

DB 품질개선사업의 대가 산정 모형연구

서용원* · 이덕희** · 정승호*** · 박건수****

A Development of Cost Estimation Model for Data Quality Analysis and Improvement Project

Yong Won Seo* · Duck Hee Lee** · Seung Ho Jung*** · Kun Soo Park****

■ Abstract ■

As interests in the quality of data in database systems are growing recently, analysis and improvement of data quality in databases have been an important issue. However, there has yet to be a clear agreement on how to reasonably calculate the total cost of such project. In this paper, based on real project data and budget statistics, we develop a model to estimate the cost for quality analysis and improvement project of a database. We first conduct statistical analysis to build our basic model. Throughout this analysis, we have identified factors that determine the scale of works required to conduct the project and eventually determine the cost. In addition, we have identified factors that determine the complexity of the project. These factors can adjust the cost determined by the scale of works. Our model is verified and improved by surveys on experts. We apply our model to newly conducted projects and observe that our model estimates the cost of each project reasonably well.

Keyword : Data Quality Improvement Project, Cost Estimation Model, IT Project Budgeting Plan

1. 서 론

공공데이터 개방법에 따라 공공기관의 데이터 품질에 대한 관심도가 높아져 다수의 기관들이 데이터품질 진단 및 개선사업을 진행하고 있다. 하지만, DB 품질개선사업의 적절한 예산산정의 기준이 부재하여 공공기관이 정보화 사업을 계획하거나 실제 구축 사업을 추진하는 데 있어 어려움을 겪고 있다.

데이터 품질 진단 및 개선사업은 실제 사업 수행 시에 고려하지 못한 다양한 위험요소들로 인하여 적절한 사업의 예산 수립이 쉽지 않다. 예를 들어, 품질 진단 및 개선사업의 특정 활동을 수행하기 위한 비용은 사업 대상의 데이터 크기(또는 테이블의 수나 크기)에 주로 영향을 받을 수 있으며, 다른 활동의 비용은 주로 사업 대상 데이터가 다른 데이터와 연관된 개수에 주로 좌우될 수 있다. 또한 특정 활동은 사업의 크기나 난이도에 상관없이 항상 일정한 비용이 발생하는 편이다. 이에 따라 데이터베이스 품질 진단 및 개선사업의 예산을 객관적이고 합리적으로 추정할 수 있는 대가모형의 필요성이 대두되고 있다.

지금까지의 데이터 품질진단 및 개선을 위한 예산산정은 주로 유사사업의 수행에 의한 경험치를 통한 인건비 산정방식이 주로 적용되었다. 이러한 방식의 비용 산출은 소요예산 비용의 객관성을 확보하기 어려워 동일한 유형의 사업도 편차가 많이 발생할 수 있는 개연성이 있다. 따라서 공공데이터 품질 진단 및 개선을 위해 예산을 산정하거나 사업규모를 파악하기 위한 기준을 제공하는 것을 목적으로, 합리성과 공정성을 확보한 체계적인 대가 산정 기준이 필요하다. 즉, 데이터베이스 품질진단 및 개선사업의 사업비 추정을 비용 산정의 핵심인 비용의 결정 요인과 영향 요인을 파악하여 객관적이고 쉽게 적용 가능한 산식을 확립하고자 한다.

본 연구는 기존에 다른 사업들에 대하여 수립된 사업비 대가 산정 모형들에 대한 분석을 통해 예산의 효율적 집행 및 DB 품질개선사업 수행에 적

용 가능한 대가 산정 기준을 구축하고자 한다. 이를 통해 공공기관과 사업자들이 동시에 참조할 수 있는 DB 품질개선사업의 대가 기준을 마련할 것으로 기대된다.

2. 문헌 연구

과거 데이터 품질 진단 및 개선사업 대가는 별도의 대가 산정기준 없이 담당자의 경험이나 선행사업의 예산을 참조하여 대가가 산정되었으며, 2012년에 수립된 대가 산정 모형에서는 '비용, 데이터 양, 테이블복잡도'의 3가지 요인에 따라 대가가 결정되고 있다(National Information Society Agency, 2012). 그러나 실제 데이터 품질 진단 및 개선사업의 대가는 각 활동별 난이도 및 특성에 따라 영향을 받는 복잡한 구조로 되어 있다. 이에 따라 본 연구에서는 이러한 사업 특성을 고려함으로써 대가 모형의 객관성과 합리성을 높이는 것에 중점을 두고 있다.

먼저, 본 연구에서 개발하는 모형의 대상사업인 정보화사업의 효율성 제고에 관한 연구를 살펴보면 다음과 같다. 강운식(Kang et al., 2007)은 서비스 요청모형을 비용모형을 전환하여 측정하는 IT 운영비용 최적화 방안을 제시하였다. 전성철(Jun et al., 2008)은 건설 프로젝트에서 주로 사용되는 획득 가치관리(EVM : Earned Value Management)를 국내 소프트웨어 개발 프로젝트에 적용함으로써, 프로젝트 기간 중의 일정 및 비용을 예측하는 방안을 제시하였다.

한편 본 연구의 대가 모형 개발에 참고할 수 있는 대가 산정 기준으로는 건설사업관리(CM : Construction Management), 정보시스템감리, 정보전략계획(ISP : Information Strategy Planning), PMO(Project Management Office) 대가 기준 등이 있다. 먼저, 건설사업관리 대가 기준을 살펴보면, 국토해양부(Ministry of Land, Infrastructure, and Transport, 2009) 및 한국건설교통기술평가원

(Korea Agency for Infrastructure Technology Advancement, 2003)은 건설사업의 규모(공사비)와 건설사업관리 적용대상 단계의 2개 요소로 대가 결정 모형을 수립하였다. 대체적으로 건설사업관리 대가는 공사비로 파악되는 건설사업 규모의 4~10% 범위에서 산출되는 것으로 나타났다. 이와 유사하게 정보시스템관리 대가 기준(National Information Society Agency, 2006; National Information Society Agency, 2008)에서는 감리 적용대상 단계와 감리 대상사업비의 2개 요소를 통해 대가를 계산한다. 본 연구에서는 이러한 대가 모형과 유사하게 먼저 사업의 규모를 결정하는 요소들을 파악하여 이들이 사업비에 미치는 영향을 분석해 대가 모형에 반영하고자 한다.

정보전략계획 대가기준(Korea Software Industry Association, 2013)의 경우, 업무별 가중치와 업무 난이도의 2개 요소로 대가를 결정하는 모형으로 수립되어 있다. 이때 업무별 가중치는 정보전략계획을 구성하는 업무들의 비용이 전체 사업의 비용에서 차지하는 비율을 계산한 것이며, 업무 난이도는 대상 사업 업무의 난이도를 결정하는 조직규모, 업무처리유형 등의 요인들을 반영한 난이도 계수를 고려하고 있다. 또한 유사 모형 중 하나인 소프트웨어 개발비 산정연구(Park et al., 2009)는 Function Point 방법(IFPUG, 2004)에 기반을 두어 애플리케이션의 특성을 난이도 보정계수로 반영하는 개발비를 연구하였다. 본 연구에서는 위의 대가 모형들을 참고하여 사업의 범위를 결정하는 작업 요소들을 나누고, 해당 요소별로 작업의 난이도를 결정하는 요소들을 파악하는 방식으로 데이터 품질 진단 및 개선사업의 대가를 산정하고자 한다.

