

해양플랜트 의장품 조달관리를 위한 배관 공정 리드타임 예측 모델에 관한 연구

함동균¹·백명기²·박중구³·우종훈^{1,†}
한국해양대학교 조선해양시스템공학과¹
서울대학교 조선해양공학과²
삼성중공업³

A Study of Piping Leadtime Forecast in Offshore Plant's Outfittings Procurement Management

Dong Kyun Ham¹·Myung Gi Back²·Jung Goo Park³·Jong Hun Woo^{1,†}
Naval Architecture and Ocean Systems Engineering, Korea Maritime and Ocean University¹
Dept. of Naval Architecture and Ocean Engineering, Seoul National University²
Samsung Heavy Industries³

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

In shipbuilding and offshore plant construction, pipe-stools of various types are installed. Moreover, these are many quantities but they must be installed in a successive manner. Due to these characteristics the pipe-stool installation processes easily tends to cause the schedule delays in the overall production processes. In order to reduce delay, the goal of this study is to predicts production's lead time before manufacturing. Through this predictions it's expected to reduce total production's lead time by improving it's process. First of all, we made MLR(Multiple Linear Regression) and PLSR(Partial Least Square Regression) model to predict pipe-spool's lead time and then compared predictability of MLR and PLSR model. If a explanatory variable is added, it will be possible to predict results precisely.

Keywords : Shipbuilding(선박건조), Production planning(생산계획), Simulation modelling(시뮬레이션모델링), Standardization(표준화), Neutral file format(중립파일포맷)

1. 서론

2010년대 심해 자원 개발에 대한 관심이 증대 되면서 해양플랜트 수요가 증가하였다. 당시 상선 수주 부진을 겪고 있던 국내 조선사들은 글로벌 해양플랜트의 수요를 거의 독점하다시피 하여 건조물량 확보에는 성공하였다. 하지만 부작용은 수주된 해양플랜트가 본격적으로 생산되기 시작하면서 나타났는데, 2014년부터 2015년까지 해양플랜트 부실로 총 8조5000억원 가량의 손실을 냈다 (Kim, 2015).

이러한 문제는 원인을 설계와 생산의 관점에서 두 가지로 나눌 수 있는데, 첫 번째는 설계엔지니어링 역량 부족으로 인한 저가수주에 따라 제작기간 중에 빈번한 설계변경을 들 수 있다 (Seong, 2014). 해양프로젝트가 대형화, 고사양화로 난이도가

높아지는 상황에서 수주를 턴키방식으로 진행하였고 건조 경험 부족과 까다로운 검사 기준은 설계가 자주 변경되는 현상을 발생시켰으며 이는 생산일정 지연과 건조비용 상승으로 이어졌다. 또한 두 번째 원인은 기존의 상선과 해양플랜트의 제조방식의 차이를 들 수 있다. 기존의 상선의 경우 후행작업에 비해 선행작업의 비중이 월등히 많았기 때문에 후행작업에 대한 설치계획 및 설치순서가 정형화되지 않았으나 해양플랜트 산업의 경우 안벽에서 이루어지는 후행의장작업의 비중이 늘어남에 따라 기존의 공정 시스템이 이를 수용하지 못하게 되어 일정지연이 발생하게 된 것으로 볼 수 있다. 본 연구는 2가지 요인 중 생산적 측면의 관점에서 문제를 다루고자 한다. 특히 의장품 중 후행작업에서 가장 많은 지연이 야기되는 배관 제작 및 설치 공정의 리드타임 예측을 통해 조달관리의 수준을 높이는데 목적으로 두고 연구를 수행하였다.



Fig. 1 Pipe spools used at offshore plant

배관 설치는 플랜트 프로젝트의 공종별 비중에서 43%를 차지하고, 공정관리 측면에서 주 공정(critical path)상에 있는 중요한 공종이다 (Kim & Shin, 2014). 또한 Fig. 1의 배관은 해양설비에서 요구하는 프로세스 특성에 따라 다양한 재질과 수량을 갖는데, 이러한 배관재들이 원활하고 효과적으로 공급되지 못해 하나의 1)spool이라도 제작에서 누락되면 전체 배관 라인을 형성할 수 없어 전체 공사 수행에 많은 지장을 주게 된다 (Im, 2012). 그러므로 배관의 효율적인 설치를 위해선 먼저 각각의 배관재들이 설치 장소에 도착시간을 예측 수 있어야 하고 만일 지연이 예측된다면 이를 해소할 수 있는 조치를 취해 주어 납기 결함으로 인해 비효율을 최소화 하여야 한다. 본 연구는 일련의 배관제작 조달 프로세스의 단계별 소요 기간 예측을 위해 다중 선형회귀분석 및 PLS 회귀분석 기법을 사용하여 예측 모델을 구성하고자 한다.

2. 관련 연구 동향

2.1 배관 공정 관련 연구

배관 공정과 관련하여 수행된 연구를 보면 배관 공급망의 개선의 관점과 배관 설치 공정순서 체계 정립의 관점에서 각각 분류될 수 있다 배관 공급망의 개선의 경우 배관이 제작에서부터 설치되는 장소까지 적기에 납품을 시킬 수 있는가에 대한 것으로 Lee (2013)은 배관의 주요공정 중 배관 의장품을 설치부서로 적기 납품하기 위한 목적으로 배관제작공정의 전체적인 핵심문제를 찾고 해결방안을 수립하기 위한 과정으로 제약이론(TOC, Theory of Constraints)을 근거로 DBR(drum-buffer-rope)시스템을 포함한 집중개선 5단계 기법을 활용하여 그 실현방안을 연구하여 구체적인 해결방안을 수립하였다. 이는 납기지연이 빈번하게 발생하는 배관제작공정에서 제약공정을 찾아내고 공정을 개선하여 납기준수율 및 후처리 공정준수율을 향상시킨 사례라 할 수 있다. 그러나 해당 연구의 경우 배관재들의 평균적인 특징을 가지고 공정 개선을 수행한다는 점에서 한계가 있다.

