경우 Output layer를 구성하는 뉴런에 대한 Update rule은 아래 수식과 같다. 이때 L은 ANN을 구성하는 layer의 수이며, $w_{ij}^{(l)}$ 은 l번째 layer를 구성하는 i번째 뉴런의 j번째 weight, $b_i^{(l)}$, $y_i^{(l)}$ 은 각각 i번째 뉴런의 bias와 Output이다. 또한 l번째 layer에 입력되는 j번째 입력 $x_j^{(l)}$ 은 $y_j^{(l-1)}$ 과 같다.

$$w_{ij}^{(L)} = w_{ij}^{(L)} - \eta \frac{\partial c}{\partial w_{ij}^{(L)}} = w_{ij}^{(L)} + \sum_{n=1}^N (t_n - y_i^{(L)}) y_i^{(L)} (1 - y_i^{(L)}) y_j^{(L-1)}$$

$$b_i^{(L)} = b_i^{(L)} - \eta \frac{\partial c}{\partial b_i^{(L)}} = b_i^{(L)} + \sum_{n=1}^N (t_n - y_i^{(L)}) y_i^{(L)} (1 - y_i^{(L)})$$

그 동안 뉴런을 구성하는 파라미터에 대한 Update rule은 트레이닝 데이터세트에 포함된 label을 이용하여 정의한 Cost function을 최소화하는 weight와 bias를 계산함으로써 유도되었다. 그러나 은닉층의 경우 label이 존재하지 않기 때문에 Cost function을 정의할 수 없다는 문제점이 존재한다. 역전파 알고리즘(Back propagation algorithm)은 이러한 문제를 해결하여 Multi-layer-ANN을 트레이닝 하기 위해 고안되었다. 역전파의 기본적인 방식은 Fig. 28와 같이 Output layer의 Output과 Cost function으로부터 계산된 오차(t와 y의 차이)를 은닉층으로 역전파하는 것이다.



자료 : <https://ko.wikipedia.org/>

Fig. 28 오차의 역전파

3.2.4 이중계절형 Holt-Winters 모형

지수평활법은 매우 단순하지만 시계열 예측에서 매우 유용한 기법으로 사용이 되고 있다. 지수평활법의 표준 계절조정 방법으로 Holt-Winters의 가법 추세성분과 승법추세 성분, 감소가법추세성분, 감소승법추세성분 등을 포함하고 있다.

Holt-Winters 모형은 Holt의 이중 모수 지수평활법을 확장시킨 방법으로 계절형 시계열 예측에 사용되는 잘 알려진 방법으로서, 지수 평활법은 계산이 쉽고 필요정보의 양이 최소화 되는 장점이 있어 유용하다. 모형은 1개의 계절주기를 고려한 모형으로서 식 (1)에서 Y_t 는 고려할 t시점의 값으로 계절주기를 s라고 하면, L_t 는 t시점의 수준의 평활상수이다.

식 (2)의 T_t 는 t 시점의 추세의 평활 상수이며 식 (3)에서 S_t 는 t시점의 계절 평활상수를 의미한다. α, β 는 수준과 추세에 대한 모수이며 γ 는 계절 평활 모수이다.

식 (4)의 F_{t+h} 는 h 시점 후에 예측한 값을 의미한다.

$$L_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (1)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (2)$$

$$S_t = \gamma(Y_t - L_t) + (1-\gamma)S_{t-s} \quad (3)$$

$$F_{t+h} = L_t + T_t \times h + S_{t+h-s} \quad (4)$$

이중 계절성을 적용한 계절형 Holt-Winters 모형은 식 (5)와 같다. 식 (6), 식 (7)은 계절 주기 s_1, s_2 의 계절상수로 t시점에서 S_t 와 D_t 로 정의된다.

$$L_t = \alpha(Y_t - S_{t-s_1} - D_{t-s_2}) + (1-\alpha)(L_{t-1} + T_{t-1}) \quad (5)$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \quad (6)$$

$$S_t = \gamma(Y_t - L_t - D_{t-s_1}) + (1-\gamma)S_{t-s_1} \quad (7)$$

$$D_t = \delta(Y_t - L_t - S_{t-s_1}) + (1-\delta)D_{t-s_2} \quad (8)$$

$$F_{t+h} = L_t + T_t \times h + S_{t+h-s_1} + D_{t+h-s_2} \quad (9)$$

이중 계절형 Holt-Winters 모형의 초기값은 식 (10)을 통하여 제시된다.

$$L_{s_1} = \frac{1}{s_1} \sum_{t=1}^{s_1} Y_t, L_{s_2} = \frac{1}{s_2} \sum_{t=1}^{s_2} Y_t \quad (10)$$

$$T_{s_1} = \frac{1}{s_1^2} \left(\sum_{t=s_1+1}^{2s_1} Y_t - \sum_{t=1}^{s_1} Y_t \right), T_{s_2} = \frac{1}{s_2^2} \left(\sum_{t=s_2+1}^{2s_2} Y_t - \sum_{t=1}^{s_2} Y_t \right)$$

$$S_1 = Y_1 - L_{s_1}, \dots, S_{s_1} = Y_{s_1} - L_{s_1}$$

$$D_1 = Y_1 - L_{s_2}, \dots, D_{s_2} = Y_{s_2} - L_{s_2}$$

평활 모수 $\alpha, \beta, \gamma, \delta$ 의 추정은 훈련자료를 통하여 추정가능하다. 과거 모형추정단계 중 1차 시점 예측에서 오차제곱합(Sum of Squared error ; SSE)을 최소화하는 과정을 통하여 계산되는데, 이에 대하여 $\alpha, \beta, \gamma, \delta$ 의 각 모수들의 값을 0에서 1까지 매우 작은 수치로 변동시켜 h시점 이후의 예측값을 F_{t+h} 와 실제값 Y_{t+h} 의 제곱합 SSE가 최소가 되는 조합으로 $\alpha, \beta, \gamma, \delta$ 의 값을 결정한다.

예를 들어 $\alpha, \beta, \gamma, \delta$ 의 초기값을 0.001로 지정한 상태에서 모수 추정을 시작할 경우, 각 모수들의 추정치는 0.001 ~ 1까지의 반복된 입력을 통하여 1000^4 의 조합된 모수 추정치 중 최소의 SSE가 되는 조합을 선택한다. 이에 대하여 SSE는 식 (11)에 의하여 계산된다.

$$\begin{aligned} SSE &= \sum_{t=1}^{n-h} (Y_{t+h} - F_{t+h})^2 \\ &= \sum_{t=1}^{n-h} (Y_{t+h} - (L_t + T_t \times h + S_{t+h-s_2}))^2 \end{aligned} \quad (11)$$

Taylor(2003)는 수정된 이중계절형 Holt-winters 모형을 제안하였다. Taylor(2003)은 오차항간의 자기상관오차를 고려한 모형으로서 기존의 이중계절형 Holt-Winters 모형의 식 (12)로부터 F_{t+h} 를 식 (13)과 같이 변형하여 정의 한다.

$$L_{s_1} = \frac{1}{s_1} \sum_{t=1}^{s_1} Y_t, L_{s_2} = \frac{1}{s_2} \sum_{t=1}^{s_2} Y_t \quad (12)$$

$$T_{s_1} = \frac{1}{s_1^2} \left(\sum_{t=s_1+1}^{2s_1} Y - \sum_{t=1}^{s_1} Y_t \right), T_{s_2} = \frac{1}{s_2^2} \left(\sum_{t=s_2+1}^{2s_2} Y_t - \sum_{t=1}^{s_2} Y_t \right)$$

$$S_1 = Y_1 - L_{s_1}, \dots, S_{s_1} = Y_{s_1} - L_{s_1}$$

$$D_1 = Y_1 - L_{s_2}, \dots, D_{s_2} = Y_{s_2} - L_{s_2}$$

$$F_{t+h} = L_t + T_t \times h + S_{t+h-s_1} + D_{t+h-s_2} + \emptyset^h [Y_t - L_{t-1} - T_{t-1} - S_{t-s_1} - D_{t-s_2}] \quad (13)$$

\emptyset 는 1차 상관계수에 조정을 나타내는 값으로 평활 모수 $\alpha, \beta, \gamma, \delta, \emptyset$ 는 1차 예측단계에서 위의 이중계절형 Holt-winters 모형과 마찬가지로 오차제곱합(Sum of Squared error ; SSE)을 최소화하는 과정을 통하여 계산된다.



