

Research Paper

공공청사 개산건적 정확도 향상을 위한 공사비 영향요인 분석

Analysis of Impact Factors for the Improvement of Conceptual Cost Estimation Accuracy for Public Office Building

조영호¹ · 윤석현^{2*}

Jo, Yeong-Ho¹ · Yun, Seok-Heon^{2*}

¹Researcher, Department of Architectural Engineering, Gyeongsang National University, Gajwa-Dong, Jinju, 52828, Korea

²Professor, Department of Architectural Engineering, Gyeongsang National University, Gajwa-Dong, Jinju, 52828, Korea

*Corresponding author

Yun, Seok-Heon
Tel : 82-55-772-1755
E-mail : gfyun@gnu.ac.kr

ABSTRACT

A Conceptual cost estimate, which is computed in the preliminary step of a project, is important for decision-making by a contractor in terms of the project budget, economic feasibility and validity analysis, and alternative comparisons. Therefore, a high error rate of a prediction model for a conceptual cost estimate can lead to various problems including excessive project expenditures and a delayed break-even point. this study proposed optimal impact factors by configuring quantitative impact factors computable in a preliminary step in various cases(combinations of impact factors). subsequently, the accuracy of different cases was comparatively analyzed by using the cases as input values of a prediction model using regression analysis. when the optimal combination of impact factors proposed in this study and other combination of impact factors were applied to the prediction model, the regression analysis-based prediction model exhibited 0.2-4.7% improvements in accuracy, respectively. the optimal combination of impact factors proposed in this study improved the accuracy of the prediction model of a conceptual cost estimate by removing unnecessary impact factor.

Keywords : conceptual cost estimate, regression analysis, impact factor

1. 서론

1.1 연구의 배경 및 목적

건설 프로젝트 기획단계에서 이루어지는 타당성 및 경제성 분석은 건설사업관리에 있어서 매우 중요한 요소이며, 이는 향후 건설 프로젝트의 진행에서도 큰 영향을 미친다. 따라서 정확한 타당성 및 경제성 분석에 있어서는 적정공사비 산정 즉 프로젝트 초기단계에서 이루어지는 예산은 건설 프로젝트의 성패를 좌우할 만큼 많은 영향을 미친다[1]. 그러나, 프로젝트 초기단계에서 공사비 예측을 위해 활용할 수 있는 정보는 매우 제한적이며, 신속하게 공사비를 예측해야 하므로 현재 국내에서는 단위면적당 공사비를 정보를 이용한 공사비 예측이 가장 많이 활용되고 있다. 이는 제한된 정보와 신속하게 공사비를 산출해야 하는 초기단계에서 활용도는 우수하나 정확도가 낮아 이후 프로젝트 진행에 있어 많은 문제점이 발생할 수 있다.

이에 국내에서는 초기단계에서 보다 정확한 공사비를 예측을 위해 사례기반 회귀분석 및 기계학습, 딥러닝 등 다양한 개산건적 방법에 관한 연구가 진행되었다. 하지만, 예측모델을 작성하는 과정에서 다양한 방법을 적용하여 정확도를 비교하는 연구는 진행되고 있으나 모델을 학습하는 과정에 공사비에 상관없는 정보 및 영향요인 간의 독립성이 확보되지 않은 영향요인



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

(속성값)을 활용하여 예측모델을 학습시켜 예측 모델의 제 성능을 발휘 못 하는 경우가 발생하고 있다. 본 연구에서는 현재 기획단계 개선견적에서 활용 가능한 영향요인 조합을 구성하여 회귀분석을 통한 최적의 영향요인 조합을 제시하고자 한다.

1.2 연구의 방법 및 범위

본 연구는 최근 3년간 공공기관에서 발주한 공공청사 프로젝트 51건을 연구 대상으로 개선견적 모델 학습 과정에 있어서 최적의 영향요인 조합을 제시하는 것을 연구의 범위로 한다.

연구의 방법으로는 우선 이론적 고찰을 통해 개선견적관련 선행연구를 고찰하고 개선견적의 범위 및 기준을 정의하고 공사비 영향요인을 선정하였다. 선정된 영향요인은 상관분석을 통해 산출되는 상관계수를 활용하여 영향요인 간의 공사비에 영향을 미치는 영향요인 식별하였다. 이후, 다양한 영향요인 조합을 구성하여 SPSS통해 회귀분석을 진행한다. 회귀분석을 통해 각 영향요인 조합의 독립성 검증 및 회귀식을 산출하였다. 독립성 검증은 다중공선성(Multicollinearity)을 활용하여 VOF, TOL계수를 활용하였다. 산출된 회귀식은 검증 데이터 셋을 활용하여 가장 정확성이 높은 영향요인 조합을 선정하도록 하였다. 연구의 흐름도는 Figure 1과 같다.