배관 공정과 관련하여 설치 공정순서 측면에서 살펴보면 Yan

1) 해양플랜트 및 육상플랜트에서 사용되는 수십 미터의 길이, 수 미터의 직경을 가진 배관을 운반 및 설치를 위해 적절한 길이로 절단한 단위 배관재를 말함.

(2012)은 호선에 탑재되는 의장품들의 조립 순서 알고리즘을 정의하고 실제로 배관에 적용하여 간섭 및 공정가능여부를 확인하는 연구를 수행하였다. 해당 연구에서는 실제 조립 시 고려되어야 하는 제약사항들을 알고리즘화하여 자동으로 조립 순서 정보를 도출하는 시스템을 구현하였다. 실제로 후행작업에 해당하는 배관 설치는 정형화된 설치계획 체계가 부재한 상황으로, 합리적인 설치 공정이 이루어지지 못하고 있는 시점에서 개선 가능성을 제시하였다는 점이 본 연구의 의미로 볼 수 있다. 다만 설치 순서를 최적화하는 부분은 모든 배관재들이 배치되어 있다는 점을 가정하고 있기 때문에 배관공정 전체 개선하기 위해선 앞서 언급했던 배관 공급망의 개선이 함께 고려되어야 한다.

2.2 제조 및 생산관련 예측 모델 구축 사례

제조분야에서 활용된 예측모델의 활용 대상을 살펴보면 많은 영역에서 센서를 활용하여 불량률을 예측하거나 특정 센서값을 예측하는 연구가 진행되고 있다. 불량률 예측과 관련하여서 Kim (2010)은 반도체 식각 공정에서의 CD값(critical dimension의 약자로 패턴의 거리, 즉 에칭이 된 부분의 폭을 말함) 예측을 위해 공정 진행 과정에서 얻을 수 있는 데이터를 활용하여 예측 모델을 구축하였다. 예측 모델 구축을 위해 온도, 압력, 전압 등의 공정 특성을 설명 변수로 하고 이 중 유의한 변수들을 선택하여 다중 선형 회귀분석과 비선형적 회귀접근 방식의 MARS(Multivariate Adaptive Regression Spline)을 사용하였다. 이를 통해 매 공정 후 거쳐야 하는 CD값 계측을 예측된 계측 값으로 대체하여 단계를 최소한으로 줄여 전체 생산 주기를 단축시키는 연구를 수행하였다. 본 논문의 경우 일반적으로 노이즈가 많이 포함되어있는 제조공정 데이터를 활용하여 비교적 높은 불량률 예측 결과를 내었다는 점에서 의미가 있다. 본 논문의 주제인 배관 공정 리드타임 예측 모델의 경우도 입력하는 과정에서 인적요소에 의해 노이즈 값이 포함되었을 가능성이 있으므로 선형 연구로써 참고할 여지가 많다고 할 수 있다. Ha (2013)은 LNG 냉매분리 공정의 효율적인 운전을 위한 특정 센서값 계측을 부분최소자승법을 이용하여 예측하는 소프트센서에 관한 연구를 수행하였다. 이 연구의 경우 여러 가지 변수들 간의 다중공선성 및 노이즈 값 발생에 대한 문제를 부분최소자승법을 사용하여 해결하였는데 본 연구에서도 이와 같은 문제를 해소하기 위해 다중 선형회귀분석과 함께 부분최소회귀분석을 활용하여 예측 모델을 구축하였다.

국내기업 사례의 경우 웅진 케미칼의 정수기 필터 사업부 예를 들 수 있는데, 해당 사업부는 제조 생산 데이터와 영업 및 기타 기업 내 운영데이터의 통합 분석 환경을 조성하여 제조에서 영업의 요구를 실시간으로 파악하고 데이터분석을 통해 적정 재고를 산출하여 생산계획에 이를 반영, 납기율 및 재고를 줄이는 결과를 도출하였다. 본 사례의 경우 제조생산 데이터와 함께 운영데이터를 종합적으로 분석하여 최적생산량을 예측하는 좋은 사례라 할 수 있다 (Jang, 2012).

또한 에너지 생산과 관련하여 한국남동발전의 발전설비 운영 효율 극대화 사례를 들 수 있는데, 해당 사례의 경우 시스템 및 연료 설비로부터 발생하는 데이터를 수집·분석하여 연료비 절

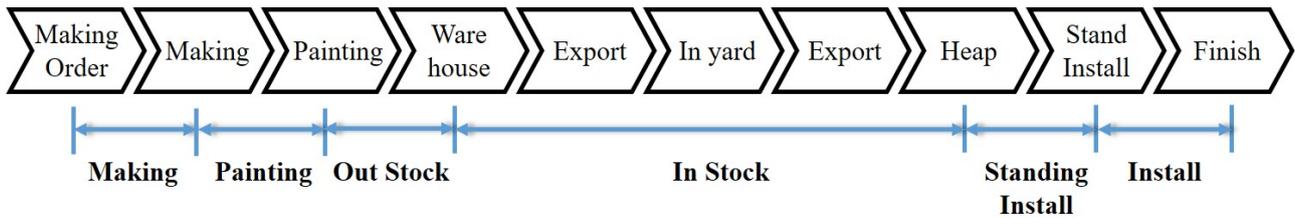


Fig. 2 Value chain of shipbuilding outfitting process

감 및 안정적인 설비 운영을 추진한 사례로 장비 고장을 예측하여 적절한 정비시기를 결정할 수 있는 예측 시스템을 구축하였다. 이로써 채선일 수를 최소화하고 운영효율을 극대화하여 2)비 계획발전손실을 0.02%를 절감할 수 있었다 (NIA, 2015).

