

## 論文

J. of The Korean Society for Aeronautical and Space Sciences 46(1), 32-40(2018)

DOI:https://doi.org/10.5139/JKSAS.2018.46.1.32

ISSN 1225-1348(print), 2287-6871(online)

## 항공기 날개의 통계적 중량 예측식 도출 연구

김석범\*, 정한규\*, 황호연\*\*

## A Study on Deriving the Statistical Weight Estimation Formula for an Aircraft Wing

Seok-Beom Kim\*, Han-Gyu Jeong\* and Ho-Yon Hwang\*\*

Department of Aerospace Engineering, Graduate School, Sejong University\*

Department of Aerospace Engineering, Sejong University\*\*

## ABSTRACT

In this research, a method of deriving statistical weight prediction formula which is used during the conceptual design phase was studied and it was programmed using Microsoft Excel and verified by applying to jet transport aircraft. The database was built while referencing the variables of conventional wing weight estimation formulas and it was used for modeling the jet transport wing weight regression equation. The model was evaluated using the K-fold cross validation method to solve the overfitting problem of the model.

## 초 록

본 논문에서는 개념설계 단계에서 주로 사용되는 통계적 중량 예측식 도출 방법에 관한 연구를 수행하였으며 Microsoft Excel을 이용해 이를 프로그램화하고 제트 여객기에 적용하여 검증하였다. 기존 중량 예측식들의 변수들을 참고하여 데이터베이스를 구축하였고 이를 사용하여 제트 여객기 날개 중량 예측식을 모델링하였다. 모델의 과적합 문제를 해결하기 위해 K-fold cross validation 방법을 사용하여 모델을 평가하였다.

**Key Words** : Weight Estimation(중량 예측), Parametric Weight Estimation(모수 추정 중량 예측), Regression Analysis(회귀분석), K-fold Cross Validation(K겹 교차검증)

## 1. 서 론

항공기 설계 시 중량 예측은 이착륙거리, 순항거리, 연료소모, 실속속도, 상승률 등 항공기 성능뿐만 아닌 중량 데이터를 기반으로 하는 비용분석에도 직접적인 영향을 주는 중요한 요소

중 하나이다. 항공기 중량을 예측하는 방법은 설계 단계에 따라 정확한 CAD 데이터를 이용한 분석적 방법, 항공기 축소 모델을 이용한 실측 방법 등 여러 방법이 존재하지만 일반적으로 초기설계 단계의 항공기 중량 예측에는 통계적 방법 중 하나인 모수 추정 방법(parametric weight

† Received : August 12, 2017 Revised : October 12, 2017 Accepted : December 18, 2017

\*\* Corresponding author, E-mail : hyhwang@sejong.edu

estimation)이 사용된다. 모수를 이용한 중량 예측방법은 몇 가지 변수를 통해 항공기 형상 중량의 경향을 비교적 정확하고 빠르게 예측할 수 있다는 장점을 가지고 있다. 그러므로 세부형상이 정해져 있지 않고 형상 변화 주기가 짧은 초기 설계 단계에 많이 사용된다. 모수 추정을 이용한 중량예측은 현재까지 개발 운용되는 항공기의 자료들을 바탕으로 데이터베이스를 구축한 후 성능, 형상, 이륙중량 등 항공기 중량에 영향을 줄 수 있는 여러 가지 변수들에 관한 회귀식을 만들어 새로운 항공기의 중량을 예측하는 방법이다.

항공기 제작사들은 자체적으로 내부의 데이터베이스를 활용해 고유의 회귀식을 개발하여 활용하고 있으며 민간에서는 일반적으로 Raymer, Roskam 등의 항공기 설계 문헌에 기술되어 있는 회귀식을 이용한다[1,2]. 하지만 이러한 기존의 회귀식들의 한계는 기술진보로 인한 항공기의 변화하는 요구사항과 설계과정 중 필요한 상황에 알맞게 적용할 수 없다는 한계가 있다. 예를 들어 항공기의 중량은 사용되는 재료의 특성에 영향을 많이 받지만 기존의 회귀식들은 최신 복합재료의 특성을 반영하지 않아 결과의 오차가 커질 수 있으며 초기 설계 과정에서 알 수 없거나 혹은 정하지 않은 변수가 포함되어 있을 경우 기존의 회귀식을 통해 항공기의 중량을 예측하는 데 어려움이 있을 수 있다. 이러한 문제들을 해결하기 위해선 변화하는 데이터베이스를 기반으로 사용자가 사용할 수 있는 변수들로 구성된 새로운 회귀식이 도출되어야 좀 더 정확하게 항공기의 중량을 예측할 수 있다.