제4장 인공지능경망을 활용한 SCFI 예측 모형 구축

4.1 예측 요인 선정

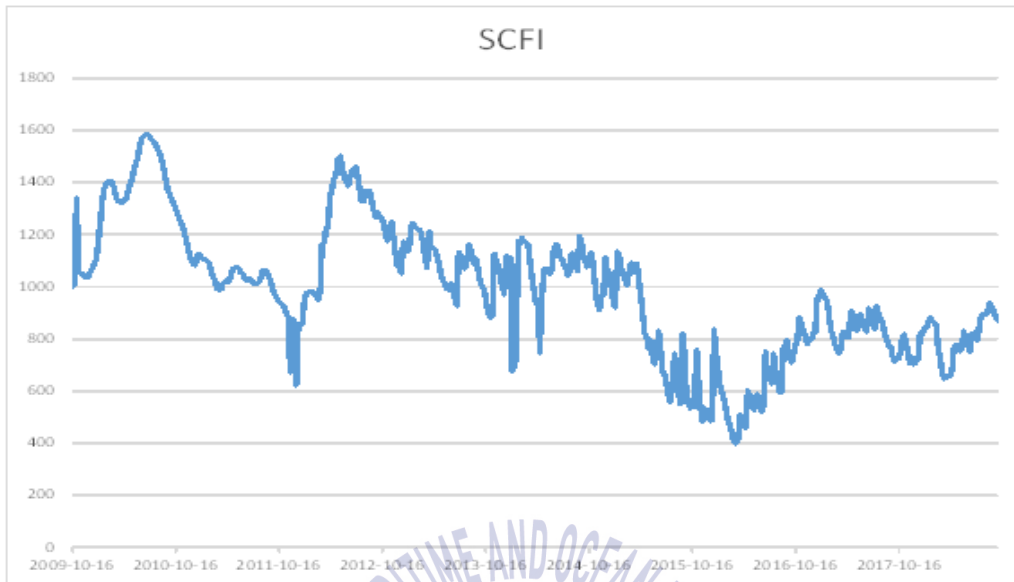
4.1.1 SCFI(Shanghai containerized Freight Index)

정기선 운임 지수의 경우 대표적으로 SCFI(Shanghai containerized Freight Index)와 CCFI(China containerized Freight Index)가 있다. SCFI와 CCFI의 차이점을 살펴보면 SCFI는 상해발 주요 항로가 포함되며 대표적인 항로로는 유럽, 지중해, 미국 서안, 미국 동안, 페르시아만, 호주·뉴질랜드, 서아프리카, 남아프리카, 남아메리카, 서일본, 동일본, 동남아시아, 한국, 대만 및 홍콩 등 15개 항로이다.

반면 CCFI의 경우 운임지수를 산출하는 주요 항로는 11개 항로인데, SCFI와 달리 일본항로가 서일본과 동일본 항로로 나뉘지 않고 통합되어 있으며, 페르시아만, 대만 및 홍콩 항로가 없다. SCFI와 마찬가지로 중국발 수출 운임이나 주요 항만이 전 중국에 걸쳐 있다는 특징이 있다. SCFI와 또 다른 것은 SCFI가 선물거래의 기초자산으로 활용되는 데 반해, CCFI는 그렇지 않다는 점이다.(김태일, 정봉민, 2012) 또한 SCFI는 각종 선행연구 상 분석에 활용되고 있으며 이외 실무차원에서 SCFI가 커버하는 구간에 운송 서비스를 제공하는 선사, SCFI를 이용한 각종 파생거래상품 등장 등 상대적으로 CCFI에 비해 활용도 및 분석의 적합성이 높다고 할 수 있다. 이처럼 SCFI의 경우 CCFI에 비해 세계적인 주요항로를 포함하고 있으며 선물거래의 기초자산으로 활용되기 때문에 본 연구에서 수행할 정기선 운임지수 예측에 적합한 운임지수라 판단하였다.

Table 12 SCFI 자료 기초통계량분석

기간	자료 수	최소값	최대값	표준편차	평균
2009. 10.16 ~ 2018. 9.31	450개	400.43	1583.18	254.95	979.80



자료 : 상해해운교역소(<http://www.sse.net.cn/home>)를 참고하여 재구성

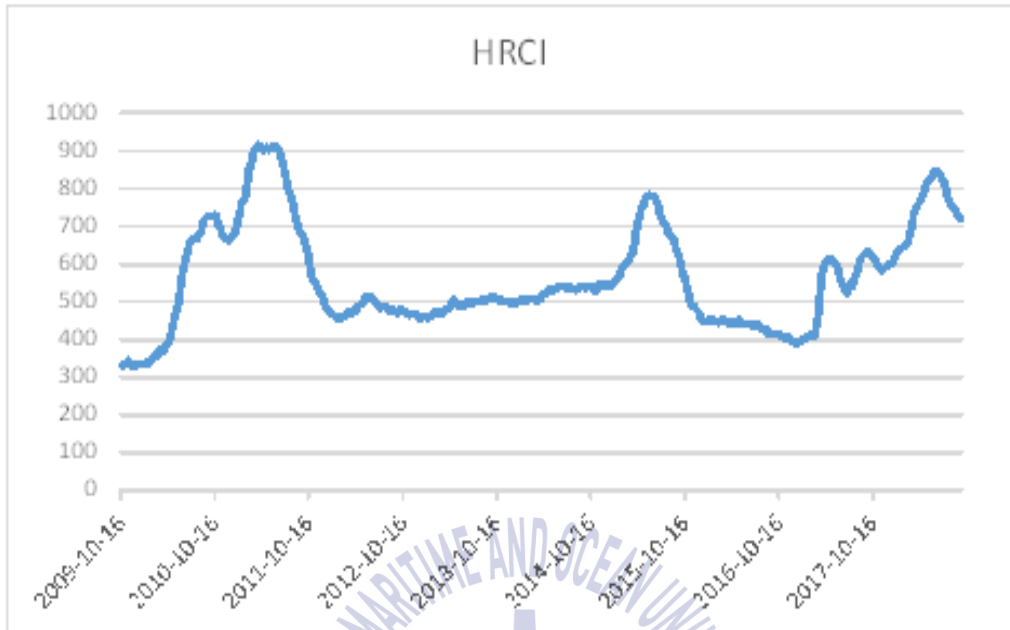
Fig. 29 SCFI 추이(2009.10 ~ 2018.9)

4.1.2 HRCI(Howe Robinson Containership index)

HRCI는 영국의 해운컨설팅 및 브로커 업체인 ‘Howe Robinson C. I.’ 사가 최초 1997년 1월 15일 지수를 1,000 기준으로 매주간 단위로 발표하는 대표적인 컨테이너선 용선료 지수이다. 지수는 여러 가지 컨테이너 선형의 기간 용선료를 선형에 따라 가중 평균하여 지수화한 것으로 정기선 해상운임과 선박의 기간 용선료는 같은 방향으로 움직인다. 해상운임이 상승 국면이면 용선 수요가 증가하여 용선료도 상승하게 된다.(박동주, 2017)

Table 13 HRCI 자료 기초통계량분석

기간	자료 수	최소값	최대값	표준편차	평균
2009. 10.16 ~ 2018. 9.31	450	329.4	916.1	139.19	563.24



자료 : 상해해운교역소(<http://www.sse.net.cn/home>)를 참고하여 재구성

Fig. 30 HRCI 추이(2009.10 ~ 2018.9)

4.1.3 Bunker 380CST Index

선박 운항에 영향을 미치는 중요한 요인 중 하나는 연료 가격이다. 실질적으로 Bunker 가격 하락은 선사에게 긍정적인 영향과 부정적인 영향을 동시에 미친다. 통상 매출의 30%를 Bunker 비용으로 소비하는 정기선사의 경우 Bunker 가격 하락은 비용 절감을 실현하였지만 반대로 운임하락과 용선료 하락을 동시에 유발하였다. 이에 따라 2015년에 발생한 SCFI 지수의 급락의 원인 중 하나로 작용하였다.(박동주, 2017)

반대로 Bunker 비용 상승은 해상운임의 증가와 화주의 물류비 상승으로 전가되고 이는 다시 제품가격의 상승으로 이어지며, 영향의 크기는 미미하지만 제품가격의 상승으로 이어질 수 있다.

따라서 본 연구에서 중요한 운임지수 결정요인이라 판단하여 주 선박연료유 종류인 380CST 기준으로 6개 지역 16개 항만의 평균 Bunker 가격인 Bunker 380CST Index를 활용하였다. 기존 일부 대표적인 Bunkering 항만의 평균가격을 사용하는 것 보다 주요 Bunkering 항만의 특정 유종의 평균가격을 활용하는 것이 보다 항로별 다양성을 확보할 수 있는 방안이라 판단하였다.