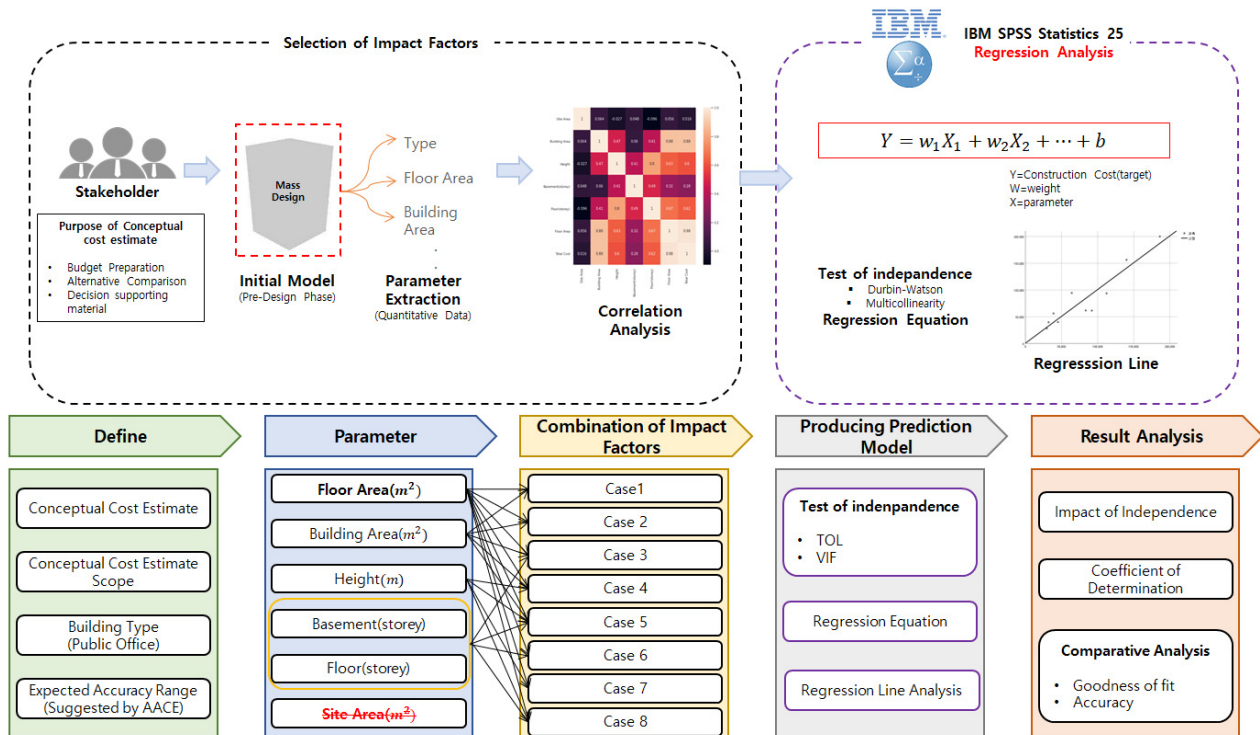


Figure 1. Research flow

2. 이론적 고찰

2.1 선행연구 분석

개선견적은 각 기관 및 전문가에 따라 그 정의가 다르다 이에 선행 연구를 고찰하면서 개선견적의 범위 및 정의하고 개선견적 모델 학습에 있어서 활용되고 있는 영향요인을 파악하고자 공사비 예측 모델 작성과 관련된 연구 문헌을 고찰하였다.

Song[2]은 공동주택을 대상으로 하여 개산견적 시 영향요인을 다양하게 반영하지 못하는 점을 문제 삼아 각 독립변수와 종속변수의 관계를 산점도를 활용하여 독립변수와 종속변수와의 상관관계를 산출하였다. 이후 헵타그램 분석을 활용하여 골조 개산수량의 기준이 되는 단위면적당 수량산출식을 제시하였다. 해당 연구에서는 설계년도, 층수, 평형, 세대수, 연면적을 활용하여 골조 수량산출식을 제시하였다.

Kim et al.[3]은 프로젝트 초기단계에서 발주자의 견적 및 공사비 관리 업무를 보조하기 위해 공공청사의 물량산출 방정식을 제시하였다. 해당 연구에서는 구조체의 부위별 단위면적당 수량을 산출하기 위하여 연면적, 대지면적, 건축면적, 층수, 구조형식, 기초형식, 데크 유무를 영향요인으로 구성하여 단위면적당 산출식을 제시하였다.

Park and Kim[4]은 회귀분석 및 인공신경망을 활용한 개산견적 방법이 학습율, 학습 횟수, 노드수 변경 등 연구자 경험에 따라 고려해야 할 요소가 많은 것을 해소하기 위해 SVR방법을 활용하여 공동주택 공사비를 산출하였다. 해당 연구에서 사용된 영향요인으로서는 연면적, 층수, 총 세대수, 공기, 지붕 형식, 기초형식, 지하층, 마감수준을 활용하여 총 공사비를 예측하였다. 하지만, 사용된 영향요인 중 마감수준에 대한 등급 분류는 기준이 제시되어 있지 않아 견적자 따라 정확도의 차이가 발생할 수 있다고 판단된다.

Son et al.[5]은 교육시설을 대상으로 프로젝트 개념단계에서의 인공신경망을 활용한 공사비 예측 방법을 제시하였다. 해당 연구에서는 활용된 영향요인으로서는 대지면적, 건축면적, 연면적, 학급 수, 지하층 면적, 지상층 면적, 예정 공사기간을 활용하여 회귀분석과 인공신경망의 정확도를 비교하였다. 하지만, 해당 연구에서 사용한 학습모델의 데이터 수가 부족한 것으로 판단된다.

Yim et al.[6]은 공공건축물공사를 대상으로 공사규모에 따른 공사비 산출방법을 제시하였다. 해당 연구에서는 공사규모에 따라 데이터를 분류하여 회귀분석을 통한 예측 모델을 구축하였다. 하지만, 해당 연구에서 사용된 공사비 영향요인은 연면적 하나로 건축물의 특징을 반영하는 요소가 제한적이라고 판단된다.