3. 해양플랜트 배관의장 조달관리

3.1 해양 플랜트 의장 조달관리 프로세스

일반적으로 해양플랜트의 의장 조달관리 프로세스는 조선과 동일하며 해양플랜트의 경우 전체 공정 중 의장 작업의 비율이 조선보다 큰 것이 특징이다. 우선 조선소의 의장품 조달관리 시스템 구성요소를 보면 다음 Table 1와 같다 (Okumoto, 2012).

Table 1 Main component of outfitting procurement management system

Component	Description
Process schedule	Register ship's number, Job code, Work start day, Placement part day, Arranging day, Forwarding day, Changing day to database
Process demand from design	Register Product Info. and Delivery Info. in procurement database
Process Order from purchase	Decision cost base on Demand document, drawings and Spec.
Process for delivery deadline	Re-register Info. on changing delivery day
Process for warehousing inquiry	Register inspection Info.

여기에서 '설계에서 요구처리' 요소의 경우 물건 납입에 필요한 정보로 조정일수라는 것을 정하게 되는데, 이때 초기치는 조선소마다 다르지만 대개 장착개시일의 35~39일 전으로 설정된다. 이 조정일은 각 조선소의 입지환경, 유통환경 및 거래처와 계약조건 등에 다르게 매겨지는데 일반적으로 일회 요청 시 매우

2) 주어진 일정 시간 동안 갑작스런 고장 등 계획되지 않은 발전기 정지를

많은 주문량이 발생하기 때문에 한 번에 요청되는 건에 대해선 조정일수가 일괄적으로 부여되고 있다. 이 과정에서 필연적으로 납기일을 만족시킬 수 없게 되고 이러한 납기 지연은 의장품 설치 공정에도 영향을 미치는데, 가령 용접부서에서는 당일 설치하기로 한 의장품을 설치하기 위해 작업장에 갔는데 만일 의장품이 도착하지 않게 되면 그 기간 동안 공수가 허비되고 설치공정을 지연시키게 되는 것이다. 그러므로 주문 처리가 이루어지는 과정에서 의장 공급망 사이의 리드타임을 예측할 수 있다면 위의 문제처럼 공수를 낭비하는 것을 줄이고 의장품 전체 리드타임을 최적화하는데 기여할 수 있을 것이다.

3.2 배관 제작 및 설치 공정의 이해

해양플랜트 의장은 크게 배관 의장과 철의장으로 나눌 수 있는데 이 중에서 배관의장이 차지하는 비중이 상대적으로 큰 편으로 해양플랜트의 경우 적게는 3만개에서 많게는 10만개의 배관으로 이루어져 있다. 따라서 호선 건조 중 선형 의장 공사의 대부분을 차지하는 배관 의장 공사가 효율적으로 진행될 경우 선박 건조의 공기 단축을 이끌어 낼 수 있다. 배관의 주요 공정 중 배관제작과 설치간의 관계를 살펴보면, 호선 건조과정에서 특정 호선의 블록(block)이 형성되는 과정에서 그 작업장에 따라서 배관 의장 설치일정이 달라지는데, 일반적으로 최초 블록의 형태를 띠는 대조립에서부터 판별조립, 선형의장, 의장설치에 이르기까지 건조의 전 과정을 통해 배관설치작업이 이루어지고 있다 (Lee, 2013). 이렇게 배관 공정은 건조과정에서 한 부분에 한정되어 있지 않고 전 공정에 걸쳐 있기 때문에 다소 복잡한 구조의 공급망을 가지고 있다. 배관 공정의 일반적인 구조를 분석하면 Fig. 2와 같다. 아래의 그림에서 M/O(Making Order)발행일에서 제작일(making)까지의 기간을 제작기간(making)로 정의하였는데 이 공정은 외주제작업체로 M/O를 발행한 뒤 외주제작업체에서 배관의 제작이 완료되는 시점까지의 공정을 의미한다. 제작이 완료된 시점부터 도장협력사로 보내져 도장이 완료되는 시점을 도장기간(painting)으로 정의하였고 도장이 완료되는 시점에서 사내로 입고되는 시점까지를 사외재고(out stock)정의 하였는데 사외재고의 경우 사외 외부창고에 적치되어 있는 기간을 의미한다. 입고일부터 설치 작업을 위해 작업장 근처로 적치되는 적치일까지를 사내재고(in stock)로 정의하였고 적치일부터 불출요청을 받아 배관을 갑판(deck) 위로 올라가는 시점인 On-Deck까지를 설치대기(standing Install)로 정의 하였다. 마지막으로 On-Deck에서 설치가 완료되는 날까지를 설치(install)로 정의하였다. 위에

정의된 6개의 공정에서 2장에서 언급한 공급망 개선에 해당하는 부분은 제작일에서 사내재고까지로 볼 수 있고 설치 공정과 관련된 부분은 설치대기와 설치와 관련되어 있다고 볼 수 있다.