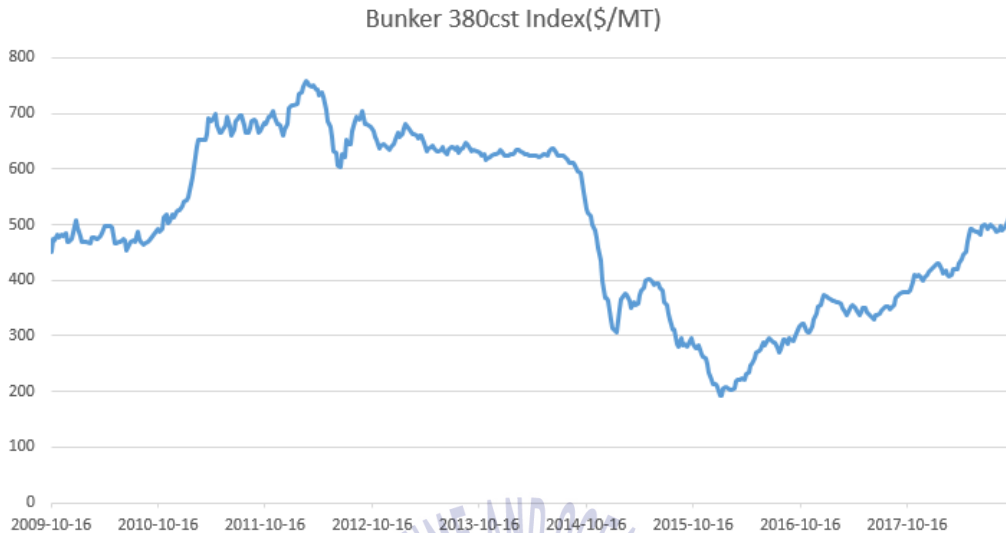
Table 14 Bunker 380CST Index 해당 주요 항만

지역	항만
Asia&Oceania	Busan
	Kaohsiung
	Singapore
Caribbean & C. America	Maracaibo
Europe	Antwerp
	Istanbul
	Piraeus
	Rotterdam
Middle East & Africa	Fujairah
	Istanbul
	Las Palmas
North America	Houston
	New Orleans
South America	Maracaibo
	Rio de Janeiro
	Santos

자료 : <http://www.bunkerindex.com>을 참고하여 제작성

Table 15 Bunker 380CST Index 자료 기초통계량분석

기간	자료 수	최소값	최대값	표준편차	평균
2009. 10.16 ~ 2018. 9.31	450	193.14	757.63	151.86	497.97



자료 : <http://www.bunkerindex.com>을 참고하여 제작성

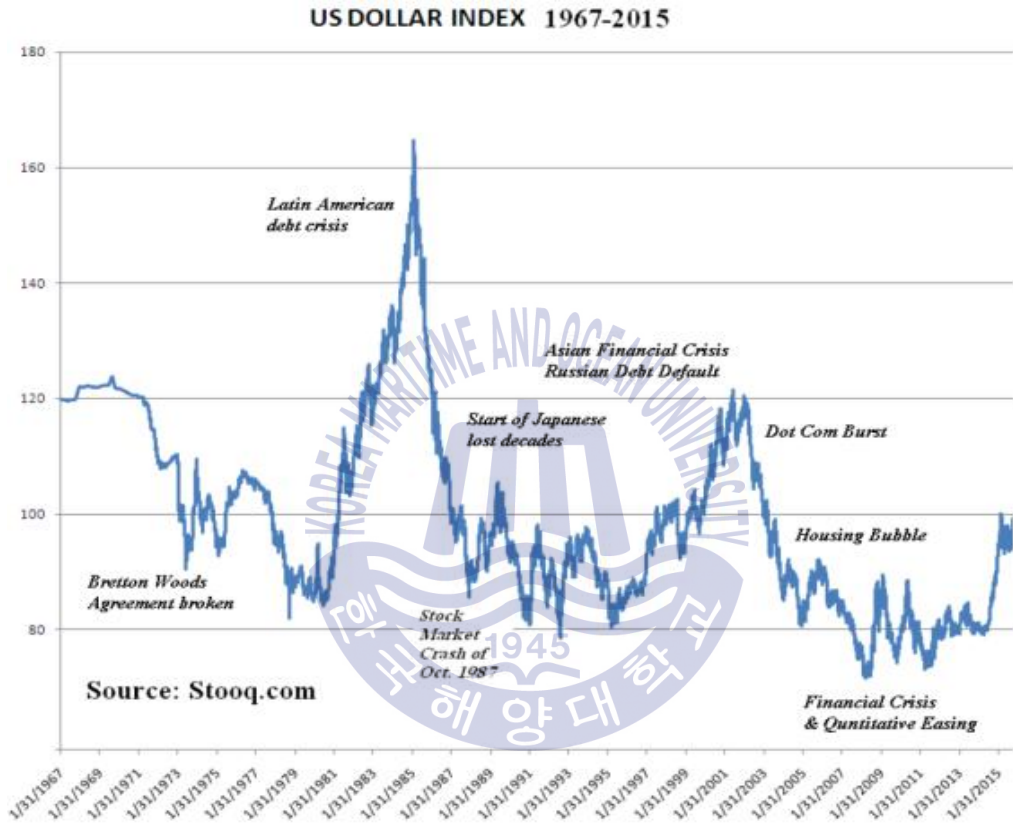
Fig. 31 Bunker 380CST Index 추이

4.1.4 U.S. Dollar Index

통화 가치가 비교적 안정적인 주요 6개국의 통화 대비 미국 달러의 가치를 지수화한 것으로 주요국의 화폐 가치 대비 달러 가치의 움직임을 파악하기 위하여 만들어진 지수이다. 통상적으로 줄여서 ‘USDX(U.S. Dollar Index)’ 라 한다. 6개국 통화에는 유로, 일본 엔, 영국 파운드스털링, 캐나다 달러, 스웨덴 크로나, 스위스 프랑이 포함된다. 1973년 3월을 기준점(100)으로 하여 미국 연방준비위원회 FRB(Federal Reserve Board of Governors)에서 발표한다. 만약 이 지수의 수치가 110이면 달러의 가치가 주요 6개국 화폐 가치보다 1973년 3월 대비 약 10% 상승했음을 의미한다.

6개국의 통화 가치는 가중평균을 이용하여 계산하며 이때 가중치는 각 국가의 경제수준 혹은 미국과의 교역비중 등을 바탕으로 정해진다. 2017년 기준 가중치는 유로 57.6%, 일본 엔 13.6%, 파운드스털링 11.9%, 캐나다 달러 9.1%, 스웨덴 크로나 4.2%, 스위스 프랑 3.6%이다. 1973년부터 2017년까지 6개국 통화 대비 달러 가치가 가장 높았을 때는 U.S. Dollar Index: 164.7으로 1985년 1월 라틴아메리카에서 부채 위기가 발생했을 때이며, 가장 낮았을 때는 U.S. Dollar Index 70.69 2008년 리먼사태로 글로벌금융위기가 발생했을 때이다.

또한 U.S. Dollar Index의 경우 Asian Financial Crisis, Russian Debt Default 및 2008 Financial Crisis & Quantitative Easing 등 세계 경제 상황을 반영하고 있다. 컨테이너 해운시장과 달러가치는 밀접한 연관을 가지고 있으며 기존 해당 국가의 달러환율을 활용하는 방안보다 U.S. Dollar Index를 활용하는 것이 더욱 적합할 것이라 판단하였다.



자료 : <https://en.wikipedia.org>

Fig. 32 U.S. Dollar Index 추이

Table 16 U.S. Dollar Index 자료 기초통계량분석

기간	자료 수	최소값	최대값	표준편차	평균
2009. 10.16 ~ 2018. 9.31	450	73.11	103	8.22	86.93

4.2 예측 요인 상관 분석

상관 분석(Correlation analysis)은 확률론과 통계학에서 두 변수간에 어떤 선형적 관계를 가지고 있는지 분석하는 방법이다. 두 변수는 서로 독립적인 관계이거나 상관된 관계일 수 있으며 이때 두 변수간의 관계의 강도를 상관관계(Correlation, Correlation coefficient)라하며 상관관계의 정도를 파악하는 상관계수(Correlation coefficient)라 한다.

단순히 두 개의 변수가 어느 정도 강한 관계에 있는가를 측정하는 단순상관분석(simple correlation analysis), 3개 이상의 변수들간의 관계에 대한 강도를 측정하는 다중상관분석이 있다. 다중상관분석에서 다른 변수들과의 관계를 고정하고 두 변수만의 관계에 대한 강도를 나타내는 것을 편상관계분석(partial correlation analysis)이라고 한다. 이때 상관관계가 $0 < \rho \leq +1$ 이면 양의 상관, $-1 \leq \rho < 0$ 이면 음의 상관, $\rho = 0$ 이면 무상관이라고 한다. 하지만 0인 경우 상관이 없다는 것이 아니라 선형의 상관관계가 아니라는 것이다.

본 연구에서는 보편적으로 사용되는 피어슨 상관 계수(Pearson correlation coefficient 또는 Pearson's r)를 활용하였다. r의 경우 X와 Y가 함께 변하는 정도/X와 Y가 각각 변하는 정도 계산할 수 있으며 결과의 해석은 r 값은 X 와 Y 가 완전히 동일하면 +1, 전혀 다르면 0, 반대방향으로 완전히 동일하면 -1 을 가진다. 결정계수 (coefficient of determination) 는 r^2 로 계산하며 이것은 X 로부터 Y를 예측할 수 있는 정도를 의미한다. 일반적으로 r값의 구간에 따른 해석은 Table 17과 같다.

Table 17 r값 구간에 따른 상관관계

r값 구간	상관관계 해석
-1.0과 -0.7 사이	강한 음적 선형관계
-0.7과 -0.3 사이	뚜렷한 음적 선형관계
-0.3과 -0.1 사이	약한 음적 선형관계
-0.1과 +0.1 사이	거의 무시될 수 있는 선형관계
+0.1과 +0.3 사이	약한 양적 선형관계
+0.3과 +0.7 사이	뚜렷한 양적 선형관계
+0.7과 +1.0 사이	강한 양적 선형관계

본 연구에서 예측 요인으로 활용하는 변수인 SCFI, HRCI, US Dollar Index, Bunker 380 CST Index 간 상관관계는 Table 18과 같다. 종속변수인 SCFI와 독립변수인 HRCI 간에는 거의 무시될 수 있는 선형관계를 나타내고, 종속변수인 SCFI와 독립변수인 HRCI과 US Dollar Index간에는 뚜렷한 음적 선형관계를 나타낸다.