Jung et al.[7]은 스마트 교육시설을 대상으로 기획·설계단계에서 산출이 가능한 연면적, 건축면적, 주차장 면적, 교실 등을 활용하여 초기단계에서의 총 공사비를 예측하였다. 해당 연구에서는 사용되는 영향요인은 상관분석 및 회귀분석을 통한 독립성 검증이 이루어지지 않아 인공신경망 예측 모델간의 넓은 오차율 범위가 발생한 것으로 판단된다.

국내 연구 사례를 분석해본 결과 영향요인으로 사용되는 공통적인 정량적인 요소로는 연면적, 지상 층, 지하 층, 대지면적, 건축면적 등이 있으며, 정량적인 요소로 사용되는 마감수준, 대지 상태 등은 견적자의 능력 및 판단에 따라 많은 오차가 발생할 수 있다고 판단된다. 따라서 본 연구에서는 정량적인 요소에 대한 영향요인을 대상으로 연구를 진행한다.

2.2 개산견적의 범위 및 정의

개산견적의 정의는 각 기관 및 전문가에 따라 그 정의가 다르다. 하지만 건설 프로젝트의 초기단계에서 제한된 정보로 예산을 산정하고 그에 따른 타당성 및 경제성 평가에 활용된다는 점에서는 그 의미가 동일하다. 따라서 본 연구에서는 선행 연구된 개산견적의 범위를 분석하여 활용할 수 있는 영향요인을 분석 및 개산견적의 시기에 따른 활용도에 대해 분석하고 해당 연구에 적합한 범위를 정의하고자 한다.

Kim et al.[8]은 개산견적의 범위를 설계가 시작되기 전 단계인 기획단계부터 개산견적의 범위로 정의하였다. 이를 통해, 공사비를 사전에 미리 예측 및 분석하여 향후 발생할 수 있는 제반의 문제를 사전에 파악하여 대처할 수 있으며, 설계가 진행되는 과정에서 최적의 대안을 선정할 수 있다고 주장하였다.

Seo et al.[9]은 개산견적의 범위를 기획단계부터 시작하지만 견적의 첫 단계가 개산견적을 활용한 방법이 아닌 단위면적당 단가를 활용한 개념견적을 계산한 후 개산견적을 진행하는 것으로 정의하였다. 해당 연구에서는 기획단계에서 개념견적을 진행하지만 의사결정이 큰 파급효과를 가지고 있으며, 후속적으로 진행되는 의사결정의 평가기준이 된다고 주장하였다[9].

Son et al.[10]은 개산견적의 범위를 기본설계 단계 이후로 정의하였다. 해당 연구에서는 수량산출을 활용한 개산견적 방

법을 사용하였으며, 이는 기획단계에서의 불확실성한 요소를 제거할 수 있다고 하였다.

Park et al.[11]은 BIM을 활용한 개산견적 방법을 사용하였으며, 초기계획 및 개념설계단계에서 정보의 제약으로 인해 오차범위가 넓어 타당성 분석 및 예산 추정에 정확도를 향상시키고자 하였다. 해당 연구에서는 BIM모델을 치수산출이 가능한 수준의 모델을 작성하여 개산견적에 필요한 변수를 산출하여 정확도를 향상시키고자 하였다. 해당 연구에서 개산견적의 범위는 BIM모델의 LOD를 고려하였을 때 기본설계 이후가 그 범위에 해당된다고 판단된다.

국내의 개산견적의 문헌을 고찰해본 결과 기획단계부터 개산견적을 진행하는 경우는 프로젝트 예산 편성 및 향후 프로젝트를 진행함에 있어서 발주자의 대안선정에 의사결정을 지원해줄 수 있으며, 주로 활용되는 영향요인은 연면적, 지하층, 지상층 건축면적, 대지면적이 활용되고 있다.

개념설계 이후에 개산견적을 수행하는 경우에는 주로 연면적, 지하층 면적, 지상층 면적, 실 정보, 치수정보를 이용한 수량산출 기반의 예측 모델을 활용하였으며, 기획단계에서의 불확실성 요소를 제거하여, 더욱 정확한 예측을 할 수 있게 된다. 하지만 국내의 건설 프로젝트 특성상 예산산정이 기획단계에서 이루어지며, 경제성 및 타당성 분석이 진행되기 때문에 개산견적 시점이 프로젝트가 시작하는 기획단계에서부터 진행되어야 할 것으로 판단된다. 따라서 본 연구에서는 개산견적의 범위를 기본설계 도면이 생성되기 전의 기획단계로 정의하고자 한다.

2.3 예측모델의 적합성 평가

국내에는 개산견적의 예측 정확도에 대한 평가 기준이 없으며, 미국 AACE(American Association of Cost Engineering, 이하 AACE)에서는 ‘2013 Cost Estimate Classification System’에서 견적 정확도 평가기준을 제시하고 있다. Figure 2는 AACE에서 제시하는 오차범위 기준이다. 여기서는 프로젝트 진행도에 따라 5단계로 나누어 오차범위를 제시하였다.