위의 분석한 설치 공정에서 가장 이슈가 되는 부분은 설치 (install) 단계에서 계획된 설치 날짜까지 배관들이 설치되는 장소까지 공급이 되지 않아 공수가 손실되고 계획 지연이 발생하는 부분이다. 이러한 문제를 해결하기 위해서는 구획별 설치 계획 시 배관들이 적치되는 날짜를 예측하여 가장 먼저 설치 준비가 되는 구획부터 설치되도록 하여야 하는데, 본 연구를 통해 M/O가 발행되는 시점부터 미리 리드타임을 예측하여 설치 순서를 계획하고 이를 통해 공정 흐름을 최적화하기 위한 목적으로 연구를 수행하였다.

4. 변수정의 및 분석방법

4.1 분석절차

본 연구의 분석절차는 Fig. 3과 같이 데이터 전처리, 공정시간 생성, 케이스 선택, 상관분석, 회귀분석의 순서로 진행하였고 분석과정에서 필요시 케이스 선택을 달리하여 Raw 데이터에 최적화 되도록 분석하였다.

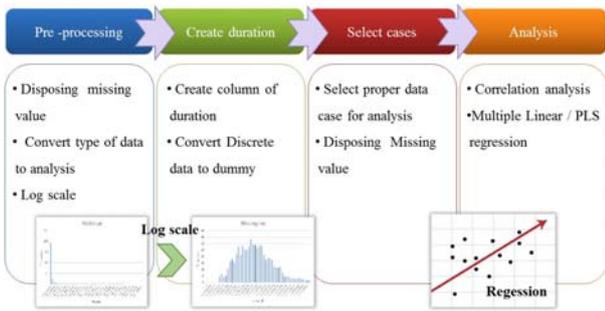


Fig. 3 Procedures to analyze spool data

Table 2 Definition of variable for analysis

Type	Description	Type	Description
Independent Variable	Making term(day) = Lead time to make pipe at vendors	Dependent Variable	Diameter = Diameter of pipe(inch)
	Painting term(day) = Lead time to paint pipe at vendors		Length = Length of pipe(mm)
	Out Stock term(day) = Lead time to be heaped at out-warehouse		Weight = Weight of pipe(kg)
	In Stock term(day) = Lead time to be heaped or moved in shipyard		Number of parts(piece)
	Standing install term(day) = Lead time to get somewhere to be installed		Number of Joint parts(piece)
	Install term(day) = Lead time to be installed		Term from 'DesignCompletionDay' to 'M/O(Making Order)'(day)
		Continuous data	Term from 'DeliveryRequestedDay' to 'M/O'(day)
		Discrete data	Vendors for making
			Heaped warehouse to completed spool
			Character of materials
			Tack welding stage before installing = Pre-weld Stage

먼저 현업 공급망 시스템의 Raw 데이터를 결측값 제거, 데이터 형식 변환, 로그스케일 등의 전처리 과정을 수행하여 분석 가능한 조건을 만들어 주었다. 또한 범주형 자료 4종에 대해서 더미변수를 생성하고 수치화하였다 (Lee & Im, 2015).

분석은 상용 통계 툴인 SPSS를 사용하여 다중선형회귀분석 및 PLS(Partial Least Square) 회귀분석을 수행하였고 분석과정에서 상관분석표와 유의확률 및 R²값을 참고하여 회귀분석을 수행하였다. 변수 선택과정에서 파이프 기본 스펙과 관련된 정보는 모든 공정에 반영하였고 범주형 자료의 경우 분석 결과를 확인한 뒤 반영 여부를 결정하였다.

4.2 변수정의

본 연구에서 활용한 데이터는 국내 조선소의 배관 제작에서 설치에 이르는 공급망 데이터로 공정완료날짜정보 및 파이프 사양, 제작정보 등이 포함되었다.

이중 파이프의 사양과 관련된 데이터인 직경, 길이, 재질, 중량, 부재수, Joint수의 경우 제품의 고유의 특성이 공정에 미치는 영향을 반영하고자 선정하였고 제품 고유의 특성 외에 외적으로 미치는 영향을 반영하기 위해 가접장, 설계완료날짜, 최종납품요청일, 외주제작업체, 외부창고와 같은 제작정보를 선정하였다.

선정된 정보들을 성격에 따라 종속변수, 독립변수로 분류하고 독립변수 안에서 변수 특성에 따라 수치형 변수와 범주형 변수로 분류하였다. 그 중 종속변수의 경우 전체 가치사슬을 6개의 공정으로 구분하고 위의 Table 2의 Making term부터 Install term까지 공정기간을 생성하여 사용하였다. 각 공정별로 살펴보면 제작공정의 경우 M/O 발행일부터 제작된 날짜까지, 도장공정은 제작일부터 도장일까지, 사외재고공정은 도장일부터 입고일 까지 사내재고공정의 경우 입고일부터 적치일 까지, 설치대기공정의 경우 적치일에서 On Deck일까지 설치의 경우 On Deck에서 설치완료일까지로 정의하고 기간일 수를 생성하였다.

4.3 분석모형의 이론적 고찰

본 연구에서는 배관의 6가지 공정기간을 예측하는 예측모형을 구축하기 위해 다중선형회귀분석과 PLS(Partial Least Square:일부 최소제곱)을 사용하였다.

다중회귀분석은 단순선형회귀의 확장으로 설명변수가 두 개 이상인 회귀모형에 대한 분석으로 회귀모형 식 (1)과 같다.

$$y = X\beta + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I) \quad (1)$$

여기서, 오차항은 평균이 0이고 분산이 일정한 정규분포를 따르고 서로 독립이라고 가정한다. 회귀계수벡터(β)와 오차항 벡터(ϵ)는 $\beta = (\beta_0, \beta_1, \dots, \beta_p)'$, $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_p)$ 이다. β 는 최소자승법으로 추정하며 잔차의 제곱합(Q)이 최소가 되도록 회귀계수를 구한다.