Table 18 예측 요인 상관관계 분석 결과

	SCFI	HRCI	US Dollar Index	Bunker 380CST Index
SCFI	1.0000	-0.0627	-0.6256	0.6024
HRCI	-0.0627	1.0000	-0.0679	0.1243
US Dollar Index	-0.6256	-0.0679	1.0000	-0.8533
Bunker 380CST Index	0.6024	0.1243	-0.8533	1.0000

변수간 관계의 직선성을 알아보는 산점도의 경우 Fig.33과 같다.

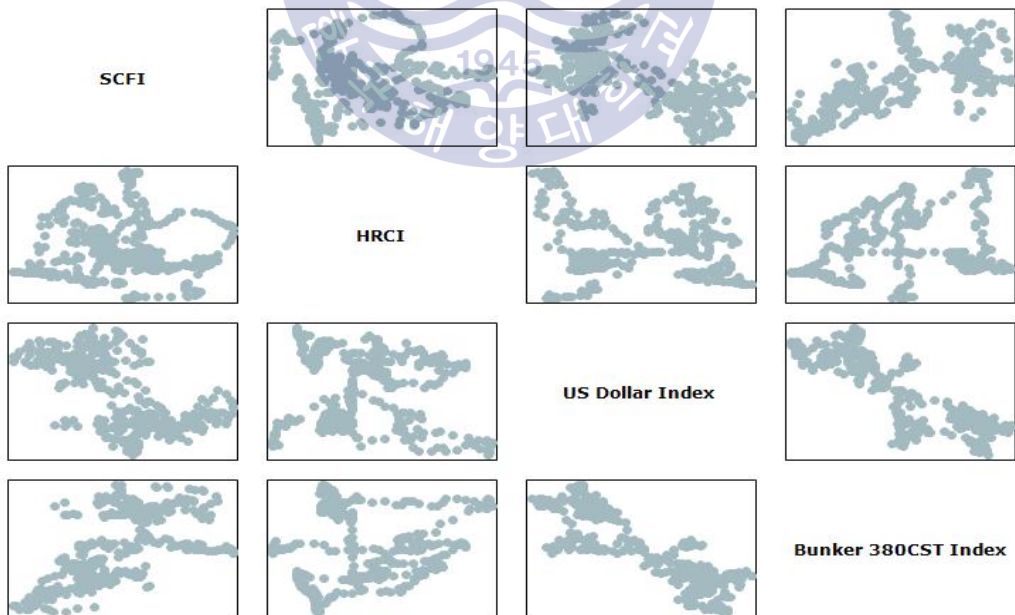


Fig. 33 예측요인 산점도

4.3 최적 인공신경망 예측 모형 구축

4.3.1 최적 인공신경망 예측 모형 구축 과정

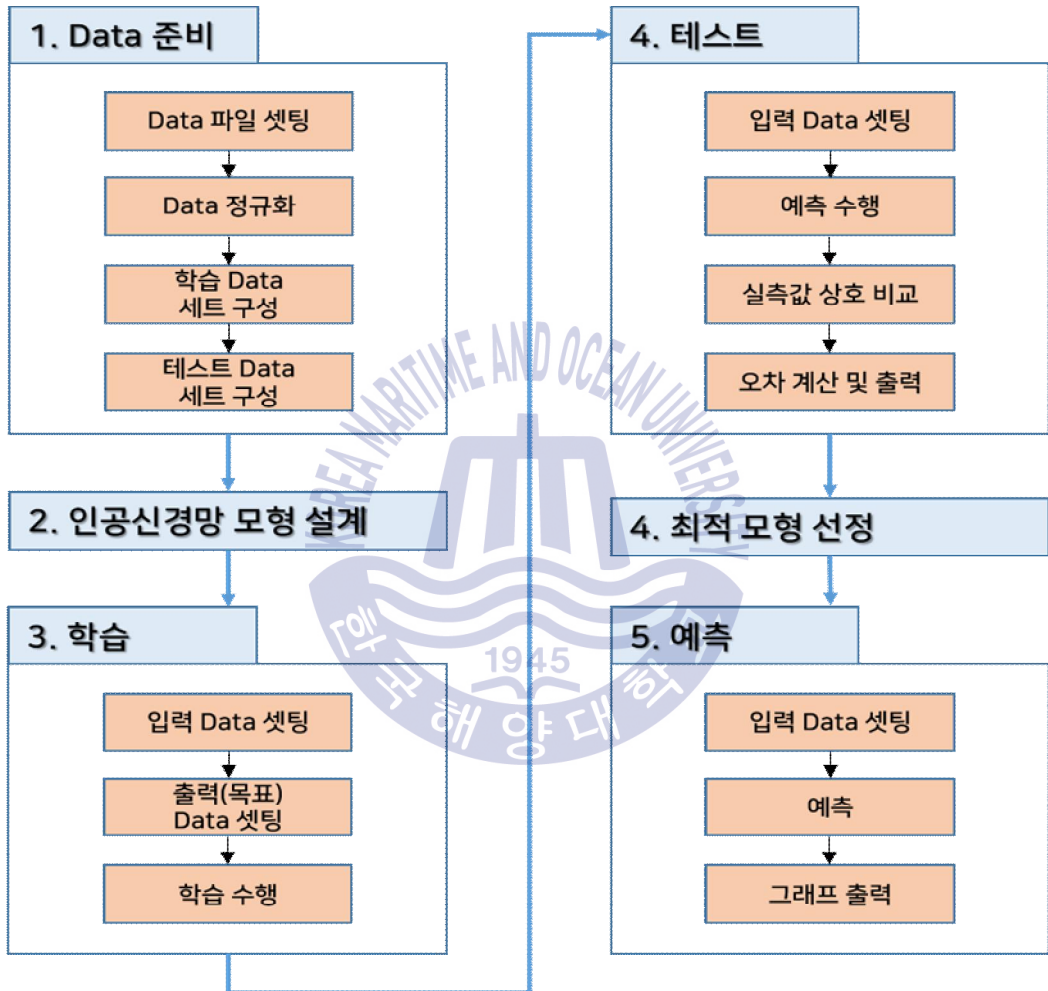


Fig. 34 최적 인공신경망 예측 모형 구축 과정 요약

인공신경망 모형의 성능을 결정하는 은닉층과 은닉뉴런의 수, 은닉층과 출력층의 활성화함수, 인공신경망 알고리즘 등을 조절하여 적절하게 설정해야한다. 하지만 선행연구에 따르면 이러한 설정에 대한 최적화 방법은 정해진 것이 없으며, 실험적으로 오차를 줄여나가는 시행착오법을 사용한다.

은닉층과 은닉노드 구성의 경우 1 ~ 3개의 은닉층과 4 ~ 12개의 은닉노드를 설정하면서 가장 우수한 평가기준을 충족하는 구성을 파악하였다.

평가 기준은 평균 제곱근 오차(Root Mean Square Error ; RMSE), 평균 제곱 오차(Mean Square Error ; MSE), 절대 오차평균(Mean Absolute Error ; MAE) 및 평균 절대 백분비오차(Mean absolute percentage error ; MAPE)를 활용하였다.

활성화함수로는 선형함수(Linear function), Sigmoid, tanh 함수 등이 존재하며 본 연구에서는 Sigmoid 함수를 활용하였다. 출력층의 활성화함수의 경우 선형함수를 사용하였는데 이는 출력변수의 값이 상수로써 의미를 가지는 회귀 모형을 훈련하기 때문이다. 오차함수의 경우 잔차제곱합(Sum of Squared Errors ; SSE)를 활용하였으며 인공신경망의 알고리즘은 가중치를 가지는 탄력적 역전파(Resilient backpropagation)을 활용하였다.

최적의 인공신경망 모델을 도출하기 위해 위에 언급한 Hyper-Parameters 뿐만 아니라 학습 계수인 Learning Rate, 중단 기준으로 에러 함수의 편미분(partial derivatives)을 위한 임계값인 Threshold 등을 조절하였다.