ESTIMATE CLASS	Primary Characteristic	Secondary Characteristic		
	MATURITY LEVEL OF PROJECT DEFINITION DELIVERABLES Expressed as % of complete definition	END USAGE Typical purpose of estimate	METHODOLOGY Typical estimating method	EXPECTED ACCURACY RANGE Typical variation in low and high ranges
Class 5	0% to 2%	Concept screening	Capacity factored, parametric models, judgment, or analogy	L: -20% to -50% H: +30% to +100%
Class 4	1% to 15%	Study or feasibility	Equipment factored or parametric models	L: -15% to -30% H: +20% to +50%
Class 3	10% to 40%	Budget authorization or control	Semi-detailed unit costs with assembly level line items	L: -10% to -20% H: +10% to +30%
Class 2	30% to 75%	Control or bid/tender	Detailed unit cost with forced detailed take-off	L: -5% to -15% H: +5% to +20%
Class 1	65% to 100%	Check estimate or bid/tender	Detailed unit cost with detailed take-off	L: -3% to -10% H: +3% to +15%

Figure 2. Cost estimate accuracy range(AACE)

Class 5단계는 프로젝트 초기단계에서 매우 제한된 정보를 기반으로 작성되며, 빠르게 공사비를 산출하는 단계이다. 프로젝트가 0~2%정도 진행되었을 때 시행되며, 주로 프로젝트의 위치 및 발주자의 요구사항을 취합해서 산출하는 단계이다. 해당 단계에서 사용되는 주된 목적은 시장 조사, 초기 실행 가능성 평가, 대체 계획 평가, 예산 책정 등 프로젝트의 경제성 평가의 기준이 된다. 본 연구에서는 초기 단계의 개산견적을 대상으로 하고 있으므로, Class 5단계에 기준을 사용하고자 한다. Class 5의 허용오차 범위는 프로젝트의 기술적 복잡성 및 정보의 위험에 따라 -20 ~ -50%에서 30 ~ 100%를 기준으로 한다.

3. 영향요인 분석

3.1 개산건적의 영향요인

영향요인은 프로젝트 진행정도에 따라 산출되는 정보가 달라진다. Figure 3에서 보는 바와 같이, 개산건적의 영향요인은 크게 정량적 영향요인과 정성적 영향요인으로 구분된다. 정량적 영향요인은 해당 프로젝트 진행 정도(설계단계)에 산출이 가능한 정보를 활용하는 경우이며, 정성적 영향요인은 해당 단계에서 산출이 불가능한 정보를 영향요인으로 활용하기 위해 사용된다.

일반적으로 마감 재료 및 기계설비 등의 정보는 프로젝트가 진행된 기본 설계 단계 이후에서 산출이 가능하다. 그러나 이러한 영향요인을 건적 주체의 기준에 따라, 등급으로 구분하여 해당 단계에서 적용이 불가능한 영향요인을 정성적 영향요인으로 활용할 수 있다. 이는 프로젝트의 특성을 예측 모델에 최대한 반영하여 정확도를 향상시키기 위해 사용할 수 있다. 하지만, 정량적 영향요인만을 활용하는 경우 산출되는 값이 일관성을 갖게 되지만, 정성적 영향요인은 건적 주체에 따라 그 분류가 달라 예측 모델의 일관성이 저하시킬 수 있다. 즉, 정량적 요인은 개산건적 예측 모델을 활용하는 건적 주체의 부담을 줄여줄 수 있다. 따라서 본 연구에서는 기획단계에서 활용할 수 있는 정량적 영향요인으로 활용할 수 있는 대지면적, 연면적, 건축면적, 높이, 지상층수, 지하층수를 대상으로 최적의 공사비 영향요인을 제시하고자 한다.

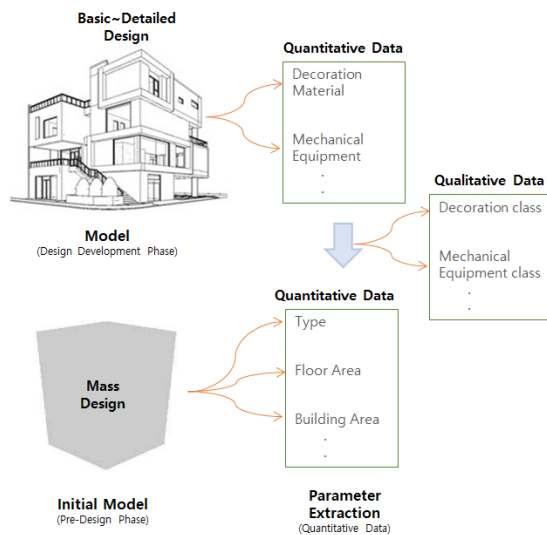


Figure 3. Process for selecting impact factors of conceptual cost estimate

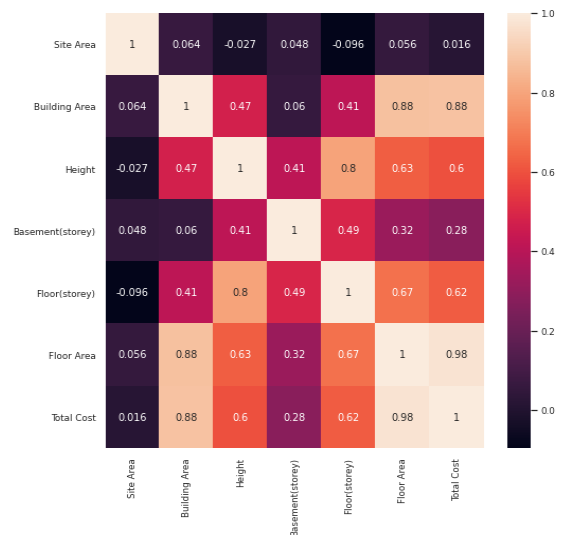


Figure 4. Coefficient of correlation

3.2 영향요인 결정계수

기획단계에서 산출되는 정량적 영향요인이 공사비에 미치는 상관계수를 확인하기 위하여 앞서 제시된 6가지 영향요인을 대상으로 총 공사비와의 상관분석을 진행하였다. 분석대상의 영향요인으로는 대지면적, 건축면적, 높이, 층(지하층, 지상층), 연면적, 총 공사비를 분석한 결과는 Figure 4와 같다.