항공기 그룹별 중량 예측식을 도출하기 위한 데이터베이스에는 기존의 유사 항공기들의 그룹별 중량자료, 형상자료, 성능자료 등이 반드시 필요하다. 일반적으로 항공기의 형상자료와 성능자료는 각종 문헌 및 인터넷 매체에 상대적으로 공개된 것이 많지만 그룹별 중량자료는 항공기 제작사의 기밀 사항으로 민간에 공개된 자료가 극히 적다.

본 연구에서는 비교적 민간에 공개된 자료가 많은 제트 여객기 날개의 데이터베이스를 구축하고 이를 사용하여 새로운 회귀식을 만드는 과정을 연구하였다. 또한 공개된 자료의 부족으로 인해 데이터베이스의 양이 충분하지 않을 경우 발생할 수 있는 모형의 신뢰도 문제를 해결하기 위해 K-Fold Cross Validation 방법을 적용하고 주어진 자료를 최대한 활용하여 모형을 평가하는 방법으로 제트 여객기 날개 중량 예측식을 도출하였다.

## II. 회귀분석 모델링

### 2.1 다중 비선형 모델

항공기 구성 요소의 중량 회귀식의 모델은 일반적으로 식 (1)과 같은 지수함수 형태로 표현된다[3].

$$W_{wing} = \beta_0 X_1^{\beta_1} X_2^{\beta_2} X_3^{\beta_3} \cdots X_n^{\beta_n} \quad (1)$$

여기서  $\beta_i$ 는 회귀계수,  $X_i$ 는 독립변수,  $W_{wing}$ 은 종속변수인 항공기 중량을 의미한다.

회귀식을 결정하기 위해서는 최소 제곱법을 사용하여 회귀계수를 계산해야 한다. 최소 제곱법은 선형 형식의 회귀계수를 결정하는 방법으로 위와 같은 지수 형식 모델에 적용을 할 수 없다. 따라서 식 (1)에 로그를 취하여 선형형태로 만든 후 식 (2)와 같은 선형식에서 최소 제곱법을 사용해야 회귀계수를 계산할 수 있다[4].

$$\ln(W_{wing}) = \ln\beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \cdots + \beta_n \ln X_n \quad (2)$$

이를 j개의 데이터 집합에 대하여 행렬로 표시하면 식(3)과 같이 표현할 수 있다.

$$\begin{bmatrix} \ln W_{wing_1} \\ \ln W_{wing_2} \\ \vdots \\ \ln W_{wing_n} \end{bmatrix} = \begin{bmatrix} 1 & \ln X_{1_1} & \ln X_{2_1} & \cdots & \ln X_{n_1} \\ 1 & \ln X_{1_2} & \ln X_{2_2} & \cdots & \ln X_{n_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \ln X_{1_j} & \ln X_{2_j} & \cdots & \ln X_{n_j} \end{bmatrix} \begin{bmatrix} \ln \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \quad (3)$$

행렬식으로 표현된 식 (3)을  $y = X\beta$ 로 표시하면  $\beta$ 의 최소제곱추정량은  $(X'X)^{-1}$ 이 존재할 때 식 (4)와 같이 간단히 나타낼 수 있다[5].

$$\beta = \begin{bmatrix} \ln \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} = (X'X)^{-1} X' y \quad (4)$$

식 (4)에서 구한 회귀계수는 식 (2)에 대입되며 지수를 취함으로써 최종적으로 식 (1)과 같은 회귀식을 얻을 수 있게 된다. 다중 비선형 모델 회귀분석 흐름도는 Fig. 1과 같다.

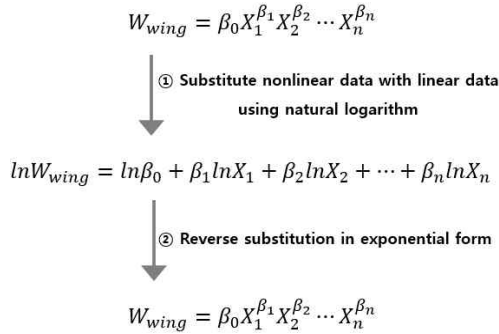


Fig. 1. Multiple nonlinear regression flow chart

2.1.1 초기 변수 설정

회귀모형의 종속변수에 영향을 미치는 다양한 변수들 중에서 어떠한 변수들이 독립변수로 모형 속에 포함되어야 하는지는 정확하게 알려진 바가 없다[5]. 일반적으로 회귀모형을 분석하는 초기 단계에서는 경험적인 판단으로 독립변수들을 설정하며 추후 검정통계 기준에 따라 최종 독립 변수들을 선택하게 된다. 이에 본 연구에서는 Raymer 통계식, Roskam 통계식, 초기 설계 변수 등을 고려하여 날개 면적(Wing area), 극한 하중계수(Ultimate load factor), 날개 길이(Wing span) 등 11가지 독립변수를 설정하였다. 또한 적절한 변수선택을 위해 모든 독립변수를 포함하는 모형에서 시작한 뒤 유의하지 않다고 판단되는 변수들을 하나씩 제거해 나가는 변수 감소법을 사용해 적합한 모형을 설정하였다[5]. 초기 설정된 독립변수 항목은 Table 1과 같다.