한편, PMO 대가기준(National Information Society Agency, 2013)의 경우, PMO 대상사업비와 PMO 적용대상 단계, 대상사업 특성(난이도)의 3개 요소를 통해 대가가 결정되는 모형이다. PMO와 관련된 구체적 지침은 현재 시행 초기 단계로써, 대가 모형 개발에 데이터 분석 중심 방법을 적용하

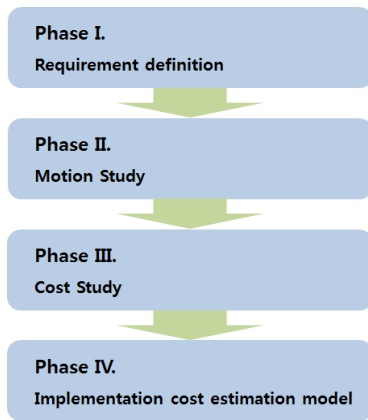
기가 부적합하다. 이에 따라 PMO 대가 산정 모형의 개발 연구(Seo et al., 2014)에서는 이러한 현실적 제약이 존재하는 상황에서 전문가 대상의 추정 에 의한 설문조사법을 활용하여 대가 산정모형에 소요되는 요인들의 추정치를 산정하고 모형을 수립하는 방법에 대해 연구하였다. PMO 대가기준과 유사하게 데이터베이스 구축사업 대가기준(National Information Society Agency, 2013)에서는 수행규모, 수행범위, 수행복잡도의 3개 요소를 통해 대가를 계산한다. 본 연구에서는 위의 대가 모형들을 참고하여 사업의 범위 및 규모를 결정하는 요소들과 해당 요소별 난이도를 결정하는 요소들을 파악하여 대가 모형 수립에 반영하고자 한다.

이상과 같이 사업대가 산정을 위한 관련 연구들이 존재하지만, 데이터 품질 및 진단사업에 적용할 수 있도록 작업 요소별 특성을 반영한 모형은 찾아볼 수 없다. 관련 연구에서 시사하는 바와 같이 기존의 대가 산정기준들은 데이터 품질 및 진단사업에 직접적인 적용이 어렵지만, 품질 및 진단사업 대가를 결정하는 요인 도출시 참고 기준으로 사용 가능하다. 앞장에서 기술하였듯이 데이터 품질 및 진단사업은 여러 종류의 단위 작업이 존재하며, 이들 작업의 비용이 사업의 규모나 사업의 난이도에 받는 영향은 상이한 것으로 나타났다. 따라서 본 연구에서는 관련 대가 산정기준들을 참고하여, 데이터 품질 진단 및 개선사업의 특성을 반영하여 사업의 작업 요소별로 대가 규모 및 난이도 결정 요소들을 고려한 대가 산정기준을 개발하고자 한다.

3. 대가 산정절차

본 연구에서는 합리성과 공정성을 확보한 DB품질개선사업 대가 산정 기준의 수립을 목적으로 한다. 이를 위해 DB 구축대가, PMO 대가 등에 적용되어 온 5단계의 방법론을 활용하여 대가 모형을 수립한다.

대가 산정방식의 수립은 <Figure 1>에 정리된 Phase I~IV의 총 4단계로 진행된다. 먼저 'Phase I. 요구사항정의' 단계에서는 대상 사업의 작업 단위를 사업 대가에 영향을 주는 최소 단위로 세분화한다. 수립된 작업 단위는 대가 산정 대상 사업의 범위를 결정한다.



<Figure 1> Methodology to Develop Cost Estimation Model

'Phase II. 작업요소 모형수립(Motion Study)' 단계에서는 Phase I에서 수립된 작업 요소별로 사업 수행에 필요한 작업의 규모를 결정하는 요소들을 파악하여 대가 결정 요소로 정의한다. 대가 결정 요소의 도출은 2013년 데이터 품질 및 개선사업 데이터 분석과 전문가 사전조사 인터뷰를 통해 작업 요소별로 사업 대가를 결정하는 대표 결정 요소를 하나씩 도출하여 초기 모형의 수립에 사용하였다. 또한 작업요소 모형 수립 단계에서는 사업의 수행 난이도를 결정하는 난이도 보정요소를 파악하여 결정 요소와 함께 초기 모형 수립에 사용하였다. 즉, 비슷한 규모의 작업이라도 더 높은 난이도를 가진 작업들이 포함된 사업의 대가가 상대적으로 더 크게 산정될 수 있도록 보정 요소들을 대가 모형에 사용하였다.

'Phase III. 작업량 산출(Time Study)' 단계에서는 앞 단계에서 도출된 대가 결정 요소 및 대가 보정 요소들이 사업에 투입된 공수의 양과 어떤 관계

를 가지고 있는지 각 작업 요소별로 파악하였다. 관계식의 도출을 위해 2013년 사업 데이터의 통계적 분석과 전문가 설문을 통한 델파이 분석(Lins-tone and Turoff, 2002) 을 병행하였다.

'Phase IV. 대가모형 수립'에서는 수립된 모형으로 산정한 사업의 투입공수로부터 사업비와의 관계를 파악하여 최종 사업의 대가를 산정하는 단계이다. 이를 위해 사업의 투입 직급별 노임단가, 기술료율 및 제정비율을 추정하여 최종 대가 산정 모형식을 도출하였다.

4. 작업모형 및 영향요인 도출

4.1 작업요소 도출

본 연구에서는 각 사업 별로 작업의 단계를 세분화하여, 데이터 품질 진단 및 개선 사업에 포함되는 작업의 범위를 명확히 선정하였다. 이는 사업별로 수행되는 작업의 범위가 서로 다를 수 있음을 대가 산정에 고려할 수 있게 해준다. 작업요소의 도출은 기본적으로 공공정보 품질관리 매뉴얼의 품질진단 및 개선절차를 기반으로 하여 사업대가에 영향을 미치는 작업 단위의 특성이 같은 단위로 재편하였다. 도출된 작업 요소들은 사업의 수행 규모 및 사업의 난이도를 결정하는 기본 단위로 사용되며, 선정된 대가 작업요소는 <Table 1>과 같다.

4.2 결정요인 도출

본 장에서는 앞에서 수립된 작업 요소별로 사업의 규모에 영향을 주는 결정요인을 도출한다. 결정요인의 도출은 우선 1차로 소규모 전문가 자문 회의를 통해 작업요소별 대표적 대가 결정요인을 파악하고, 2013년 사업자료의 분석을 통해 적합성을 분석하였다. 이후 2차례의 전문가 설문 조사를 통해 수립된 결정요인의 적절성을 최종적으로 검증하였다.

이상의 절차를 통해 간단하고 객관적으로 사업의 발주 단계에서 파악 가능한 규모 지표로 테이블수,

조직규모, 연계기관 수 세 가지를 파악하였다. 특히 작업 대상 테이블 수는 사업의 수행 활동별로 상이할 수 있기 때문에 테이블 수는 전체진단테이블 및 심층진단테이블의 두 가지로 세분하였다. 선정된 규모결정 요인은 <Table 1>과 같다.

<Table 1> Project's Scale Factors

Factors		Definition
Number of tables	Total tables	Total number of tables excluding temporary, backup, history, package tables.
	Tables for in-depth diagnosis	Public open, master, and other tables where quality issues are raised and agreed to be investigated upon the contractor and owner.
Organization scale		Total number of divisions that use the database covered in the project.
Related organization number		Total number of external organizations that are related to the database covered in the project.

4.3 보정요인 도출

보정요인은 결정요인에 의해 산정된 사업의 투입공수를 증가 또는 감소시킬 수 있는 사업 환경적 요인으로 데이터의 복잡도, 구조의 복잡도, 조직의 복잡도 등을 들 수 있다. 본 연구에서는 객관적으로 쉽게 측정 가능하고 규모 결정요인에서 고려되지 않은 요소로 전문가 조사를 통해 <Table 2>

<Table 2> Project Complexity Adjusting Factors

Factors	Definition
Existence of data model	Complexity of the project increases if there is no data model.
Data model up-to-date	Complexity of the project increases if the data model is not updated.
Existence of data standard	Complexity of the project increases if the data standard is not developed.
Operation of quality management system	Complexity of the project increases if the quality control system management is not operated.

와 같이 데이터 모델의 보유 여부, 데이터 모델의 현행화 여부, 데이터 표준 존재 여부, 품질관리 시스템 운용 여부 등의 4가지 보정 요인을 파악하였다.