분석 결과 연면적의 상관 계수가 0.98로 가장 높았으며, 그 외 건축면적, 지상층, 높이, 지하층, 대지면적 순으로 상관계수가 높았다. 0.016으로 상관계수가 가장 낮은 대지면적의 경우 총 공사비와의 상관성이 부족하다고 판단되어 이후 분석 단계에서 제외하였다.

3.3 영향요인 특성

Table 1은 기획단계에서 활용할 수 있는 영향요인을 수직, 수평으로 분류한 표이다.

Table 1. Horizontal and vertical cost estimate factors

Type	Factor
Horizontal Factor	Total Floor Area, Building Area
Vertical Factor	Height, Storey(Floor, Basement)

건물의 수평적인 특성을 나타내는 영향요인에는 연면적과 건축면적이 있다. 연면적은 건축물의 바닥면적의 합계를 나타내기 때문에 건물의 대략적인 규모를 파악하기에 가장 적합한 영향요인으로 판단된다. 건축면적은 건축물의 외벽 또는 중심선으로 둘러싸인 부분을 나타내며 일반적으로 1층 바닥면적을 가리킨다. 건축면적은 총 공사비와 높은 상관관계를 갖고 있으나, 건물의 전체적인 규모를 설명하기에는 부족할 수 있다. 건물의 수직적인 특성을 나타내는 영향요인에는 높이, 층(지상, 지하)이 있다. 지상층과 지하층의 상관계수는 각각 0.62, 0.28의 상관계수로 분석되었다. 이는 건축물의 규모가 커지는 경우 지하층 보다 지상의 층수가 증가하는 경우가 많아, 지하층의 상관계수가 낮게 나타나는 것으로 판단된다.

상관분석 결과, 연면적은 수평적인 영향요인이지만 수직 영향요인과의 상관계수 또한 높은 것을 알 수 있다. 따라서, 건축물의 전체적인 규모를 설명하기에 있어 연면적이 가장 높은 설명력을 가지고 있다는 것을 알 수 있다. 하지만 예측 모델의 입력값으로 활용되는 경우 연면적은 건축물의 특성을 전부 반영하기에는 부족하다고 판단되기 때문에 본 연구에서는 연면적을 중심으로 영향요인의 최적의 입력변수 조합을 제시하고 그 정확도를 분석하고자 한다.

3.4 예측모델 설정

상관분석 결과, 상관계수가 가장 높은 영향요인인 연면적을 기본으로 포함하는 8가지의 영향요인 조합을 구성하였다. 구성된 8가지의 영향요인은 Table 2와 같다.

Case 1은 제일 높은 설명력을 가지고 있는 연면적만으로 구성하였으며, Case 2~5는 연면적과 건축면적으로 영향요인을 조합하였는데, Case 5는 모든 영향요인을 포함시켰다. Case 6~8은 연면적과 건축면적 간의 상관계수가 높은 것을 고려하여 건축면적을 제외하고 연면적을 중심으로 나머지 영향요인을 조합하여 구성하였다.

Table 2. Combination of factors

Type	Impact factors
Case 1	TA
Case 2	TA + BA
Case 3	TA + BA + BS + FS
Case 4	TA + BA + H
Case 5	TA + BA + H + BS + FS
Case 6	FA + BS + FS
Case 7	FA + H
Case 8	FA + H + BS + FS

TA = Total Area, BA = Building Area, H = Height, BS = Basement(storey), FS = Floor(storey)

3.5 학습 데이터 셋 구성

데이터 셋은 51건의 일반 청사 프로젝트를 대상으로 하였으며, 전체 프로젝트 중 41건은 학습데이터로 사용하고 10건은 검증용 테스트 데이터로 분류하였다.

4. 회귀분석

연구에서 설정한 공사비 영향요인 조합에 대한 회귀분석 결과는 Table 3과 같다.