설명변수의 개수가 관측치보다 많거나, 설명변수들 간에 선형관계가 있는 경우 모수추정에 어려움이 발생할 수 있는데, 이러한 현상을 다중공선성(multicollinearity)이라고 한다. 다중공선성을 진단하는 방법으로는 분산팽창계수(VIF)를 이용하며, 다중공선성의 문제를 해결하기 위해서는 사전에 변수 선택 과정에서 상관관계수가 높은 변수의 유무를 확인하여 차원 축소를 하는 것이 바람직하다. 그러나 이 경우도 반응변수를 설명하는데 중요한 변수를 제외시킬 가능성이 있다. 설명변수를 제거하지 않는 방법으로 능형회귀, PLS회귀분석 등이 있다 (Kim, 2012).

PLS는 1966년 Herman Wold에 의해 사회과학, 특히 경제분야에 적용하기 위해 개발되었다. PLS는 주성분을 추출하는데 있어 설명변수뿐 만 아니라 반응변수간의 공분산을 추가로 고려하기 때문에 반응변수를 보다 효율적으로 예측할 수 있다는 장점을 지닌다 (Kim, 2012).

PLS 모델은 X(설명변수)와 Y(공정기간)의 상관관계를 이용하여 모델링을 한다.

$$X = \sum_a^A t_a p_a^t + E_A \quad (2)$$

$$Y = \sum_a^A u_a q_a^t + F_A \quad (3)$$

t_a 는 Y를 고려한 축소공간에서의 이력을 나타내는 a번째 주 성분 벡터, u_a 는 X를 고려한 Y축소공간에서의 이력을 나타내는 a번째 주 성분 벡터이다. p_a 는 $X^t X$ 의 고유벡터인 loading 벡터, E_A 는 잔차행렬(Residual matrix)이고, Y도 비슷한 의미를 갖는다. 그리고 X와 Y의 주 성분 벡터 사이의 내적관계(inner relation)로 회귀계수($b = u^t t / t^t t$)를 구해 모델링한다.

배관 제작 및 설치 공정은 의장품들 중에서도 변동성이 크고 주관적 판단에 의한 인적 오류의 가능성이 큰 공종이다. 따라서 차원감소 효과뿐만 아니라 다중공선성하 뛰어난 예측 성능을 가지고 잡음이 포함된 데이터를 분석하는데 효과적인 PLS 회귀 방법을 다중회귀분석과 함께 사용하여 분석한 뒤 비교할 것이다.

5. 분석결과

5.1 다중 선형회귀분석 모델

다중 선형회귀분석 시 동시입력 방식으로 수행하였고 설명력(R^2) 및 수정된 설명력(adjusted R^2)의 값은 높으면서 설명변수 개수를 최소화하는 방향으로 반복적으로 실행하면서 최적의 값을 찾는 형태로 수행하였다. 입력정보는 앞서 정의한 변수들과 같고 최종 선택된 변수들은 Table 3와 같다.

분석과정에서 사외재고의 경우 부재수, Joint수, 설계완료날짜-M/O일에 해당하는 계수가 유의확률 초과로 회귀식에서 제외되었다. 또한 사내재고, 설치대기, 설치의 회귀식에서 외주제작업체 중 삼성중공업거제와 성광에 해당하는 회귀계수가 케이스 선택으로 인한 결측값이 발생하여 제외되었다.

Table 3 Input data information of regression

Process	Multi linear	PLS
Making term	Dia, Length, Weight, Number of parts, Number of Joint parts, DesignCompletionDay-M /O(day), DeliveryRequestedDay-M /O(day), Making Vendor, Pre-weld Stage	Dia, Length, Weight, Number of parts, Number of Joint parts, DesignCompletionDay-M /O(day), DeliveryRequestedDay-M /O(day), Making Vendor, Pre-weld Stage, Warehouse, Material (Except Material=LT)
Painting term	Same Upside	Same Upside (Except Material=LT)
Out Stock term	Dia, Length, Weight, Number of parts, Number of Joint parts, DesignCompletionDay-M /O(day), DeliveryRequestedDay-M /O(day), Making Vendor, Pre-weld Stage, Warehouse	Same Upside (Except Material=LT)
In Stock term	Same Upside (Except Making vendor 6,7)	Same Upside (No except)
Standing install term	Same Upside (Except Making vendor 6,7)	Dia, Length, Weight, Number of parts, Number of Joint parts, DesignCompletionDay-M /O(day), DeliveryRequestedDay-M /O(day), Pre-weld Stage, Material (Except Material=LT, DSS, CA)
Install term	Same Upside (Except Making vendor 6,7)	Same Upside (Except Material=LT, DSS, CA)

5.2 PLS 회귀분석 모델

PLS회귀분석 모델의 경우 설정된 '잠재 요인의 최대 수'만큼 프로그램이 반복해서 분석을 수행하는데 본 연구에서는 기본 20으로 설정하여 분석한 뒤 설명이 최대가 되는 잠재 요인 수를 찾고 그 수만큼 다시 분석을 수행하여 회귀계수를 구하였다. PLS 회귀분석은 유의확률로 설명변수를 제거하는 다중 선형회귀분석과 달리 VIP(Variable Importance in the Projection) 크기를 사용하는데 VIP가 0에 가까운 경우 제외시킬 수 있다 (Jung, 2012).