5. 소요공수 산정 모형 도출

5.1 초기모형 도출

본 연구에서는 2013년도에 시행된 데이터 품질 진단 및 개선사업을 초기 모형의 수립에 활용하여 전문가 델파이 분석의 기초자료로 활용하였다. 2013년에는 1차 및 2차에 걸쳐 총 15개의 DB 품질개선사업이 시행되었으며, 각 사업별로 작업 요소별 공수를 확보하여 모형의 분석에 사용했다. 다음의 <Table 3>~<Table 5>는 기존 수행 사업의 자료를 요약하고 있다.

<Table 3> Project Data in 2013

No.	First project	Second project
1	Ministry of Foreign Affairs	Gyeonggi Province
2	Ministry of Government Legislation	National Agricultural Products Quality Management Service
3	Defense Agency for Technology and Quality	Korea Agency of Education, Promotion and Information Service in Food, Agriculture, Forestry and Fisheries
4	Mistry of Public Safety and Security	Ministry of Food and Drug Safety
5	Anti-Corruption and Civil Rights Commission	Ministry of Security and Public Administration
6	Korea Employment Information Service	National Election Commission
7	Korea Environment Corporation	Korea Environmental Industry Technology Institute
8	Game Rating And Administration Committee	-

<Table 4> Project's Cost Factors Value in 2013

No.	Total tables	Tables for in-depth diagnosis	Organization scale	Related organization number
1	100	12	67	2
2	122	13	22	24
3	300	19	34	23
4	27	8	31	2
5	92	49	56	248
6	212	11	27	18
7	80	13	41	0
8	86	18	13	2
9	311	22	218	4
10	136	136	22	2
11	48	48	23	11
12	170	170	15	4
13	74	41	44	0
14	27	27	44	5
15	96	96	33	2

<Table 5> Project complexity adjusting factors in 2013

No.	Has data model (Y/N)	Data model updated (Y/N)	Level of data model	Data standardized (Y/N)	Has quality management system (Y/N)
1	N	N	3	N	N
2	Y	Y	1	N	N
3	N	N	3	N	N
4	Y	N	2	N	N
5	Y	N	2	N	N
6	Y	Y	1	Y	N
7	N	N	3	N	N
8	N	N	3	N	N
9	Y	N	2	N	N
10	Y	N	2	N	N
11	Y	N	2	N	N
12	Y	N	2	N	N
13	N	N	3	N	N
14	Y	N	2	N	N
15	Y	N	2	N	N

2013년 사업 자료의 공수를 분석하여 진단활동 및 개선활동의 각 작업요소별 기본 공수 산정식을 아래와 같은 선형 모형을 초기 모형으로 도출하였다. 즉, 하나의 작업요소 i 에 대하여 아래와 같은 수식을 사용한다.

$$Y_i = \sum_i (a_i + b_i X_i) \bar{c}$$

- Y_i : 작업요소 i 의 공수
- a_i, b_i : 작업요소 i 의 계수
- X_i : 작업요소 i 의 결정요인 규모
- \bar{c} : 난이도 보정계수 c_1, c_2, c_3 의 평균값 (초기 모형에서는 1로 설정)

기존 사업 자료에서는 도출된 보정요인의 영향이 정확히 파악이 되지 않은 상황이기 때문에, 동시에 각 계수를 추정하는 것이 어렵다. 따라서 초기 모형에서는 모든 난이도 보정계수의 평균값이 1이 된다고 가정한 상태에서, a, b 의 값을 우선적으로 추정한 후 전문가 델파이(Delphi) 설문을 통해 난이도 보정계수를 추정하였다.

초기 모형의 공수 산정식에서 작업요소 i 의 계수는 다음과 같은 절차에 따라 산정되었다. 우선 난이도 보정계수의 평균값이 1이 된다고 가정한 상태에서, 각 계수 값을 추정한다. 이때 작업요소별 고정적으로 발생시키는 a_i 와 결정요인 규모 X_i 에 따라 공수를 증가시키는 b_i 를 산정한다. 본 모형의 a_i 와 b_i 는 모두 양의 값을 가져야만 하지만, 수집된 자료의 산포가 매우 크기 때문에 최소제곱법(least square)을 단순하게 적용시키는 경우 양의 a_i, b_i 값의 추정에 어려움이 발생했다. 따라서 양의 a_i, b_i 값을 추정하기 위해, 기본 공수 산정식 $Y_i = a_i + b_i X_i$ 을 $(Y_i - a_i)/X_i = b_i$ 와 같이 변형시켜, X_i 에 대한 회귀식 기울기를 0으로 만드는 a_i 값을 추정치로 사용하였다.

구체적인 절차는 아래와 같다. 우선 임의의 고정된 값인 a_i 에 대하여 $(Y_i - a_i)/X_i = b_i$ 이기 때문에, $(Y_i - a_i)/X_i$ 는 X_i 에 대해 독립적으로 움직임을 이

용한다. 즉, 기존 사업 자료에 대해서 $(Y_i - a_i)/X_i$ 를 X_i 에 대한 회귀식으로 계산하면 기울기는 0이 되어야 한다. 따라서 기울기를 0으로 만드는 a_i 값을 계산하여 a_i 값의 추정치로 사용하였다.

또한 추정된 a_i 값을 $(Y_i - a_i)/X_i = b_i$ 에 대입하여, 기존 사업 자료에 대한 $(Y_i - a_i)/X_i$ 값을 산정한 후

각 사업별 $(Y_i - a_i)/X_i$ 의 평균값으로 b_i 를 추정하였다. 산정된 작업요소별 각 계수의 추정치는 <Table 6>에 요약했으며, 이를 델파이 설문조사를 위한 사전조사치의 도출에 사용하였다.

도출된 모형을 2013년도 수행된 사업 자료에 적용해 본 결과 <Figure 2>와 같이 분석 추정치가

<Table 6> Estimations Using Project's Data(1st, 2nd) in 2013

Phase	Task element	Constant (a_i)	Decision variable (X_i)	Slope of decision variable (b_i)
Diagnosis	DB operation and current state analysis	0.58		
	Value(profiling) diagnosis	0.45		0.409
	Link consistency diagnosis	0.50	Related	0.3688245
	Structure diagnosis	0.36	Total	0.3048635
	Business rule diagnosis	0.76	In-depth	4.718E-17
	Quality management level diagnosis	0.34	Organization	-1.48E-17
	Standardization level diagnosis	0.53	Total	0.1297998
	Performance diagnosis	0.78	In-depth	0.5238888
	Error cause analysis	0.55	Total	0.2140508
	Business impact analysis	0.61		
	Improvement opportunities	0.64		
Improvement	Defining quality improvement direction	0.42		
	Quality improvement planning	0.42		
	Standardization guideline	0.45		
	Word standardization	0.53	Total	-7.4E-17
	Domain standardization	0.45	Total	0
	Terminology standardization	0.47	Total	0
	Code standardization	0.42	Total	0
	Quality management guideline	0.35	Organization	0.3020361
	Quality management system	0.82	Organization	-1.41E-16
	Quality management training planing	0.35		
	Basic data cleansing	0.22	Total	0.4032247
	Basic improvements for application programs	0.94	Total	0
	Data modeling guide	0.37		
	Data model improvement	0.29	Total	0.126705
	Data model update	0.34	Total	0.1350266
	Performance tuning	0.52	In-depth	1.8615946
Link improvement guide	0.50			

- Note) • Total : Total tables.
 • In-depth : Tables for in-depth diagnosis.
 • Organization : Organization scale.
 • Related : Related organization number.

기존 수행사업 소요 공수를 비교적 근사적으로 반영하는 것으로 나타났다.



〈Figure 2〉 Comparison of Preliminary Analysis Model and Actual Project Data

분석 추정치와 기존 수행사업 소요 공수의 상관 계수는 약 0.6으로 매우 높은 상관관계를 보였으며, 평균 제곱근 편차는 3.10이었다.