Table 3. Regression coefficients by each Case

Type		Unstandardized Coefficient		Standardized Coefficient	t	TOL	VIF
		B	SE	β (Beta)			
Case1	intercept	284.736	3,797.64		0.075		
	FA	15.726	0.359	0.990	43.817	1.000	1.000
$R^2 : 0.980, F : 1,919.894, P\text{-value} : 0.000$							
Case2	intercept	-107.209	3,904.44		-0.27		
	FA	15.335	0.825	0.965	18.589	0.193	5.185
	BA	1.993	3.774	0.027	0.528	0.193	5.185
$R^2 : 0.980, F : 942.336, P\text{-value} : 0.000$							
Case3	intercept	9,105.9	10650.8		0.855		
	FA	16.446	1.470	1.035	11.191	0.063	15.964
	BA	-1.647	5.688	-0.023	-0.290	0.087	11.430
	FS	-1,721.6	2,248.9	-0.030	-0.766	0.352	2.840
	BS	-2,422.0	5,306.0	-0.015	-0.456	0.495	2.022
$R^2 : 0.981, F : 457.357, P\text{-value} : 0.000$							
Case4	intercept	-3,511.0	7,104.93		-0.494		
	FA	15.112	0.918	0.951	16.461	0.159	6.308
	BA	2.306	3.846	0.032	0.600	0.189	5.291
	H	181.879	315.951	0.017	0.576	0.621	1.610
$R^2 : 0.980, F : 617.281, P\text{-value} : 0.000$							
Case5	intercept	9,184.2	10,489.6		0.876		
	FA	17.027	1.501	1.072	11.341	0.058	17.180
	BA	-4.778	6.001	-0.066	-0.796	0.076	13.117
	H	573.946	394.565	0.053	1.455	0.391	2.556
	FS	-3,870.0	2,662.1	-0.067	-1.454	0.244	4.102
	BS	-4,805.4	5,476.5	-0.030	-0.877	0.450	2.221
$R^2 : 0.982, F : 377.651, P\text{-value} : 0.000$							
Case6	intercept	6,997.4	7,675.7		0.912		
	FA	16.047	0.504	1.010	31.847	0.520	1.924
	FS	-1,456.0	2,027.6	-0.025	-0.718	0.422	2.367
	BS	-1,538.7	4,287.4	-0.010	-0.359	0.739	1.354
$R^2 : 0.981, F : 625.263, P\text{-value} : 0.000$							
Case7	intercept	-2,565.6	6,869.0		-0.374		
	FA	15.589	0.455	0.981	34.245	0.634	1.578
	H	155.112	310.133	0.014	0.500	0.634	1.578
$R^2 : 0.980, F : 941.615, P\text{-value} : 0.000$							
Case8	intercept	3,839.1	8,018.3		0.479		
	FA	15.904	0.513	1.001	31.027	0.494	2.023
	H	461.291	366.440	0.043	1.259	0.449	2.227
	FS	-2,776.9	2,269.0	-0.048	-1.224	0.332	3.011
	BS	-2,104.9	4,277.7	-0.013	-0.492	0.731	1.369
$R^2 : 0.981, F : 476.754, P\text{-value} : 0.000$							

4.1 회귀계수

본 연구에서는 SPSS를 활용하여 앞서 설정한 8가지 Case에 대해 다중회귀분석을 진행하였다. 분석 결과, 8가지 예측 모델의 설명력(R^2)값은 0.980 이상으로 높은 설명력을 가지고 있는 것으로 판단된다. 이후 영향요인 간의 독립성을 검증하기 위해, Case별 VIF계수를 분석하였다. 상관분석에 영향요인 간의 상관관계수가 높았던 연면적과 건축면적의 조합(Case 2)에서는 두 영향요인 간의 VIF계수가 5.185로 나타나, 독립성이 확보되는 것을 확인하였다. 그러나 연면적, 건축면적에 총 영향요인을 활용한 Case 3과 5의 경우 연면적과 건축면적의 VIF계수가 10을 넘어 독립성이 확보되지 않는 것으로 판단된다. 이렇게 분석된 회귀계수 결과는 Table 5, 6과 같다.

4.2 회귀식

제시된 회귀계수 결과를 기반으로 하여 회귀식을 도출한 결과는 Table 4와 같다.

Table 4. Regression formula

Type	Formula
Case 1	1,570,098.521*FA + 65,632,768.03
Case 2	1,530,658.542*FA + 200,972.363*BA + 25,651,059.78
Case 3	16,46,075.876*FA + 172,507.221*BA - 237,389,340*BS + -182,560,935*FS + 972,930,963.8
Case 4	1,511,212.914*FA + 230,562.884*BA + 18,187,300*H - 351,073,211.862
Case 5	1,702,707.966*FA - 477,769.434*BA + 57,393,573.03*H - 386,999,157*BS - 480,557,114*FS + 918,444,076
Case 6	1,604,264.349*FA - 155,090,478*FS-144,529,938*BS+754,104,594.731
Case 7	1,558,872.957*FA +15,510,630.35*H - 256,538,918.478
Case 8	1,590,433.980*FA + 46,127,935.39*H - 277,681,833*FS - 210,512,121*BS +383,923,907.3

5. 모델 정확도 검증

5.1 사례 검증

Table 4는 다중회귀분석을 통해 독립성이 확보되지 않은 영향요인 조합(Case 3, 5)을 포함하여, 각 영향요인 조합의 회귀식을 보여주었다. 이들 회귀식의 검증을 위하여, Test 데이터 10건을 적용하여, 각 영향요인 조합의 정확도를 비교 분석하고자 하였다. 이때, 오차율의 산정식은 식 (1)과 같다.

$$Error\ Rate(\%) = \frac{(Y_i - \hat{Y}_i)}{Y_i} \quad (1)$$

Y_i : 실제 값(W)

$\hat{Y}_i$: 예측 값(W)

회귀모델의 정확도 검증을 위한 테스트 셋은 Table 5와 같다.

Table 5. Test dataset

Project	Cost (100,000₩)	TA (m ²)	BA (m ²)	FS	BS	H (m)
1	186,078	12,686	2,758	6	1	28
2	32,065	2,501	975	3	0	13
3	112,317	5,938	3,418	3	1	15
4	83,512	3,899	953	4	2	15
5	29,576	1,769	661	3	1	14
6	45,070	2,520	872	3	0	14
7	38,872	3,546	716	3	1	17
8	64,123	3,885	3,574	4	1	21
9	91,906	5,980	2,099	4	1	19
10	139,944	9,917	2,169	5	1	23

5.2 정확도 검증

5.1에서 제시한 테스트 데이터 셋 10건을 회귀모델에 적용시킨 결과는 Table 6과 같다.