분석결과에서 재질에 해당하는 변수가 다중 선형회귀분석에서는 공차한계로 인하여 사용하지 못하였으나 PLS 회귀분석에서는 다중공선성이 문제가 되지 않으므로 종속변수로 사용하였다. PLS 분석에서는 최대한 많은 변수들을 설명변수로 투입하여 회귀계수를 도출하였다. 다만 설치대기와 설치 공정의 경우 잠재요인 수 결정과정에서 종속변수 추가에 따른 설명력 증가가 미미하여 외주제작업체 및 외부참고 변수를 제외하였다. PLS 회귀분석의 경우 PLS 모듈이 명목형자료를 자동으로 더미변수로 바꾸어 준 뒤 분석하기 때문에 다중 선형회귀분석과 같이 직접 코딩할 필요는 없으나 지정 규칙이 다소 상이하므로 차후 회귀식 검증 시 유의해야 할 필요가 있다. 또한 설치대기와 설치 회귀식에서 재질 중 CA와 CS가 케이스 선택으로 인한 결측으로 인하여 제외되었다.

5.3 예측 모델 비교 분석 및 평가

앞서 도출한 회귀분석 모델은 Raw 데이터에 추출한 학습용 데이터를 분석하여 도출한 것이기 때문에 예측치 정확성을 검증하기 위해서는 검증용 데이터에 적용하여 정확도를 확인해보기로 한다.

검증용 데이터는 Raw 데이터에서 50% 랜덤 추출을 하여 학습용 데이터와 분리하였고 검증용 데이터의 실제값을 회귀식을 사용한 예측값과 비교하는 방식으로 검증을 수행하였다. 실제값과 예측값은 오차비율이라는 개념을 사용하였는데, 오차비율(식 (4))이란 예측값과 실제값의 차이와 실제값의 비율로 정의한다.

$$\text{오차비율} = \frac{|\text{예측값} - \text{실제값}|}{\text{실제값}} \quad (4)$$

또한 오차비율과 함께 오차비율이 30% 미만인 데이터 개수도 함께 구하여 전체 데이터 중 몇 프로 정도를 비교적 정확히 예측하는지 알 수 있도록 하였다. 오차비율을 30%를 기준으로 하여 그에 미만인 데이터의 개수를 구한 이유는 별도의 절대적 평가기준을 따른 것이 아니라, 후행일정계획 시 허용 가능한 오차정도를 임의로 선정한 기준이다. 다중 선형회귀분석의 모델 평가표는 Table 4와 같다. 표의 첫 번째 컬럼에서 R²항은 설명력에 해당하는 지표로 회귀식이 종속변수를 설명하는 정도를 의미하고 1에 가까울수록 완벽한 회귀식이 된다. 예측기간평균항은 회귀식으로

예측된 값들의 평균이고 실적기간평균항은 실적 리드타임의 평균값이다. 오차비율평균은 MAPE(Mean Absolute Percentage Error)로 불리는 평가 지표로 다수의 추정모형에 대해서 상대적인 정확도 비교에 우수한 지표로 알려져 있다. 오차비율<30항은 앞서 언급했던 것과 같이 비교적 정확한(30%) 데이터들의 분포를 알기 위한 지표이고 Total data항은 평가데이터의 레코드 수를 의미한다.

Table 4에서 제작 공정의 경우 오차비율평균이 68.5%로 전체 데이터의 35%정도가 30%미만의 오차를 보이고 도장 공정의 경우 오차비율평균 53.7%에 오차비율 30%미만 데이터 42%, 사내재고의 경우 오차비율평균 102%에 오차비율 30%미만 데이터 48%, 사내재고의 경우 오차비율평균 795.2%에 오차비율 30%미만 데이터 6%, 설치대기의 경우 오차비율평균 391.04%에 오차비율 30%미만 데이터 25%, 설치의 경우 오차비율평균 418.8%에 오차비율 30%미만 데이터 17%로 나타났다. 결과에서 제작, 도장, 사내재고의 경우 오차비율이 30% 정도의 데이터의 수가 전체의 40%정도로 나타나 비교적 정확한 예측치를 보였다. 반면 사내재고, 설치대기, 설치의 공정의 경우 오차비율의 평균이 다소 높고 예측의 정확도 역시 낮은 것으로 분석되었다. 특히 사내재고의 경우 실제값과 예측값의 평균에 차이가 있는데 이는 설치대기 데이터에 설명변수로 설명할 수 없는 요소가 존재하고 데이터에 노이즈가 다수 포함되어 나타난 현상으로 보인다.

Table 4 Verification for Multiple Linear Regression results

	Making	Paint	Out Stock	In Stock	Stand Ins.	Install
R ²	0.211	0.131	0.479	0.43	0.593	0.373
Predict Avg.	26.36	34.94	0.93	0.16	93.14	47.33
Origin Avg.	29.54	36.42	1.55	19.01	74.67	59.19
Over predict Avg. (%)	68.5	53.7	102.0	795.2	391.04	418.8
Over predict < 30 (%)	881 (35%)	1046 (42%)	1210 (48%)	153 (6%)	630 (25%)	432 (17%)
Total data	2474	2474	2474	2474	2474	2474

PLS 회귀분석 검증은 선형회귀분석과 동일한 방식으로 수행하였고 검증결과는 Table 5와 같다. Table 5에서 제작 공정의 경우 오차비율평균 62.78%에 오차비율 30%미만인 데이터가 전체 데이터의 37%를 차지하고 있고 도장의 경우 오차비율평균 120.55%에 오차비율 30%미만 데이터 25%, 사내재고의 경우 오차비율평균 57.46%에 오차비율 30%미만 데이터 53%, 사내재고의 경우 오차비율평균 387.1%에 오차비율 30%미만 데이터 9%, 설치대기의 경우 오차비율평균 201.9%에 오차비율 30%미만 데이터 19%, 설치의 경우 오차비율평균 418.7%에 오차비율 30%미만 데이터 18%로 나타났다. 결과에서 제작, 도장, 사내재고는 비교적 정확한 예측치를 보였고 선형 회귀분석과 마찬가지로 사내재고, 설치대기, 설치의 경우는 낮은 정확도를 보였다.