5.2 1차 전문가 델파이(Delphi) 설문 분석

실제로 대가 모형의 분석 및 수립을 위해서는 대가 모형의 요소들이 가지는 모든 경우의 수에 대해 대상 사업에 투입되는 공수의 산정량을 파악하여야 한다. 하지만, 실제 사업 수행 자료가 많지 않은 경우는 산정량 분석에 필요한 자료를 충분히 확보할 수 없는 경우가 많다. 따라서 본 연구에서는 제한된 사업 데이터에 대한 통계적 분석과 병행하여 작업량에 대한 신뢰성 있는 추정치를 제공할 수 있는 전문가 그룹으로부터 2차에 걸쳐 전문가 델파이 분석을 실시하여 사전조사 모형을 검증 및 보완하였다. 1차 전문가 델파이(Delphi) 설문은 발주기관 15건, 수주기관 15건(총 30건)에 대해 이루어졌으며, 설문 문항은 <Figure 3>, <Figure 4>와 같이 작업요소별 대가 결정요인 및 규모별 공수 추정치 설문과 작업요소별 보정요인 및 공수에 영향을 미치는 정도 설문으로 구성되어 있다.

설문 응답 결과 결정요인이 1개인 작업요소의 경우 초기 모형과 동일하게 나타났으나, 작업요소에

따라 단일 결정요인이 아닌 복수 결정요인이 공수에 영향을 미치는 것으로 나타났다.

1) DB operation and current state analysis

Elements	Selected	Enter man-hour estimations for factor scale						
		Scale						
Factors	Number of sales	No						
		Estimations						
	Organization scale	No	Scale	less than 20	20 or more - less than 50	50 or more - less than 100	100 or more - less than 200	200 or more
		Estimations						
Related organization number	No	Scale	less than 4	4 or more - less than 20	20 or more - less than 50	50 or more - less than 100	100 or more	
	Estimations							
	No	Scale						
	Estimations							

〈Figure 3〉 1st Questionnaire Form to Estimate Cost Factor's Coefficients

1) DB operation and current state analysis

Elements	Selected	Effective?	Degree of effect on man-hours(% or MM)
Existence of data model	No effect	(Select if effective)	
Data model up-to-date	No effect	(Select if effective)	
Existence of data standard	No effect	(Select if effective)	
Operation of quality management system	No effect	(Select if effective)	
0	No effect	(Select if effective)	
0	No effect	(Select if effective)	
0	No effect	(Select if effective)	

〈Figure 4〉 1st Questionnaire Form to Estimate Complexity Adjusting Factor's Coefficients

작업요소별 보정요인의 설문 결과를 분석한 결과 작업요소별 기타 공통적으로 요구하는 추가적인 보정요인은 나타나지 않는 것으로 나타나, 초기 대가 산정모형에서 도출한 보정요인이 적합한 것으로 분석되었다.

5.3 수정모형 수립

1차 설문조사 결과를 반영하여 사전조사로 파악된 모형의 수정이 이루어졌다. 모형의 수정은 작업요소 및 보정요소에 대해 모두 이루어졌으며 구체적인 수정 사항은 다음과 같다.

5.3.1 복수개의 결정요소 반영

작업요소에 따라 복수의 결정요인이 있는 경우를 반영하여 모형의 정확도를 높였다. 공수 산정식은 작업요소별로 공수에 영향을 미치는 결정요

〈Table 7〉 Results of Cost Factor Survey for Diagnosis Phase

(Unit : MM)

Task elements (Factors)	Constant	Total tables	Tables for in-depth diagnosis	Organization scale	Related organization number
DB operation and current state analysis (Constant)	3	9	11	12	13
Value(profiling) diagnosis(Total tables)	1	13	10	1	7
Link consistency diagnosis (Related organization number)	0	2	10	3	23
Structure diagnosis(Total tables)	0	13	8	5	9
Business rule diagnosis (Tables for in-depth diagnosis)	1	2	19	11	11
Quality management level diagnosis (Organization scale)	4	4	1	18	8
Standardization level diagnosis(Total tables)	0	18	3	8	13
Performance diagnosis (Tables for in-depth diagnosis)	0	5	20	1	8
Error cause analysis(Total tables)	0	12	13	12	15
Business impact analysis(Constant)	2	5	8	17	17
Improvement opportunities(Constant)	8	5	6	12	11

Note) Bold numbers indicate selected cost factors for tasks.

〈Table 8〉 Results of Cost Factor Survey for Improvement Phase

(Unit : MM)

Task elements (Factors)	Constant	Total tables	Tables for in-depth diagnosis	Organization scale	Related organization number
Defining quality improvement direction (Constant)	5	6	11	17	15
Quality improvement planning(Constant)	7	4	6	18	14
Standardization guideline(Constant)	7	11	1	16	8
Word standardization(Total tables)	4	19	1	13	8
Domain standardization(Total tables)	4	19	2	6	6
Terminology standardization(Total tables)	5	18	2	13	8
Code standardization(Total tables)	2	19	2	12	12
Quality management guideline (Organization scale)	5	1	1	22	10
Quality management system (Organization scale)	2	0	3	24	10
Quality management training planing (Constant)	11	0	1	16	5
Basic data cleansing (Tables for in-depth diagnosis)	1	11	13	7	7
Basic improvements for application programs (Total tables)	8	12	3	5	6
Data modeling guide(Constant)	20	5	2	0	1
Data model improvement(Total tables)	2	14	8	3	6
Data model update(Total tables)	1	20	2	4	4
Performance tuning (Tables for in-depth diagnosis)	5	5	16	1	3
Link improvement guide(Constant)	9	1	3	4	18

Note) Bold numbers indicate selected cost factors for tasks.

인 수에 따라 다음과 같이 표현된다. 즉, 작업요소 i 에 대하여,

- 결정요인 수가 1개인 공정

$$Y_i = a_i + b_i X_i^{c_i}$$

- Y_i : i 작업요소의 공수(MM)
- a_i, b_i, c_i : i 작업요소의 계수
- X_i : i 작업요소의 결정요인 규모

- 결정요인 수가 2개인 공정

$$Y_i = a_i + b_i X_{i_1}^{c_1} X_{i_2}^{c_2}$$

- Y_i : i 작업요소의 공수(MM)
- a_i, b_i, c_i : i 작업요소의 계수
- c_j : i 작업요소에서 결정요인 j 의 계수
- X_{i_j} : i 작업요소의 결정요인 j 의 규모

위의 공수 산정식에 따라 추정된 작업요소별 계수 값은 <Table 9>와 같다.

5.3.2 5단계 난이도 보정계수 적용

<Table 2>의 네 가지 보정 요소 모두 해당이 되는 사업의 난이도가 해당되지 않는 사업보다 수월하므로 난이도는 쉬워진다. 전문가 의견 청취 결과 대다수 사업들이 우선 데이터 모델을 확립하고, 이후 데이터 표준을 수립하며, 마지막으로 품질관리 시스템을 운용하는 경향이 있는 것으로 확인되었다. 따라서 보정계수의 난이도는 네 요소가 모두 해당이 되지 않는 경우의 가장 난이도가 높은 경우부터 시작하여 위에서부터 하나의 요소씩 해당이 되어 난이도가 단계적으로 쉬워지는 5단계로 단순화하였다.