Table 6. Regression model error rate

Project	Case Error Rate(%)							
	1	2	3	4	5	6	7	8
1	7.39	7.47	7.73	7.25	7.39	7.64	7.20	7.25
2	24.51	26.30	36.41	21.42	34.34	34.14	19.97	29.03
3	16.41	12.74	16.56	13.79	23.29	13.90	17.80	15.63
4	25.90	25.93	27.89	27.79	34.82	26.95	27.53	31.31
5	3.88	3.10	0.94	7.97	6.90	0.82	8.31	5.97
6	10.77	9.97	1.88	13.32	2.00	3.91	13.83	7.04
7	44.90	43.98	51.80	41.06	53.16	50.04	42.41	48.39
8	3.85	4.34	9.80	4.76	22.40	2.97	4.59	3.48
9	2.88	4.47	3.23	3.47	0.55	4.27	1.79	2.65
10	11.73	11.77	12.71	11.08	11.78	12.50	11.13	11.44

분석 결과, 8가지 영향요인 조합의 모델 모두 AACE에서 제시한 Class 5단계의 오차 범위를 만족하였다. 영향요인 조합의 조합의 정확도는 Case 2, 4, 1, 7, 6, 8, 3, 5순으로 높았으며, 표준편차는 Case 4, 7, 2, 1, 8, 6, 3, 5순으로 나타났다. Case 3, 5의 경우는 영향요인의 간의 종속적인 관계가 존재하기 때문에 통상적인 해석을 하였을 경우 심각한 오류가 발생할 수 있다.

본 연구에서 제시한 회귀모델을 분석한 결과, 연면적과 건축면적을 입력값으로 활용한 Case 2의 정확도가 가장 높았으며, 종속적인 관계가 있는 Case 3, 5의 회귀모델의 정확도가 가장 낮은 것으로 나타났다. 이는 영향요인 간의 독립성이 확보가 되지 않은 경우 예측 결과 값에 많은 영향을 주는 것을 확인할 수 있었으며, 회귀모델에 최대한 많은 영향요인을 활용하는 것이 정확도 향상에 직결되지 않는다고 판단된다. Figure 5는 이러한 결과를 보여주고 있다.

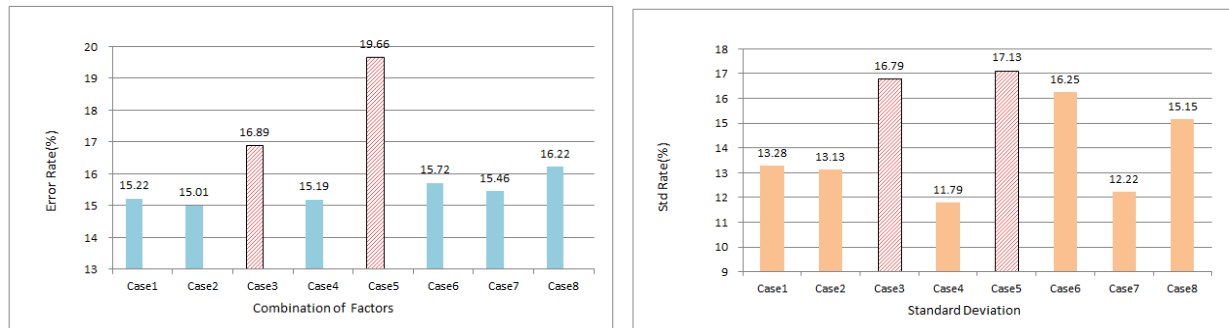


Figure 5. Average error rate & Average standard deviation rate

6. 결론

본 연구는 영향요인 조합에 따른 개산견적 예측모델의 정확도 변화를 비교 및 분석하여 최적의 영향요인 조합을 제시하고자 하였다. 기존의 선행 연구되었던 개산견적 예측모델은 최대한 많은 영향요인을 활용하여 예측 모델의 정확도를 향상시키고자 하였다. 그러나 예측 모델에 활용되는 영향요인 간의 종속적인 관계가 있는 경우 예측 모델의 일관성 및 정확도에 오히려 부정적인 영향을 미친다. 이에 본 연구에서는 상관분석 및 회귀분석을 활용하여 영향요인 간의 독립성을 검증하고 기획단계에서 산출이 가능한 정량적인 영향요인을 8가지 조합으로 구성하여 정확도를 비교하였다. 본 연구 결과를 요약하면 다음과 같다.

첫째, Case1, 2, 4, 6, 7, 8의 경우 영향요인 간의 독립성이 확보되었으며, Case 3, 5의 경우 연면적과 건축면적 간의 독립성이 확보가 되지 않았다.

둘째, 회귀계수를 기반으로 회귀식을 도출하여 정확도를 검증해본 결과, Case 2, 4, 1, 7, 6, 8, 3, 5순으로 정확도가 높았다. 이는 독립성이 확보되지 않은 회귀모델은 정확성에 많은 영향을 주는 것을 확인할 수 있었다. 즉 개산견적의 예측 모델의 정확도를 향상을 위해서는 최대한 많은 영향요인을 활용하는 것보다 최적의 영향요인을 조합하여 입력값을 활용하는 것이 더 적합하다고 판단된다. 본 연구에서는 Case 2에 활용된 연면적과 건축면적의 영향요인 조합이 정확도가 가장 높았다.