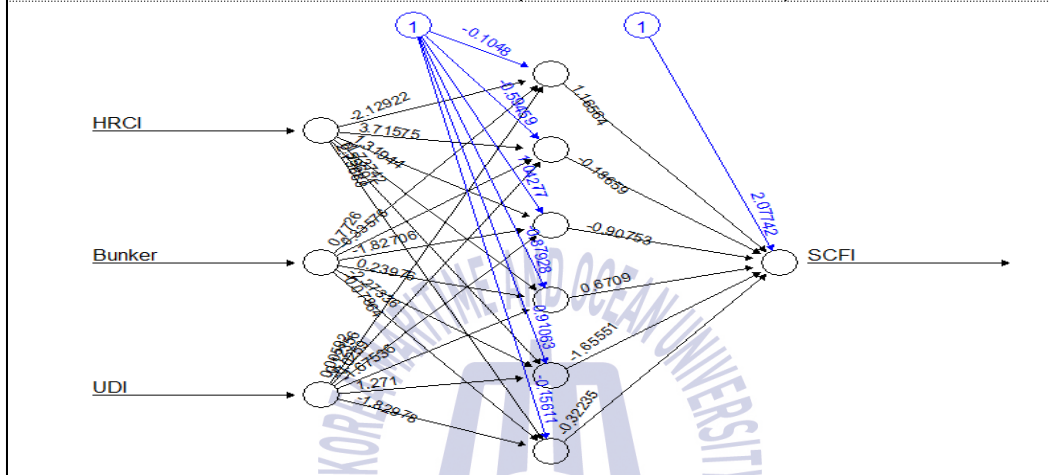
Table 19 인공신경망 예측 모형 학습결과

인공신경망 예측 모형	모형 평가	
Hidden Layer : 1 Hidden Node : 4	RMSE	0.136
	MSE	0.018
	MAE	0.116
	MAPE	0.272

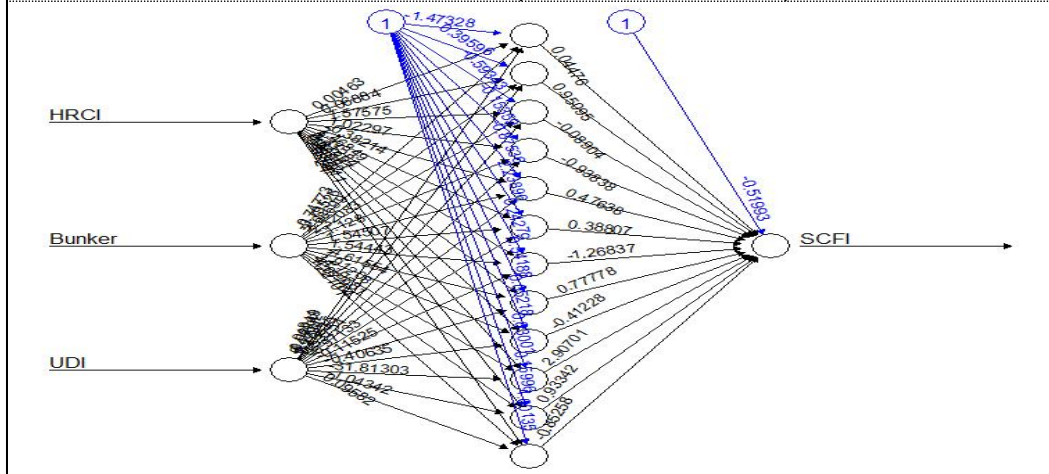
The diagram illustrates a neural network with three input nodes (HRCI, Bunker, UDI) and one output node (SCFI). There is one hidden layer with four nodes. Each input node is connected to each hidden node, and each hidden node is connected to the output node. The weights for these connections are as follows:

- HRCI to Hidden Node 1: 22.68188
- HRCI to Hidden Node 2: 13.41405
- HRCI to Hidden Node 3: 14.49229
- HRCI to Hidden Node 4: 15.44232
- Bunker to Hidden Node 1: 22.10292
- Bunker to Hidden Node 2: 33.14347
- Bunker to Hidden Node 3: 22.26687
- Bunker to Hidden Node 4: 15.40028
- UDI to Hidden Node 1: 8.22836
- UDI to Hidden Node 2: 4.16929
- UDI to Hidden Node 3: 46.57405
- UDI to Hidden Node 4: 1.90989
- Hidden Node 1 to SCFI: 1.22452
- Hidden Node 2 to SCFI: 3.28505
- Hidden Node 3 to SCFI: 1.95118
- Hidden Node 4 to SCFI: 2.2027
- Bias Node 1 (value 1) to Hidden Node 1: -8.83395
- Bias Node 1 (value 1) to Hidden Node 2: 0.74992
- Bias Node 1 (value 1) to Hidden Node 3: 8.78187
- Bias Node 1 (value 1) to Hidden Node 4: -2.2027
- Bias Node 2 (value 1) to SCFI: -0.95955

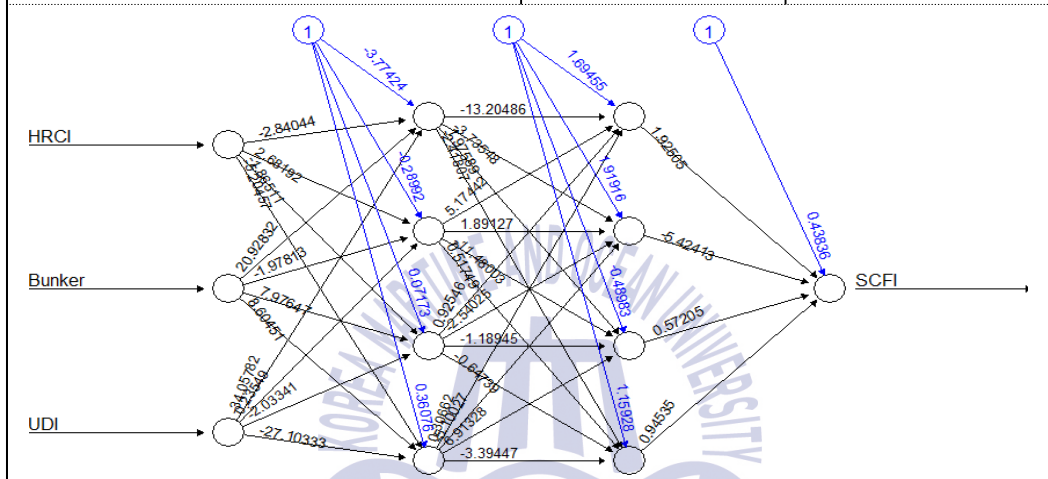
Hidden Layer : 1 Hidden Node : 6	RMSE	0.1861
	MSE	0.0346
	MAE	0.1665
	MAPE	0.4263



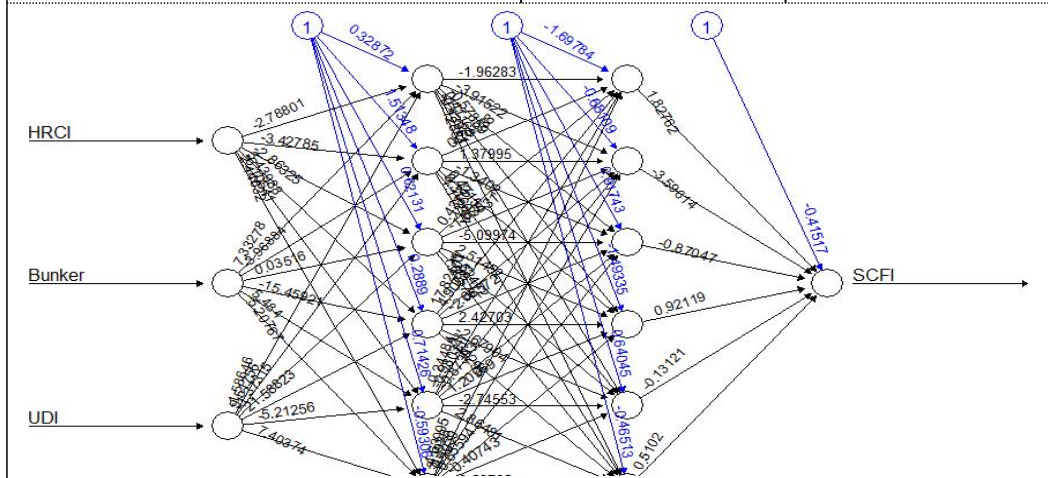
Hidden Layer : 1 Hidden Node : 12	RMSE	0.2011
	MSE	0.0404
	MAE	0.1892
	MAPE	0.4455



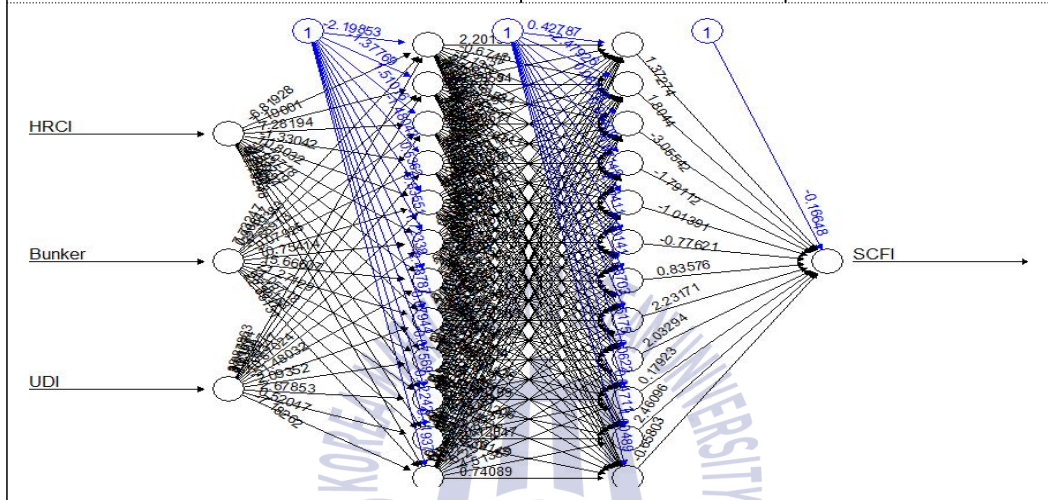
Hidden Layer : 2 Hidden Node : 4 × 4	RMSE	0.2089
	MSE	0.0436
	MAE	0.1924
	MAPE	0.4378