2013년 사업자료 분석 결과 15개 사업 중 대다수인 9개 사업이 데이터 모델만 존재하고 나머지 요소는 해당되지 않아 이 경우를 보정이 필요 없는 일반적인 경우로 적용하고, 난이도가 더 올라가고 내려감에 따라 사업의 난이도가 투입 공수에 미치

<Table 9> Results of Coefficient Estimations for Improved Model

Phase	Task element	Cost factors	a	b	c_1	c_2	
D I A G N O S I S	DB operation and current state analysis	① Organization ② -	0.376	0.087	0.438		
	Value(profiling) diagnosis	① Total ② -	0.474	0.002	1.000		
	Link consistency diagnosis	① Related ② In-depth	0.251	0.040	0.356	0.336	
	Structure diagnosis	① Total ② -	0.506	0.005	0.817		
	Business rule diagnosis	① Organization ② In-depth	0.051	0.022	0.416	0.328	
	Quality management level diagnosis	① Related ② Organization	0.071	0.017	0.318	0.358	
	Standardization level diagnosis	① Total ② -	0.570	0.001	0.971		
	Performance diagnosis	① In-depth ② -	0.320	0.111	0.483		
	Error cause analysis	① Related ② -	0.098	0.161	0.379		
	Business impact analysis	① Organization ② Related	0.113	0.035	0.408	0.289	
	Improvement opportunities	① Related ② Organization	0.176	0.016	0.389	0.527	
	I m p r o v e m e n t	Defining quality improvement direction	① Organization ② Related	0.031	0.045	0.305	0.349
		Quality improvement planning	① Related ② Organization	0.052	0.012	0.345	0.552
		Standardization guideline	① Organization ② Total	0.284	0.012	0.344	0.384
Word standardization		① Total ② Organization	0.298	0.008	0.455	0.383	
Domain standardization		① Total ② -	0.442	0.005	0.728		
Terminology standardization		① Total ② Organization	0.381	0.007	0.431	0.440	
Code standardization		① Total ② Organization	0.054	0.020	0.350	0.424	
Quality management guideline		① Organization ② -	0.410	0.007	0.821		
Quality management system		① Organization ② Related	0.260	0.041	0.376	0.279	
Quality management training planing		① Organization ② -	0.000	0.181	0.297		
Basic data cleansing		① In-depth ② Total	0.336	0.000	0.616	1.000	
Basic improvements for application programs		① Total ② -	0.615	0.014	0.603		
Data modeling guide		① Constant ② -					
Data model improvement		① Total tables ② -	0.483	0.003	0.872		
Data model update	① Total tables ② -	0.458	0.001	1.000			
Performance tuning	① In-depth ② -	0.216	0.078	0.656			
Link improvement guide	① Related ② -	0.070	0.280	0.288			

는 상대적인 크기를 설문조사를 통해 파악하였다. <Table 10>은 보정요인을 통해 구성한 난이도 수준을 보여준다.

<Table 10> Complexity Level of Adjusting Factors

Types of complexity adjusting factors	Complexity level
<ul style="list-style-type: none"> • Existence of data model (×) • Data model up-to-date (×) • Existence of data standard (×) • Operation of quality management system (×) 	Very difficult (No factor is applied)
<ul style="list-style-type: none"> • Existence of data model (○) • Data model up-to-date (×) • Existence of data standard (×) • Operation of quality management system (×) 	Difficult (Only possesses 'Data model')
<ul style="list-style-type: none"> • Existence of data model (○) • Data model up-to-date (○) • Existence of data standard (×) • Operation of quality management system (×) 	Normal (Possesses and updated 'Data model')
<ul style="list-style-type: none"> • Existence of data model (○) • Data model up-to-date (○) • Existence of data standard (○) • Operation of quality management system (×) 	Easy (All but 'Quality management system' are applied)
<ul style="list-style-type: none"> • Existence of data model (○) • Data model up-to-date (○) • Existence of data standard (○) • Operation of quality management system (○) 	Very easy (All the factors are applied)

Note) Complexity level 'Difficult' is applied as the base measure.

1차 설문 자료는 각 보정 요소별로 투입 공수에 대한 영향력을 파악하였다. 이를 바탕으로 각 보정 요소별 난이도 수준이 공수에 미치는 영향력을 <Table 11>과 같이 파악하였다.

기존 사업 자료를 대상으로 초기 모형과 수정 모형의 산정 공수를 비교했을 때, 1차 모형의 정확성이 높아지는 것으로 나타났다. 즉, 상관계수가 : 0.60(초기 모형)에서 0.63(수정 모형)으로 증가하는 것을 관찰하였다. 또한 평균 제곱근 편차는 9.61에서 10.02로 비슷하나 추정율은 115%에서 102%로 개선되었다.

5.4 2차 전문가 델파이(Delphi) 설문 분석

2차 전문가 델파이(Delphi) 설문 조사는 발주기관 7건, 수주기관 9건(총 16건)에 대해 실시하였으

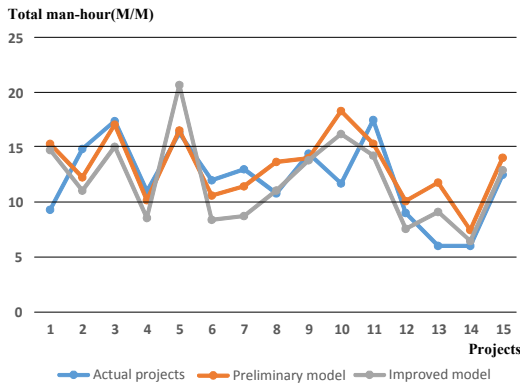
<Table 11> Results of 1st Survey for Complexity Adjusting Coefficients

Phase	Task element	Complexity level				
		Very difficult	Difficult	Normal	Easy	Very easy
D I A G N O S I S	DB operation and current state analysis	1.04	1	0.95	0.90	0.85
	Value(profiling) diagnosis	1.05	1	0.94	0.88	0.83
	Link consistency diagnosis	1.03	1	0.96	0.92	0.88
	Structure diagnosis	1.06	1	0.93	0.86	0.80
	Business rule diagnosis	1.06	1	0.94	0.88	0.82
	Quality management level diagnosis	1.03	1	0.96	0.92	0.88
	Standardization level diagnosis	1.05	1	0.94	0.89	0.83
	Performance diagnosis	1.04	1	0.96	0.92	0.88
	Error cause analysis	1.03	1	0.96	0.92	0.88
	Business impact analysis	1.04	1	0.95	0.91	0.87
I m p r o v e m e n t	Improvement opportunities	1.04	1	0.95	0.91	0.86
	Defining quality improvement direction	1.04	1	0.95	0.90	0.85
	Quality improvement planning	1.04	1	0.95	0.90	0.85
	Standardization guideline	1.06	1	0.93	0.87	0.81
	Word standardization	1.06	1	0.93	0.86	0.79
	Domain standardization	1.06	1	0.93	0.87	0.81
	Terminology standardization	1.06	1	0.93	0.86	0.79
	Code standardization	1.07	1	0.92	0.84	0.77
	Quality management guideline	1.05	1	0.94	0.89	0.84
	Quality management system	1.05	1	0.94	0.89	0.84
	Quality management training planning	1.04	1	0.95	0.90	0.85
	Basic data cleansing	1.04	1	0.95	0.90	0.86
	Basic improvements for application programs	1.05	1	0.94	0.88	0.82
	Data modeling guide	1.07	1	0.92	0.85	0.78
	Data model improvement	1.06	1	0.93	0.86	0.79
	Data model update	1.10	1	0.89	0.78	0.68
Performance tuning	1.07	1	0.92	0.84	0.77	
Link improvement guide	1.04	1	0.95	0.91	0.87	

며, 1차 설문을 통해 산정된 추정치에 대한 2차 전문가 설문 실시하여 1차 대가 산정모형을 검토 및 보완하였다. 2차 델파이(Delphi) 설문 문항은 다음과 같이 작업요소별 투입 작업인력 등급 추정치 설문, 작업요소별 대가 결정요인 및 규모별 공수 추정치, 그리고 결정요인 및 규모별 공수 추정

<Table 12> Estimation Performance Comparison of the Preliminary Model and the Improved Model

Measures	Preliminary model	Improved model
Correlation coefficient	0.60	0.63
Standard error	3.10	3.17
Average estimation performance (man-hour estimation/real input man-hour)	115%	102%



<Figure 5> Man-Hour Estimation Comparison of the Preliminary Model and Improved Model

치료 이루어졌다. 이 중 투입인력 등급 추정치는 기존 2013년 사업 자료의 작업요소별 투입 작업인력 등급의 평균 비율을 산정하여 설문 값으로 사용하였다.