선형 회귀분석과 PLS의 회귀분석 결과를 비교해보면 Fig. 4와 같이 도장을 제외하고 PLS 회귀분석이 전반적으로 높은 정확도를 보였고 도장의 경우 PLS 회귀분석 보다 정확한 예측치를 보였다. 다만 설치대기를 보면 PLS의 오차비율평균이 선형 회귀분석 결과 보다 오차가 낮게 나왔지만 오차비율 30%미만으로 근사하게 예측한 데이터 수는 오히려 선형 회귀분석이 많았다. 이러한 결과는 데이터 내에 설명변수로 설명되지 않는 극대값이 다수 존재하는 것으로 유추할 수 있다.

PLS 회귀분석과 선형 회귀분석에서 공통적으로 사내재고가 가장 낮은 정확도를 나타내는 것을 확인할 수 있는데 실제 배관 설치 작업에서 공정 지연의 많은 비중을 차지하고 있는 부분이 바로 설치대기이다. 배관 설치의 경우 배관 라인에서 배관재가 하나라도 빠질 경우 설치를 시작할 수 없기 때문에 대기기간에 임의성이 커지고 이러한 부분이 누적될 경우 아래의 결과처럼 어떤 변수에 의해 설명되지 않으면서 큰 변동폭을 가지는 특성을 띄게 된다.

Table 5 Verification for PLS Regression results

	Making	Paint	Out Stock	In Stock	Stand Ins.	Install
R ²	0.228	0.173	0.480	0.484	0.509	0.130
Predict Avg.	27.26	34.25	1.75	17.76	59.70	48.59
Origin Avg.	29.54	36.42	1.55	19.01	74.67	59.19
Over predict Avg (%)	62.78	120.55	57.46	387.1	201.9	418.7
Over predict < 30 (%)	937 (37%)	613 (25%)	1334 (53%)	211 (9%)	461 (19%)	449 (18%)
Total data	2474	2474	2474	2474	2474	2474

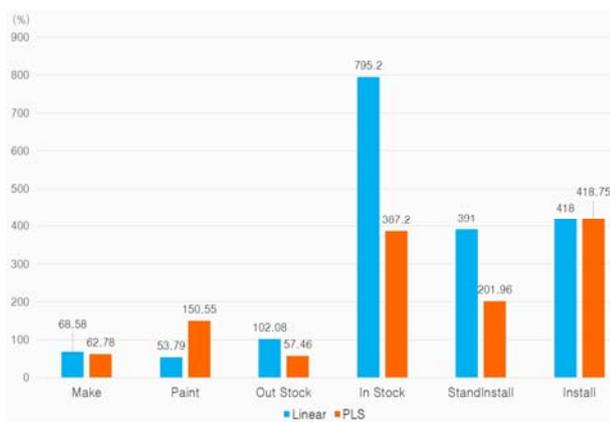


Fig. 4 Comparison graph of linear and PLS regression

5.4 분석모델 활용방안

위의 분석 모델은 단위 배관재의 리드타임을 예측하기 보다는 묶음단위의 리드타임을 예측하는데 활용하는 것이 바람직하다. 앞의 결과에서 단위 배관재에 대한 예측 정확도는 다소 낮으나

예측값의 평균은 실적데이터에 상당히 근접하는 것을 볼 수 있다. 실제로 현업에서는 제작 요청 시 설치 전까지의 리드타임을 전 실적데이터의 평균으로 일괄 배정 방식으로 진행하는데, 이 과정에서 Fig 5와 같이 특정 묶음 단위로 예측값을 구하여 계획을 수립할 경우 좀 더 최적화된 일정 관리가 가능할 것으로 보인다.

또한 공급망 관리의 특성상 공급망 데이터에 인적 오류 및 무작위성이 포함될 수 있는 여지가 많기 때문에 분석모델을 통해 얻은 효과가 단기적 또는 단위물량에 대한 예측에는 정확도가 상대적으로 떨어질 수 있으나 향후 조선소에서 나갈 방향인 관리 수준의 고도화가 진행되면 중기 예측 시스템에서 활용할 수 있을 것으로 생각한다.

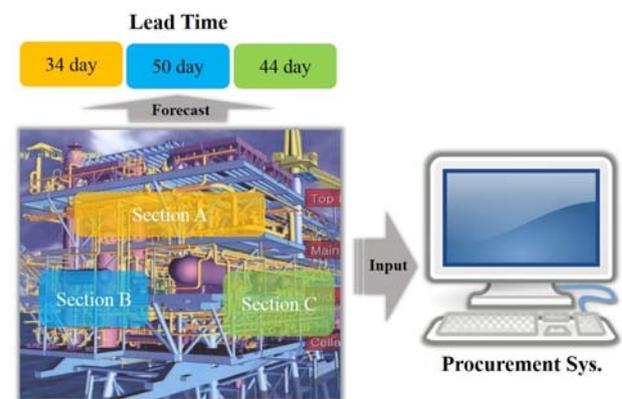


Fig. 5 An example utilizing forecast model

6. 결론

본 연구는 현재 조선 및 해양플랜트의 공정지연의 원인이 되고 있는 배관 제작 및 설치에 대한 예측모델을 구성하여 납기 지연에 의한 비효율을 최소화하고자 하였다. 예측모델을 구성하기 위해 먼저 전처리 작업을 수행하여 시스템상의 오류 및 결측값을 수정하고 본 연구에서 정의한 공정별로 기간 데이터를 생성하였으며 이 기간 데이터를 종속변수로 두고 다중 선형회귀분석과 PLS 회귀분석을 사용하여 회귀분석을 수행하였다. 전처리 작업 시 빈도분석 및 히스토그램 그래프를 통해 변수들의 극대값 여부를 확인하여 필요한 경우 로그스케일을 수행하였고 공정 분석 때마다 적절한 케이스 선택 기능을 활용하여 고유 데이터의 활용성을 높이고자 하였다.