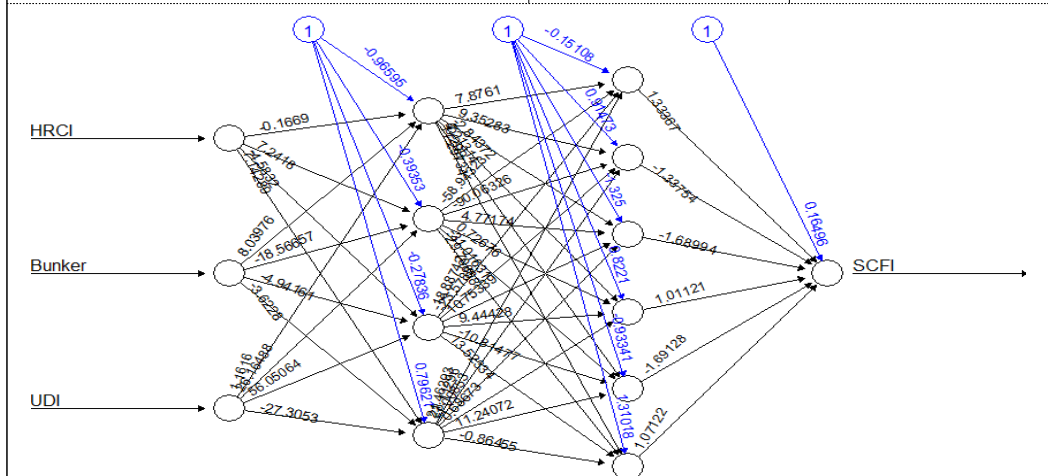
Hidden Layer : 2 Hidden Node : 6 × 6	RMSE	0.2169
	MSE	0.0470
	MAE	0.1989
	MAPE	0.4491



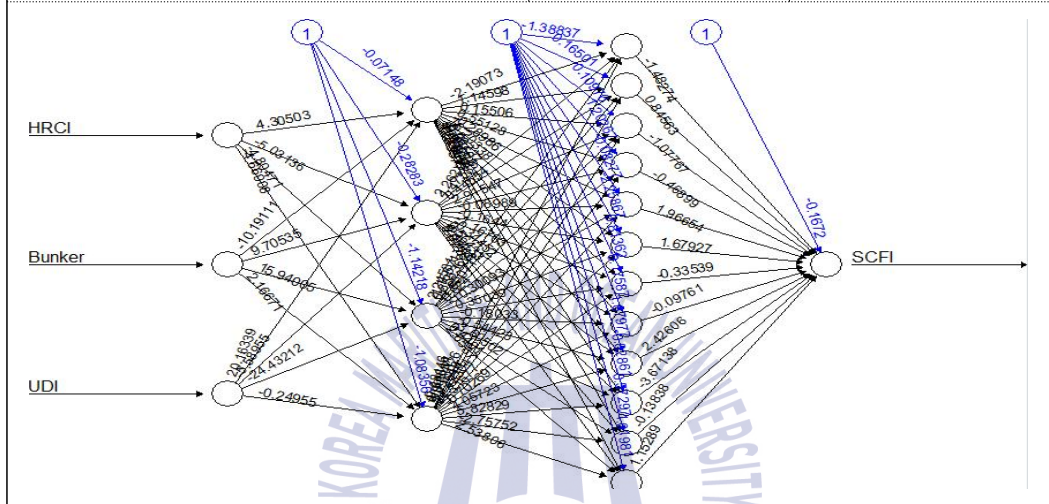
Hidden Layer : 2 Hidden Node : 12 × 12	RMSE	0.2169
	MSE	0.0470
	MAE	0.1989
	MAPE	0.4491



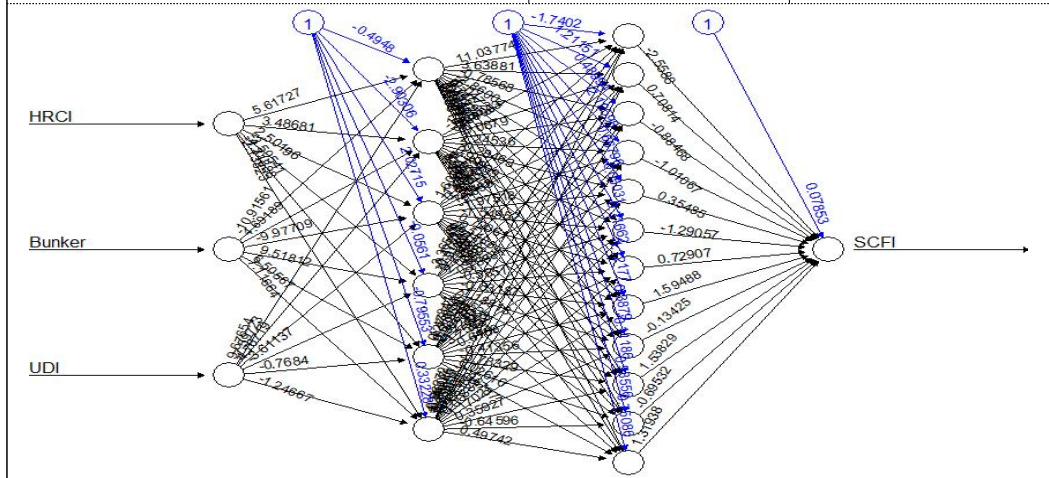
Hidden Layer : 2 Hidden Node : 4 × 6	RMSE	0.1854
	MSE	0.0343
	MAE	0.1705
	MAPE	0.3968



Hidden Layer : 2 Hidden Node : 4 × 12	RMSE	0.2046
	MSE	0.0418
	MAE	0.1867
	MAPE	0.4254



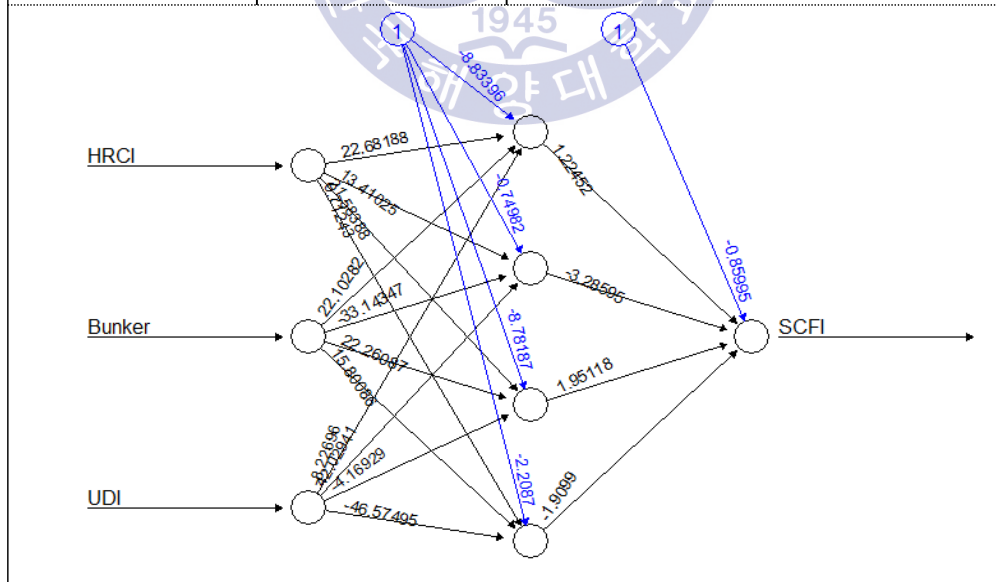
Hidden Layer : 2 Hidden Node : 6 × 12	RMSE	0.2280
	MSE	0.0520
	MAE	0.2096
	MAPE	0.4727



최적 인공신경망 모형 설계를 위해 다양한 Hyper-Parameters를 조절하여 최적의 모델을 도출하였고 구성은 아래 Table 20와 같다.