1) DB operation and current state analysis

Items	Skill level				Total
	Expert	Advanced	Intermediate	Basic	
Estimations of the research team	32.3%	23.1%	29.1%	15.5%	100.0%
Enter the estimations					0.0%

<Figure 6> Questionnaire Form for Skill Level Estimations of Work Forces

수집된 델파이(Delphi) 2차 설문응답은 1차에 비해 응답자 수가 현저히 적어 통계적인 이상치 분석법(outlier detection)을 적용해 이상치를 제거(Iglewicz et al., 1993)한 후 수주자 및 발주자별로 평

1) DB operation and current state analysis

Prior estimations	Organization scale				
	10	35	75	150	300
	0.61	0.79	0.95	1.16	1.43

→

Enter estimations	Organization scale				
	10	35	75	150	300

(a) Task Elements with Sing Cost Factor

3) Link consistency diagnosis

Prior estimations	Related organization number				
	2	12	35	75	200
5	0.34	0.42	0.49	0.57	0.70
15	0.38	0.49	0.60	0.71	0.90
35	0.42	0.57	0.72	0.86	1.12
75	0.47	0.66	0.85	1.04	1.37
200	0.55	0.82	1.09	1.35	1.81

→

Enter estimations	Related organization number				
	2	12	35	75	200
5					
15					
35					
75					
200					

(b) Task Elements with Multiple Cost Factors

<Figure 7> Questionnaire Form for Man-Hour Estimations of Cost Factor Size Scales

1) DB operate and analysis of actual state

Project complexity	Types of complexity adjusting factors	Prior estimations	Enter estimations
Very difficult	Existence of data model (X) Data model up-to-date (X) Existence of data standard (X) Operation of quality management system (X)	1.05	
Difficult	Existence of data model (O) Data model up-to-date (X) Existence of data standard (X) Operation of quality management system (X)	1.00	1.00
Normal	Existence of data model (O) Data model up-to-date (O) Existence of data standard (X) Operation of quality management system (X)	0.95	
Easy	Existence of data model (O) Data model up-to-date (O) Existence of data standard (O) Operation of quality management system (X)	0.90	
Very easy	Existence of data model (O) Data model up-to-date (O) Existence of data standard (O) Operation of quality management system (O)	0.85	

<Figure 8> Questionnaire Form for Complexity Adjusting Factor Estimations

균치를 사용하여 대가 모형의 수립 및 검증에 사용하여 최종모형의 계수를 추정하였다. 이때 최종 모형의 계수는 수정모형과 동일한 공수 산정식을 사용하여 추정하였으며 계수 추정 결과는 <Table 13>과 같다.

최종모형의 보정계수 추정 결과, 수정모형의 계수에 상당히 근사하는 값을 얻었다. 작업요소별 투입인력 직급별 투입 비율 추정을 위해 2013년 사업자료의 데이터를 기초자료로 전문가 설문 응답 자료의 이상치 제거 후 평균값을 계산하였다. 직급별 투입 비율 전문가 설문 결과, 사업 자료로 계산한 제시 값과 상당히 근사하는 값을 나타내고 있다.

<Table 16>은 2013년 기존 사업 자료를 대상으로 초기모형, 수정모형, 최종모형의 산정 공수 예측력을 비교 분석한 결과이다. 최종 모형은 수정모형과

<Table 13> Results of the Final Model's Coefficients

Phase	Task element	Cost factors	a	b	c ₁	c ₂
D I A G N O S I S	DB operation and current state analysis	① Organization ② -	0.37548	0.08696	0.43779	
	Value(profiling) diagnosis	① Total ② -	0.47363	0.00216	0.99971	
	Link consistency diagnosis	① Related ② In-depth	0.25096	0.03995	0.35585	0.33549
	Structure diagnosis	① Total ② -	0.50587	0.00535	0.81692	
	Business rule diagnosis	① Organization ② In-depth		0.05099	0.28518	0.31528
	Quality management level diagnosis	① Related ② Organization	0.07079	0.0171	0.31809	0.35779
	Standardization level diagnosis	① Total ② -	0.56987	0.00095	0.97083	
	Performance diagnosis	① In-depth ② -	0.3198	0.1113	0.48324	
	Error cause analysis	① Related ② -	0.09752	0.16114	0.37858	
	Business impact analysis	① Organization ② Related	0.11291	0.03502	0.40805	0.28852
	Improvement opportunities	① Related ② Organization	0.17612	0.01604	0.38941	0.52674
	I m p r o v e m e n t	Defining quality improvement direction	① Organization ② Related	0.03139	0.04526	0.30455
Quality improvement planning		① Related ② Organization	0.05189	0.0118	0.55183	0.34509
Standardization guideline		① Organization ② Total	0.28413	0.01163	0.34358	0.38361
Word standardization		① Total ② Organization	0.29822	0.00795	0.38286	0.45507
Domain standardization		① Total ② -	0.44198	0.00511	0.72729	
Terminology standardization		① Total ② Organization	0.38139	0.0067	0.44018	0.43083
Code standardization		① Total ② Organization	0.05428	0.01999	0.42354	0.3501
Quality management guideline		① Organization ② -	0.40952	0.00667	0.82106	
Quality management system		① Organization ② Related	0.26046	0.04148	0.37645	0.27903
Quality management training planing		① Organization ② -	0.00001	0.18098	0.29719	
Basic data cleansing		① In-depth ② Total	0.2933	0.00022	0.61058	1.00576
Basic improvements for application programs		① Total ② -	0.6152	0.01433	0.60316	
Data modeling guide		① Constant ② -	0.394			
Data model improvement		① Total tables ② -	0.4828	0.00252	0.87211	
Data model update		① Total tables ② -	0.45769	0.00092	0.99978	
Performance tuning	① In-depth ② -	0.21577	0.07843	0.65573		
Link improvement guide	① Related ② -	0.07041	0.28021	0.28849		

<Table 14> Results of Final Model's worker Skill Level Ratios

Phase	Task elements	Expert	Advanced	Intermediate	Basic
D I A G N O S I S	DB operation and current state analysis	36.1%	27.1%	25.6%	11.2%
	Value(profiling) diagnosis	24.6%	25.4%	30.5%	19.5%
	Link consistency diagnosis	36.9%	31.4%	30.5%	1.2%
	Structure diagnosis	33.7%	31.6%	22.5%	12.1%
	Business rule diagnosis	34.4%	28.4%	23.7%	13.5%
	Quality management level diagnosis	38.9%	32.5%	17.2%	11.4%
	Standardization level diagnosis	30.3%	28.9%	25.3%	15.5%
	Performance diagnosis	34.6%	36.2%	27.6%	1.6%
	Error cause analysis	36.1%	28.2%	23.7%	12.0%
	Business impact analysis	34.0%	28.9%	22.7%	14.3%
	Improvement opportunities	36.2%	29.3%	21.2%	13.3%
	I m p r o v e m e n t	Defining quality improvement direction	37.6%	28.8%	23.6%
Quality improvement planning		34.9%	26.0%	23.5%	15.7%
Standardization guideline		30.1%	30.8%	24.7%	14.4%
Word standardization		23.0%	28.7%	27.8%	20.5%
Domain standardization		29.7%	28.9%	24.0%	17.4%
Terminology standardization		30.5%	28.4%	21.9%	19.2%
Code standardization		31.8%	29.8%	23.2%	15.2%
Quality management guideline		50.6%	34.0%	12.2%	3.3%
Quality management system		42.5%	30.2%	20.9%	6.4%
Quality management training planing		47.0%	39.1%	13.0%	0.8%
Basic data cleansing		26.4%	30.6%	29.8%	13.1%
Basic improvements for application programs		28.8%	29.1%	32.4%	9.7%
Data modeling guide		39.8%	34.6%	25.2%	0.4%
Data model improvement		39.0%	34.6%	26.0%	0.4%
Data model update		19.1%	27.3%	29.2%	24.4%
Performance tuning	34.6%	37.6%	26.2%	1.5%	
Link improvement guide	57.3%	36.0%	5.6%	1.1%	