본 연구에서는 데이터를 학습용과 평가용 데이터를 분리하여 학습과 평가를 각각 따로 수행하였고 이를 통해 과적합(over fitting)을 최소화하고자 하였다.

본 연구를 통해 얻은 결론은 다음과 같다.

첫째, 배관 제작 및 설치 공정에서 조선소 사외에서 이루어지는 공정(제작, 도장, 사외제고)의 경우 대체적으로 정확한 예측이 가능하였다. 제작과 도장, 사외제고의 경우 인적요소가 포함될 여지가 적고 외주제작업체 및 외부창고마다 기본적인 특성치가 정해져 있기 때문에 비교적 정확한 예측이 가능했던 것으로 보인다.

둘째, 선박 배관재 공급망 데이터 분석에서 다중 선형회귀분석 보다 PLS 회귀분석이 더 정확한 값을 나타내었다. 이는 배관재 공급망 데이터에 인적요소(주관적 우선순위 결정, 시스템 입력 실수 등)에 의한 노이즈 및 공선성 문제가 다중 선형회귀분석에 영향을 준 것으로 보인다.

셋째, 사내재고, 설치대기, 설치의 경우 이를 더 잘 설명할 수 있는 변수들이 추가되어야 한다. 분석결과를 보면 오차비율평균이 모두 200%를 넘는 수준, 실제값에 2배를 넘는 가까운 예측치를 보이고 있기 때문에 사실상 신뢰할 수 없는 회귀분석이다. 이러한 결과의 원인으로 입력변수의 신뢰성 문제와 제한적 변수를 들 수 있는데, 변수 신뢰성 문제의 경우 차후 개별 변수에 대한 민감도 분석 및 의사결정나무 분석을 수행하여 예측 성능 향상을 저해하는 요인들을 분석하는 연구를 수행할 예정이다.

다만 평균적으로 오차가 크다는 것과는 별개로 특정비용(30%)안에 있는 예측치가 존재하기 때문에 만약 적절한 변수가 추가 될 경우 회귀식의 예측력이 올라갈 수 있는 여지는 남아있다고 볼 수 있고, 비교적 예측이 정확했던 데이터의 특징들을 분류하여 부분적으로 회귀분석을 수행한다면 좀 더 정확한 예측 모델을 만들 수 있을 것으로 보인다.

후 기

본 과제(결과물)는 교육과학기술부의 재원으로 한국연구재단의 지원을 받아 수행된 산학협력 선도대학(LINC) 육성사업과 산업통상자원부 산업핵심 기술개발사업 중소형조선소 시뮬레이션 기반 생산관리 시스템 개발 과제(과제번호:10050495)의 지원을 받아 수행되었습니다.

References

Ha, D.G., 2013. *Soft sensor design on fractionation process for LNG plant using PLS and variable selection*. Master's Thesis. Seoul National University.

Im, J.S., 2012. *Improvement for individual work order process in manufacturing piping spool of offshore facilities*. Master's Thesis. University of Ulsan.

Jang, Y.J., 2012. Utilizing Big-data's Technology in Manufacturing Field. *Journal of The Korean Institute of Communication Sciences*, 29(11), pp.30-36.

Jung, G.S., 2012. Presumed Influence Factors of Decision-Making of Mega-Projects. *Audit and Inspection Research Institute*, 19(1), pp.91-124.

Kim, D.J., 2012. *Models of local finance decisions*

using PLS regression analysis : cities and counties in Gyeongnam. Master's Thesis. Gyeongsang National University.

Kim, H.J., 2010. *A study on forecasting model of in semiconductor device fabrication*. Master's Thesis. Korea University.

Kim, J.S. & Shin, Y.T., 2014. Development Portable Pipe Spool Location-Confirm System Based UHF RFID. *KIPS Transactions on Computer and Communication Systems*, 3(10), pp.329-336.

Kim, S.S., 2015. The Three Heavy Industries Operating Loss Sharply. *Monthly Maritime Korea*, 2015(6), pp.84-87.

Lee, H.S., 2013. *A case study on improvement of pipe outfitting production process using TOC*. Master's Thesis. University of Ulsan.

Lee, H.S. & Im, J.H., 2015. *SPSS 22 Manual*. 1st Ed. Jyphyuntae: Seoul.

National Information Society Agency(NIA), 2015. *2015 global big data Casebook*. National Information Society Agency: Korea.

Okumoto Yasuhisa, 2012. *Shipbuilding Technology & Production System*. Translated by Kim, Y.S. GS-intervision: Seoul.

Seong, H.G., 2014. A Proposal for Rowth of Offshore Plant Industries. *Machinery Industry*, 2014(9), pp.41-48.

Yan, W., 2012. Automatic Generation of Assembly Sequence for the Planning of Outfitting Processes in Shipbuilding. *Journal of Ship Production and Design*, 28(2), pp.49-59.



함동균

백명기

박중구

우종훈