Table 20 최적 인공신경망 모형 구성

Hyper-parameters		최적 인공신경망 구성
Normalization Method		Min-Max [0, 1]
Train set/Test set		7:3
Hidden Layers		1
Hidden Node		4
Activation Function		Sigmoid
Error Function		SSE
Learning Rate		0.01
Learning Algorithm		Resilient back propagation
Threshold		0.01
Evaluation	RMSE	0.136
	MSE	0.018
	MAE	0.116
	MAPE	0.272



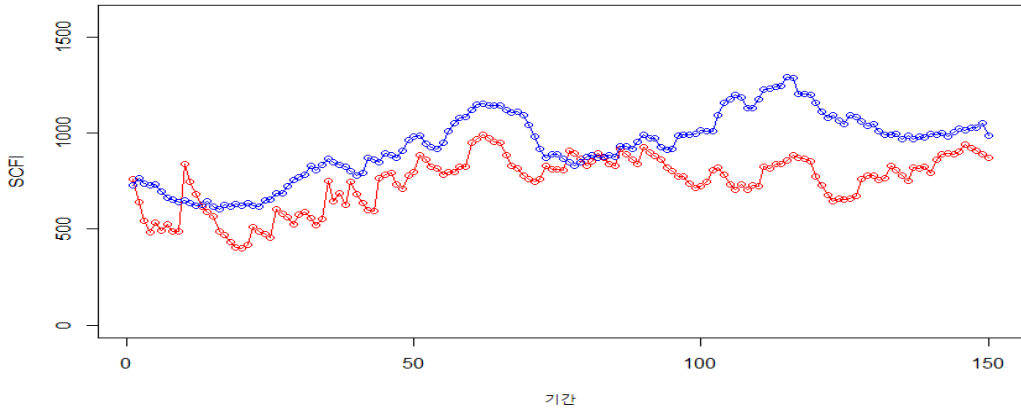


Fig. 35 최적 인공신경망 모델과 SCFI 실측값 비교 그래프

4.3.2 최적 인공신경망 모델을 통한 예측 수행

본 연구에서는 최종 선정된 인공신경망 모델을 기반으로 2018년 10월 첫째주 기준 향후 15주의 SCFI를 예측하였다. 시계열 데이터 예측 모델을 출력 노드 수에 따라 분류하면 1단계 예측(One-step ahead forecasting)과 다단계 예측(Multi-step ahead forecasting) 모형이 있다. 본 연구에서는 다단계 예측 모형을 통해 2018년 10월 첫째주 기준 향후 15주의 SCFI를 예측하였다. 예를 들어 다단계 예측 모형은 출력노드가 m 개이면, 과거 $t-n+1$ 부터 현재 t 시점까지 n 개의 실제 데이터 세트($\hat{x}_{t+1}, \hat{x}_{t+2}, \dots, \hat{x}_{t+m}$)로 향후 $t+1$ 부터 $t+m$ 시점의 데이터 세트($x_{t-n+1}, x_{t-n+2}, \dots, x_{t-1}, x_t$)를 한번에 예측한다.

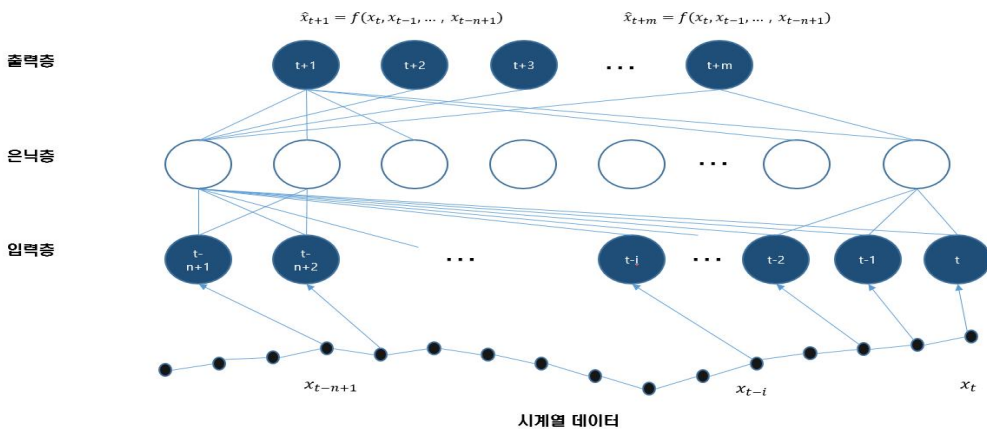
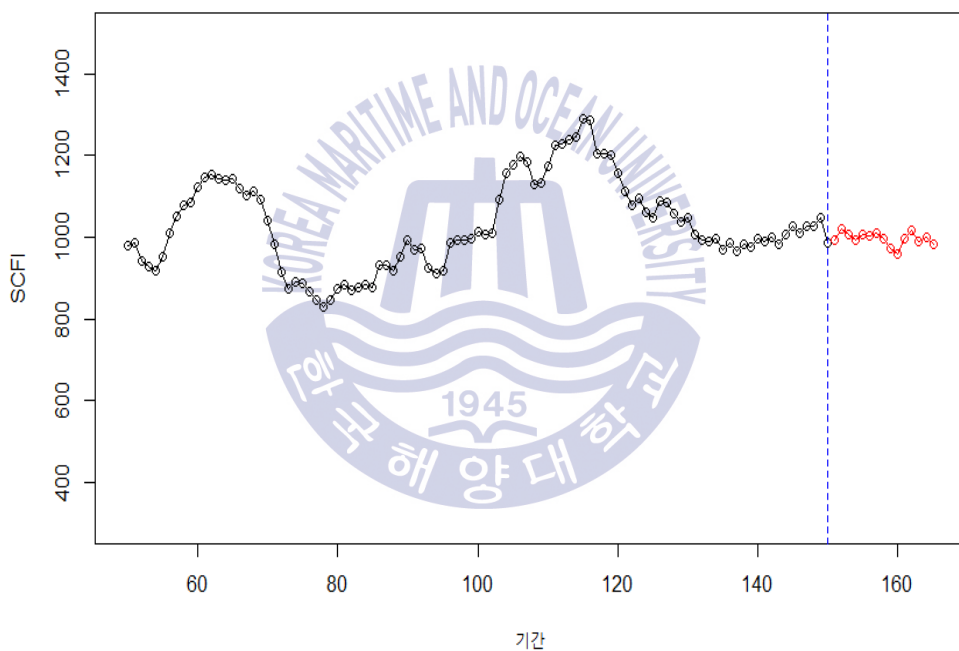


Fig. 36 다단계 예측 모형(Multi-step ahead forecasting)

총 150주의 과거 시계열 자료를 기반으로 다단계 예측 모형을 활용하여 2018년 10월부터 향후 15주간의 SCFI를 예측하였다. 본 연구에서 2018년 10월부터 2019년 1월(15주)간의 예측을 수행한 이유는 3/4분기 성수기 이후 운임상승세 유지 및 춘절 전 물동량 증가로 인한 운임상승 현상이 반영되는지 파악하기 위함이다.

반면 정통적으로 컨테이너 해운시장의 운임상승 시기임에도 불구하고 예측결과 2018년 10월부터 2019년 1월간 최대 1,010포인트, 최소 992포인트로 일시적인 소폭 상승세를 나타내었지만 전반적으로 하락하는 추세를 나타내었다.



주 : 파란 실선 이후 빨간색 그래프가 향후 15주 SCFI 예측값을 나타냄

Fig. 37 다단계 예측 모형을 통한 SCFI 예측값(2018년 10월 ~ 2019년 1월)

예측 결과를 바탕으로 2018년 10월 첫째주부터 2018년 12월 둘째주 까지 실제 SCFI 값과 비교해본 결과 상대적인 수치의 차이는 존재하나 점차 하락하는 동일한 추세를 나타내었다.

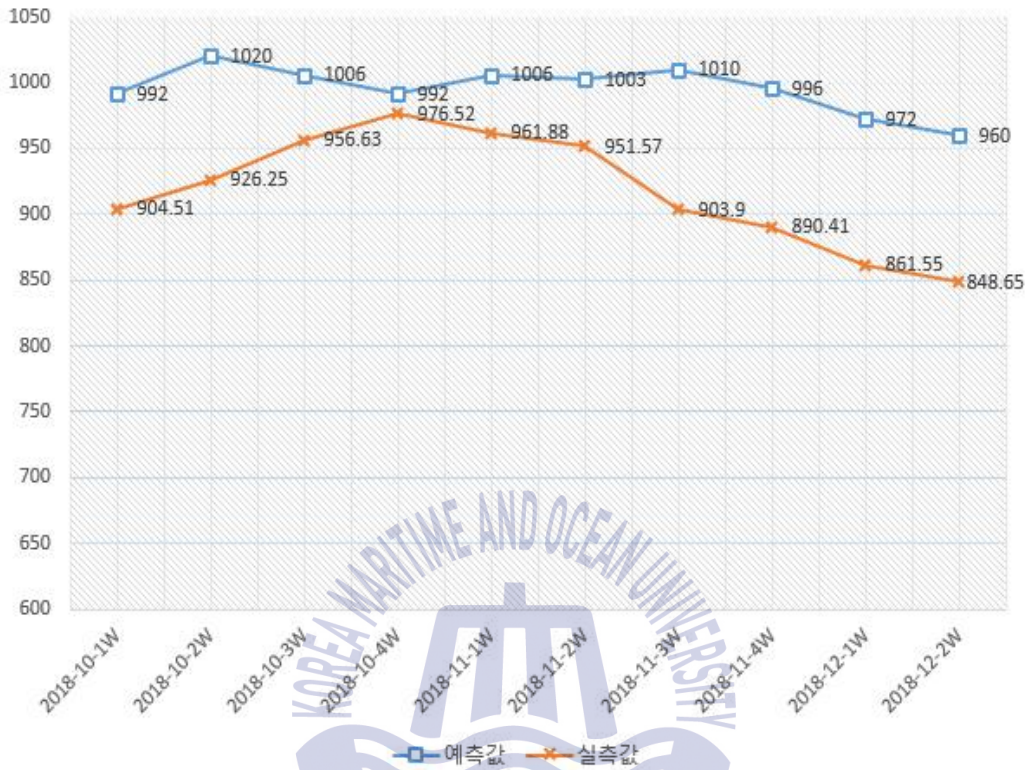


Fig. 38 SCFI 실측값 및 예측값 비교 그래프(2018년 10월 첫째주 ~ 2018년 12월 둘째주)

제5장 결론

5.1 연구결과 및 시사점

컨테이너 해운시장은 급격한 환경변화와 다양한 경제적·환경적 요인들로 인해 불확실성이 점차 증대되고 있다. 이러한 정기선 해운의 시황을 대표하는 SCFI는 2016년 3월 18일 400.3으로 최저를 기록하였으며 이후 등락을 반복하며 컨테이너 해운시장의 불확실성과 급격한 변동성을 반영하고 있다.