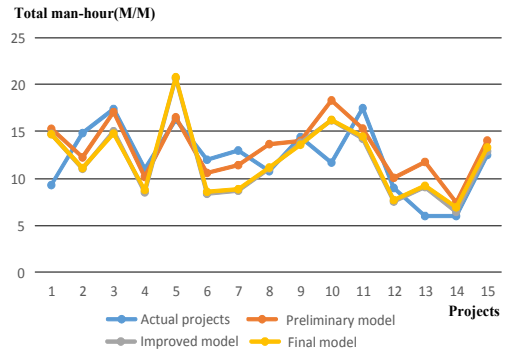
<Table 15> Complexity Adjusting Coefficients of Final Model

Phase	Task element	Complexity level				
		Very difficult	Difficult	Normal	Easy	Very easy
D I A G N O S I S	DB operation and current state analysis	1.06	1	0.93	0.86	0.81
	Value(profiling) diagnosis	1.08	1	0.93	0.83	0.77
	Link consistency diagnosis	1.07	1	0.93	0.88	0.82
	Structure diagnosis	1.09	1	0.92	0.86	0.79
	Business rule diagnosis	1.07	1	0.93	0.86	0.80
	Quality management level diagnosis	1.04	1	0.95	0.91	0.86
	Standardization level diagnosis	1.10	1	0.94	0.88	0.83
	Performance diagnosis	1.04	1	0.95	0.91	0.87
	Error cause analysis	1.05	1	0.95	0.90	0.86
	Business impact analysis	1.05	1	0.94	0.90	0.86
	Improvement opportunities	1.05	1	0.95	0.90	0.86
I m p r o v e m e n t	Defining quality improvement direction	1.05	1	0.95	0.90	0.85
	Quality improvement planning	1.04	1	0.95	0.90	0.85
	Standardization guideline	1.10	1	0.94	0.89	0.83
	Word standardization	1.06	1	0.93	0.87	0.80
	Domain standardization	1.08	1	0.93	0.88	0.82
	Terminology standardization	1.08	1	0.93	0.86	0.81
	Code standardization	1.08	1	0.92	0.83	0.77
	Quality management guideline	1.04	1	0.94	0.89	0.84
	Quality management system	1.05	1	0.94	0.89	0.84
	Quality management training planing	1.04	1	0.95	0.90	0.85
	Basic data cleansing	1.08	1	0.92	0.87	0.82
	Basic improvements for application programs	1.07	1	0.94	0.88	0.82
	Data modeling guide	1.07	1	0.92	0.86	0.80
	Data model improvement	1.15	1	0.91	0.85	0.79
	Data model update	1.19	1	0.88	0.80	0.73
Performance tuning	1.06	1	0.92	0.86	0.79	
Link improvement guide	1.04	1	0.95	0.90	0.86	

비슷한 예측력을 보이고 있으며, 평균 추정율은 103%로 전체적으로 총 소요공수를 적절히 예측하고 있는 것으로 나타났다.

<Table 16> Estimation Performance Comparison of All Models

Measure	Preliminary model	Improved model	Final model
Correlation coefficient	0.60	0.63	0.62
Root-mean-square deviation	3.10	3.17	3.14
Average estimation performance (man-hour estimation/real input man-hour)	115%	102%	103%



<Figure 9> Comparison of the Man-Hour Estimations of All Models

6. 대가 모형 검증

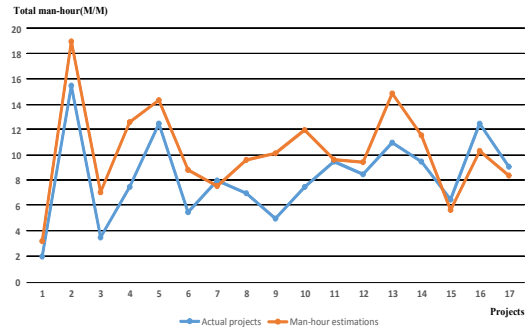
본 모형의 검증을 위해 2014년에 시행된 데이터 품질 진단 및 개선사업의 사업 자료 17건에 최종 모형을 적용해 비교 분석하였다. <Table 17>에 2014년에 시행된 사업의 자료를 요약하였다.

최종 모형으로 산정된 공수 추정치를 사업 자료와 비교 분석한 결과 2014년 사업 자료로 파악된 실제 투입공수와 근사한 값을 나타내고 있다. 사업의 실제 투입공수와 모형으로 추정된 투입공수 사이의 상관관계는 최종 모형의 수립에 사용된 2013년 사업 데이터상의 실제 투입 공수와 모형으로 추정된 공수

<Table 17> Project Data in 2014

Items	Organization name(DB name)
1	Korea National Arboretum
2	Ministry of Land, Infrastructure and Transport
3	Ministry of Agriculture, Food and Rural Affairs(Animal medicine and medical supplies)
4	Ministry of Agriculture, Food and Rural Affairs(Agricultural and stockbreeding products quarantine)
5	Ministry of Agriculture, Food and Rural Affairs(Imported beef)
6	Ministry of Agriculture, Food and Rural Affairs(Vegetation quarantine)
7	Korea Forest Service(Trail)
8	RURAL DEVELOPMENT ADMINISTRATION(Korea agricultural pests)
9	National Agricultural Products Quality Management Service(GAP and excellent food authentication information)
10	National Agricultural Products Quality Management Service(agricultural and stockbreeding products)
11	National Institute of Environmental Research(National source of pollution investigation system)
12	Korea Agro-Fisheries Trade Corporation (Agricultural and stockbreeding products price information)
13	Korea Water Resources Corporation (Real time water service information)
14	Korea Environment Corporation (National water service)
15	Korea Coast Guard(Power operated water leisure crafts registration)
16	National Emergency Management (Situation notice system)
17	National Institute of Animal Science (Korea standard feed composition)

사이의 상관계수인 0.62보다 오히려 더 높은 0.68로 나타났으며, 평균 제공근 편차도 3.29로 2013년 사업 데이터의 3.14와 비슷하여 본 모형이 사업의 투입 공수를 절절하게 추정하고 있음을 검증 할 수 있었다.



<Figure 10> Comparison of Man-Hour Estimations of the Final Model and Actual Project Data in 2014

7. 결 론

본 연구에서는 수·발주기관 전문가 및 이해관계자들에 대한 설문응답 및 의견수렴을 통해 합리적이고 공정한 DB 품질개선사업 대가 산정 모형을 수립하였다. 이를 위해 우선 DB 품질개선사업 대가 산정 모형 수립에 참고할 수 있는 유사 제도의 대가 현황 및 기존 DB 품질개선사업 대가 사례 분석을 통해, 대가 모형 수립에 필요한 시사점들을 도출하였다. 데이터 품질 및 진단사업은 기존의 관련 대가 기준이 다루고 있는 작업의 내용과 상이하고, 사업의 규모나 난이도로부터 받는 영향이 상이한 다양한 종류의 단위작업으로 구성되어 있어, 기존의 대가 모형을 직접적으로 적용하기가 어렵다. 본 연구에서는 수행범위, 결정요인 및 보정요인으로 구성되는 기존의 대가 모형의 틀을 참고하되 단위 활동별 결정요인, 보정요인 및 비용 산정식을 각 활동별로 정하는 새로운 형태의 모형을 통해 대가 산정의 정확도를 높이고자 하였다.

대가 산정을 위하여 본 연구에서는 DB 품질개선사업의 도입 단위와 작업요소 등을 고려하였으며, 객관적이고 체계적인 방법론을 적용하여 대가 모형의 신뢰성을 제고하였다. 기존 소요공수 산정 모형의 개선을 위해, 전문가 및 이해관계자 설문을 통해서 작업요소별 결정요인 및 난이도 보정요인을 도출하고, 델파이(Delphi) 기법 등의 활용을 통

해 적합하고 신뢰성 있는 대가 모형을 수립하였다. 대가 산정모형 수립 이후 DB 품질개선사업 대가 산정방식에 대한 기준 해설서와 대가 산정 예시 등을 통해 적용 용이성을 확보하였다.