본 연구는 회귀분석을 통해 산출되는 다중공선성 계수를 통해 영향요인 간의 독립성을 검증하는 방법 및 최적의 영향요인 조합을 제시하였다. 이는 기존 선행연구에서 많이 사용되는 영향요인 조합인 CASE5의 비해 약 4.65% 정도 정확도가 향상된 것을 확인하였다. 따라서 개산견적 예측 모델 구축에 앞서 본 연구에서 제시한 조합 및 프로세스를 활용할 경우 보다 정확한 예측이 가능한 것을 판단된다.

향후 연구결과를 활용하여 인공지능망 기반의 예측 모델에 적용시켜 정확도를 향상시킬 수 있는 연구가 필요하다고 판단된다. 또한 본 연구는 개산견적의 범위를 기획단계에서 활용가능 정량적인 영향요인을 대상으로 분석하였지만, 향후 기획단계 뿐만 아니라 기본설계 단계에서 산출이 가능 영향요인을 활용하여 최적의 영향요인 조합을 제시하는 연구가 필요하다고 판단된다.

요약

본 연구는 기획단계에서 이루어지는 개산견적 예측 모델의 정확도를 향상시키기 위하여 최적의 영향요인 조합을 제시하였다. 이에 기획단계에서 활용이 가능한 정량적인 영향요인을 선정하여 상관분석 통해 공사비에 가장 많은 영향을 주는 연면적을 중심으로 8가지의 영향요인 조합을 설정하였다. 8가지 영향요인 조합을 다중회귀분석을 통하여 VIF계수 및 회귀식

을 도출하였다. VIF계수를 통해 연면적, 건축면적과 층 영향요인을 함께 사용할 경우 연면적과 건축면적 두 영향요인 간의 종속적인 관계를 확인하였다. 이에 독립성이 예측 모델 정확도의 관계를 분석하기 위하여 실 사례 프로젝트 10건을 회귀식에 대입하여 정확도를 분석하였다. 분석결과, 독립성이 확보가 안 된 영향요인 조합은 다른 영향요인에 비해 정확도 떨어지는 것을 확인할 수 있다. 따라서 최대한 많은 영향요인을 활용하는 것보다 최적의 영향요인 조합을 선정하는 것이 예측 모델의 정확도를 향상시킬 수 있다고 판단되며, 본 연구에서는 연면적과 건축면적을 활용하였을 경우 정확도가 가장 높은 것을 확인하였다.

키워드 : 개산건적, 회귀분석, 영향요인

Funding

Not applicable

Acknowledgement

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport(Grant 21AATD-C163269-01).

ORCID

Yeong-Ho Jo,  <http://orcid/0000-0001-9326-8206>

Seok-Heon Yun,  <http://orcid/0000-0001-5439-4111>

References

1. Ahn YS, Song KR, Heo JM. Improving the accuracy of screening of cost estimating in early construction project phase. *Journal of the Architectural Institute of Korea Structure & Construction*. 2003 Nov;19(11):133-40.
2. Song KR. A study on the standard for approximate quantity of framework of apartment housing for using scatter diagram & heptagram method. *Journal of the Architectural Institute of Korea Structure & Construction*. 2005 Oct;21(10):167-74.
3. Kim SM, Cho JH, Lee JS, Chun JY. Cost estimating model of structural elements using approximate quantity survey in preliminary building design phase. *Journal of the Architectural Institute of Korea Structure & Construction*. 2009 Dec;25(12):155-64.
4. Park UY, Kim GH. A study on predicting construction cost of apartment housing projects based on support vector regression at the early project stage. *Journal of the Architectural Institute of Korea Structure & Construction*. 2007 Apr;23(4):165-72.
5. Son JH, Kim SK, Kim JO. A study on the analysis and estimation of the construction cost by using artificial neural network in the btl projects for educational facilities. *Journal of the Architectural Institute of Korea Structure & Construction*. 2008 Jun;24(6):136-42.
6. Yim JH, Park JM, Kim OK. Cost estimating method of public building construction through construction scale. *Journal of the Korea Institute of Building Construction*. 2015 Jun;15(3):307-16. <https://doi.org/10.5345/JKIBC.2015.15.3.307>
7. Jung SH, Gwon OB, Son JH. A study on the analysis and estimation of the construction cost by using deep learning in the smart

- educational facilities -Focused on planning and design stage-. Journal of the Korean Institute of Educational Facilities. 2018 Nov;25(6):35-44. <https://doi.org/10.7859/kief.2018.25.6.035>
8. Kim KH, Park CS, Chang SH. Web-based cost planning program for high-rise office building. Korean Journal of Construction Engineering and Management. 2005 Apr;6(2):69-79.
 9. Seo JK, Kim SM, Cho JH, Chun JY. Supporting model of cost planning using case-based reasoning in pre-design phase-focused on the case of public office building. Journal of the Architectural Institute of Korea Structure & Construction. 2010 Apr;26(4):159-68.
 10. Son BS, Park MS, Lee HS, Lim DH. Analysis the schematic cost estimating model based on quantity variation in building projects using the case study. Journal of the Architectural Institute of Korea Structure & Construction. 2008 Jun;24(6):109-18.
 11. Park YJ, Won SK, Han CH, Lee JB. A study on 3D BIM collaborative approximate estimating model of structural work for apartment projects. Journal of the Architectural Institute of Korea Structure & Construction. 2011 Jun;27(6):123-30.