과거에 그러하였듯 시장의 급격한 변화를 파악하지 못하고 이에 대한 적절한 대응방안을 마련하지 못한다면 2008년 리먼 브라더스 사태와 EU 경제위기처럼 악순환을 반복하게 될 것이다. 따라서 이러한 악순환을 반복하지 않기 위해서는 컨테이너 해운시장의 향후 시장 변화의 예측과 이에 대한 적절한 대응방안 마련이 필요하다.

본 연구에서는 정기선 해운의 시황을 반영하는 운임지수인 SCFI를 인공신경망(Artificial Neural Network : ANN) 모델을 활용하여 향후 운임지수 변화를 예측하였다. SCFI에 영향을 미치는 독립변수로는 대표적인 컨테이너 용선료 지수인 HRCI(Howe Robinson Containership Index), 주된 선박연료유 종류인 380CST의 평균가격인 Bunker 380CST Index 및 통화가치가 비교적 안정적인 주요 6개국의 통화 대비 미국 달러의 가치를 지수화한 U.S. Dollar Index를 활용하였다. 연구에 활용된 자료는 SCFI 발표가 시작된 2009년 10월부터 2018년 9월까지 전체 450주의 시계열 자료를 활용하였다. 종속변수로 선정한 HRCI, Bunker 380CST Index 및 U.S. Dollar Index 또한 SCFI와 동일한 기간의 자료를 확보하여 분석에 활용하였다.

최적 인공신경망 모델 설계를 위해 은닉층 구성, 알고리즘, 활성화함수, Learning Rate 등 다양한 Hyper-Parameters를 조정하여 실험적으로 오차를 줄여나가는 시행착오법을 사용하였다. 인공신경망 모델의 평가 기준은 평균 제곱근 오차(Root Mean Square Error ; RMSE), 평균 제곱 오차(Mean Square Error ; MSE), 절대 오차평균(Mean Absolute Error ; MAE) 및 평균 절대백분비오차(Mean absolute percentage error ; MAPE)를 활용하였다. 최적 모델로 선정된 인공신경망의 경우 1개의 은닉층과 4개의 은

닉노드를 가지며 활성화 함수는 Sigmoid함수, 알고리즘은 Resilient back propagation을 적용하였다. 인공신경망 예측모델의 성능을 평가하기 위한 기준은 RMSE = 0.136, MSE = 0.018, MAE = 0.116, MAPE = 0.272로 평가 되었다.

Hyper-Parameters 조정을 통한 시행착오법으로 도출된 최적 인공신경망 모델을 통해 SCFI 예측을 수행하였으며 전통적인 컨테이너 해운시장의 특징인 3/4분기 성수기 이후 운임상승세 유지 및 춘절 전 물동량 증가로 인한 운임상승 현상 여부를 파악하기 위해 15주간(2018년 10월 ~ 2019년 1월) 기간설정을 통해 SCFI를 예측하였다. 예측결과 최대 1,010 포인트 최소 992포인트로 일시적인 소폭상승세를 나타내지만 전반적으로 지속적인 하락세를 나타내는 것으로 예측되었다. 또한 SCFI 예측치와 실측치의 추세비교를 위해 2018년 10월 첫째주부터 2018년 12월 둘째주까지 실제 SCFI 값과 비교해본 결과 상대적인 수치의 차이는 존재하나 점차 하락하는 동일한 추세를 나타내었다.



5.2 연구의 한계점 및 향후 연구방향

컨테이너 해운시장 운임의 경우 부컨테이너 해운시장의 완전경쟁가격결정이론이 적용되지 않는 반독점적 시장의 성격을 가지고 있으며 상대적으로 부컨테이너 해운시장에 비해 수요·공급의 변동이 운임에 미치는 영향이 적다고 할 수 있다.

따라서 이러한 컨테이너 해운시장의 특징을 가만하여 본 연구에서는 수요·공급 측면이 아닌 상대적으로 거시적인 경제요인을 변수로 활용하여 컨테이너 운임지수를 예측하였다. 하지만 향후 보다 개선된 연구를 위해 노후선박 해체량, 신조선 발주량 및 주요 정기선 화물 운송량 등 수요·공급 요인을 고려하여 컨테이너 해운시장의 예측 연구를 수행한다면 보다 다양한 요인을 고려한 연구가 이루어 질 것으로 판단된다.

또한 본 연구에서 활용한 인공신경망(Artificial Neural Network : ANN) 모델 뿐만 아니라 전통적인 시계열 분석법인 ARIMA(Autoregressive integrated moving average) 모델, 계절형 시계열 분석 방법으로 활용될 수 있는 Holt-Winters 모형 등을 활용하여 컨테이너 운임 지수 예측력에 관한 연구가 수행된다면 보다 다양한 연구결과의 비교가 가능할 것이라 생각된다.



참고문헌

<국내문헌>

- 강순일, 2018. 인공지능경망(ANN)모델을 활용한 건화물선 운임 예측. 석사학위논문. 한국해양대학교
- 김래현, 2014. 인공지능경망을 이용한 KOSPI 200 단기예측에 관한 연구. 석사학위논문. 국민대학교
- 김석환, 2016. 컨테이너 해상운임과 운임결정 요인의 상관관계 연구. 석사학위논문. 중앙대학교
- 김창범, 2015. 개입 승법계절 ARIMA와 인공지능경망 모델을 이용한 해상운송 물동량의 예측. 한국항만경제학회, 제31권, 제1호. pp.69-84.
- 김태일, 정봉민, 2012. 컨테이너선 시장 가격지수 동향과 전망. 한국해양수산개발원, 해양수산 5호, pp.128-143.
- 류동근, 2007. 해상운송의 이해. 다솜출판사.
- 박동주, 2012. 컨테이너 해운시장 공급요인과 상해운임지수 변동성간의 관계 분석. 석사학위논문. 한국해양대학교
- 박호진, 2002. 정기선 해운의 운임결정요인과 안정화 방안에 관한 연구. 석사학위논문. 한국해양대학교
- 양병철, 2015. 국제해운 시장의 정기선 운임결정 요인에 관한 연구. 박사학위논문. 연세대학교
- 이상원, 2002. 주가 예측을 위한 최적 인공지능경망 모델 선택에 관한 연구. 석사학위논문. 인제대학교
- 이슬기, 정성관, 이우성, 박경훈, 2011. 인공지능경망을 이용한 도시기온 예측 모형 구축. 대한국토도시계획학회, 제46권, 제 182호. pp.129-143.

이현호, 2013. 컨테이너 해상운임 예측 모형. 석사학위논문. 아주대학교

정용관, 윤영섭, 1998. 인공신경망모형을 이용한 주가의 예측가능성에 관한 연구. 한국재무관리학회, 제15권, 제2호, pp.369-399.

한국선주협회, 2018. 해사통계집.

<국외 문헌>

Awad, M., Pomares, H., Rojas, I., Salameh, O. and Hamdon, M., 2009. Prediction of time series using RBF neural networks: a new approach of clustering. The International Arab Journal of Information Technology, Vol. 6, No. 2, pp.138-143.

Dreiseitl, S., Machado, L.O., 2002. Logistic regression and artificial neural network classification models: a methodology review. Journal of Biomedical Informatics, Vol. 35, Issues 5-6, pp.352-359.

Oludolapo, O.A., Jimoh, A.A. and Kholopane, P.A., 2012. Comparing performance of MLP and R.B.F neural network models for predicting South Africa' s energy consumption, Journal of Energy in Southern Africa, Vol. 23, No. 3, pp.40-46.

Rout, M., Majhi, B., Mohapatra, U.M. and Mahapatra, R., 2012. Stock indices prediction using radial basis function neural network, Swam, Evolutionary, and Memetic Computing Lecture Notes in Computer Science, Vol. 7677, pp.285-293.

West, D., 2002. Neural network credit scoring models, Computers & Operations Research, Vol. 27, pp.1131-1152.