본 연구를 통해 수립된 대가 모형은 분석의 기반이 된 2013년도 실시 사업 15개를 합리적으로 예측하고 있음을 확인하였다. 또한 최종 모형을 2014년도에 새로 실시된 18개 사업에 적용한 결과 실제 사업비와 매우 근접한 추정치를 도출하는 것으로 나타나 모형의 적절성을 검증할 수 있었다. 따라서 본 연구에서 제안한 대가 모형이 향후 공공 및 관련 분야의 DB 품질개선사업 운용에 대한 합리적이고 적절한 예산 수립을 가능하게 할 것으로 기대한다.

한편으로, 본 연구에서는 DB 품질개선사업의 사업 대가를 합리적으로 예측하는 모형을 도출하였으나, 여전히 적용상의 몇 가지 한계 및 제한점을 가지고 있다. 첫째로 본 연구의 모형에서는 보정요인으로서 데이터 모델, 데이터 표준 및 품질관리시스템 등 데이터 품질 관리 수준의 현황 측면을 주로 고려하고 있다. 그러나 실제의 사업비에 영향을 미치는 환경적 요인은 대상 데이터베이스의 기술적 복잡도나 대상 조직의 복잡도 등 다양한 요인이 있을 수 있다. 이후 충분한 데이터를 바탕으로 다양한 측면에서의 보정요인의 도출과 이의 사업비에 대한 영향에 대한 연구가 필요할 것으로 생각된다. 둘째로, 본 모형에서 고려하지 못한 요인이 사업비에 영향을 미칠 가능성을 배제할 수 없다. 이후의 적용 과정에서 본 모형으로 설명되지 않는 사업비의 변동 요인은 수발주 전문가와의 지속적인 협의에 의해 조정되어야 할 필요가 있을 것이며, 이러한 조정 요인들이 대가 모형에 반영될 수 있도록 해야 할 것이다. 마지막으로, DB 품질개선사업은 현재까지 초기 확산 단계로서, 본 연구의 모형 수립을 위한 사업 데이터의 확보가 용이하지 않았으며, 본 연구에서는 2013~2014년도에 수행한 한정된 데이터만을 대상으로 모형의 수립 및 검증을 진행하였다. 따라서 향후 연구에서는 2015년 및 이후의 축적된 사업 데이터를 활용하여 현행 모형의

재검증 및 산정식의 개선이 이루어져야 할 것으로 생각된다.

References

- IFPUG, Function Point Counting Practices Manual*, 2004, <http://www.ifpug.org>.
- Iglewicz, B. and D., Hoaglin, "Volume 16 : How to Detect and Handle Outliers", *The ASQC Basic References in Quality Control : Statistical Techniques*, Edward F. Mykytka, 1993.
- Jun, S.C. and J.H. Kim, "Case Studies of the Software Development Applying EVM(Earned Value Management)", *Korea Society of IT Services Journal*, Vol.7, No.3, 2008, 129-143. (전상철, 김자희, "소프트웨어개발 프로젝트에서의 획득가치관리(EVM) 적용 사례 연구", *한국IT서비스학회지*, 제7권, 제3호, 2008, 129-143.)
- Kang, U.S., K.H. Bae, and H.S Kim, "A Cost Optimization Model of IT Operation Service by Improving Service Request Management Process", *Korea Society of IT Services Journal*, Vol.6, No.3, 2007, 87-109. (강운식, 배경한, 김현수, "서비스 요청 관리 프로세스 개선을 통한 IT 운영비용 최적화 방안", *한국IT서비스학회지*, 제6권, 제3호, 2007, 87-109.)
- Korea Agency for Infrastructure Technology Advancement, "A study on the CM promotion plan and task procedure development considering CM types", 2003. (한국건설교통기술평가원, "CM 형태별 활성화 방안 및 업무절차서 개발 연구보고서", 2003.)
- Korea Software Industry Association, "SW project cost estimation guide", 2013. (한국정보통신산업진흥원, "SW사업 대가산정 가이드", 2013.)

- Linstone, H.A. and Turoff, M. (Editors), *The Delphi Method—Techniques and Applications*, 2002, <http://www.is.njit.edu/pubs/delphibook/>.
- Ministry of Land, Infrastructure, and Transport, “Estimating method of construction management cost”, Ministry of Land, Infrastructure, and Transport, 2009.
(국토해양부, “건설사업관리 대가기준”, 2009.)
- National Information Society Agency, “IS audit standard guide”, 2006.
(한국정보화진흥원, “정보시스템 감리기준 해설서”, 2006.)
- National Information Society Agency, “A estimated standard on the supervision cost for the information system”, 2008.
(한국정보화진흥원, “정보시스템 감리대가 산정기준”, 2008.)
- National Information Society Agency, “Public Information Quality Management Manual”, 2012.
(한국정보화진흥원, “공공정보 품질관리 매뉴얼”, 2012.)
- National Information Society Agency, “A Preliminary Study of Data Quality Diagnosis and Improvement Cost Estimation Method”, 2012.
(한국정보화진흥원, “품질진단 및 개선 소요예산 산정(안) 개발”, 2012.)
- National Information Society Agency, “PMO Application and Operation Guide for E-Government Projects”, 2013.
(한국정보화진흥원, “전자정부사업관리 위탁(PMO) 도입·운영 가이드”, 2013.)
- National Information Society Agency, “Database Construction Cost Estimation Guide”, 2013.
(한국정보화진흥원, “데이터베이스 구축비 대가기준 가이드”, 2013.)
- Park, C.K., W.J Kim, and Y.W. Seo, “Improving the Application Type and Quality/Characteristics Adjustment Factors of the Korea Software Cost Estimation Standard”, *Korea Society of IT Services Journal*, Vol.8, No.2, 2009. 43-69.
(박찬규, 김우제, 서용원, “소프트웨어 개발비 기준의 애플리케이션 유형과 품질 및 특성 보정요소 개선”, *한국IT서비스학회지*, 제8권, 제2호, 2009. 43-69.)
- Seo, Y.W. and D.H. Lee, “A Study on the Development of PMO Cost Estimation Model”, *Korea Society of IT Services Journal*, Vol. 13, No.4, 2014, 169-188.
(서용원, 이덕희, “PMO 대가산정모형의 개발”, *한국IT서비스학회지*, 제13권, 제4호, 2014, 169-188.)

◆ About the Authors ◆



Yong Won Seo (seoyw@cau.ac.kr)

Professor Yong Won Seo is currently a professor of MS and OM at College of Business and Economics, Chung-Ang University. He received his Ph.D. in Industrial Engineering from Seoul National University, and served in the National Information society Agency (NIA). He is interested in SCM, Service Operation, and Public Informatization.



Duck Hee Lee (leedh_81@outlook.com)

Duck Hee Lee is currently the doctoral course of MS and OM at Graduate School of Business, Chung-Ang University. He received his master's degree in Business management from Chung-Ang University. His research area includes SCM and Service Operation.



Seung Ho Jung (jsh@nia.or.kr)

Seung Ho JUNG is currently a senior researcher at National Information society Agency (NIA). He graduated from Korea University where he received BS in software engineering. He also received Ph.D. in business administration from Dong Guk university. His interest includes data quality management, big data, e-government, and cloud computing.



Kun Soo Park (kunsoo@business.kaist.ac.kr)

Kun Soo Park is an assistant professor of Operations Strategy and Management Science at KAIST College of Business in Seoul, Korea. He received BS and MS from Seoul National University and MS and Ph.D from Columbia University. His research interest includes game theoretic analysis of global supply chain and strategic management of information technology (IT).