III. 데이터베이스 구축

3.1 주익 중량 및 성능 데이터

항공기 개발사들은 항공기의 전체적인 중량 및 형상 데이터만을 공개할 뿐 그룹별 세부 중량 데이터의 유출은 보안사항으로 엄격하게 통제한다. 따라서 민간에서 항공기의 그룹별 세부 데이터를 수집하는 일은 매우 어려운 일이다. 이에 본 연구에서는 Roskam, Sforza의 저서를 통해 21개의 기존 제트 여객기 주익 중량 데이터를 수집하였으며 기본적인 성능 데이터 및 형상데이터는 문헌 및 인터넷 매체를 통해 수집하였다[2,3,6]. 수집된 항공기의 목록 및 데이터는 Table 2에 요약 제시하였다.

Table 1. Variables and values

Aircraft	Wing Group Weight (lbs)	Wing area (ft <sup>2</sup> )	Takeoff weight (lbs)	...
DC-9-30	11400	1001	108000	...
DC-10-10	48990	3861	430000	...
DC-10-30	58859	3959	555000	...
MD-80	15560	1270	140000	...
737-200	10613	980	115500	...
727-100	17764	1700	160000	...
707-121	86402	5500	710000	...
A300-B2	44131	2799	302000	...
BAC 1-11	9643	1003	87000	...
L-1011	46233	3590	430000	...
⋮	⋮	⋮	⋮	⋮

Table 2. Independent variable list

	Independent variable
1	Wing area
2	Ultimate load factor
3	Wing span
4	Taper ratio
5	Takeoff weight
6	Zerofuel weight
7	Thickness ratio
8	1/4 Wing sweep back angle
9	Root chord length
10	Aspect ratio
11	Altitude

3.2 주익 형상 데이터

본 연구에서는 문헌 및 인터넷 매체에 공개되지 않은 형상 데이터는 항공기 삼면도의 기하학적 분석을 통하여 수집하였다. 즉, Windows의 MS Paint Tool을 통해 얻은 픽셀 값들과 실제 알려진 형상 데이터 값들의 비율을 활용하여 수집되었다. 본 논문에서는 데이터베이스에 수집된 항공기 중 Boeing사의 B-720 항공기를 예로 들어 설명한다[7]. 주익 형상 데이터 중 테이퍼비(Taper ratio), 1/4 후퇴각(Sweepback angle), 뿌리 코드 길이(Root chord length)를 얻기 위해 필요한 픽셀 값들의 위치와 픽셀 좌표는 Fig. 2, 3과 Table 3에 제시하였다.

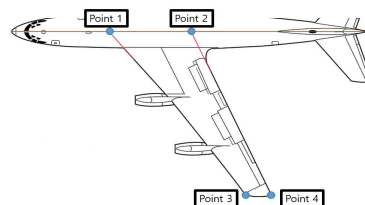


Fig. 2. Main wing pixel point (1)

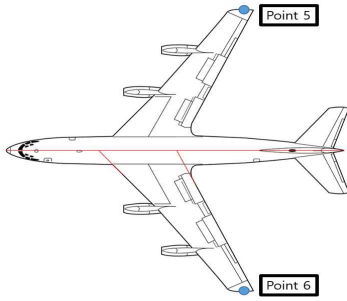


Fig. 3. Main wing pixel point (2)

Table 3. Pixel point values (Unit: Pixel)

	Horizontal	Vertical
Point 1	1717	1293
Point 2	2096	1293
Point 3	2342	2083
Point 4	2465	2083
Point 5	2405	490
Point 6	2405	2097

### 3.2.1 테이퍼 비(Taper ratio)

테이퍼 비는 끝단 코드(Tip chord) 길이와 뿌리 코드(Root chord) 길이의 비율이며 Point 1, 2의 가로 픽셀 차이와 Point 3, 4의 가로 픽셀 차이의 비율로 식 (5)와 같이 나타낼 수 있다.

$$\lambda = \frac{C_{tip}}{C_{root}} = \frac{P_{4H} - P_{3H}}{P_{2H} - P_{1H}} \quad (5)$$

여기에서  $C_{tip}$ ,  $C_{root}$ 는 각각 끝단 코드, 뿌리 코드의 길이이며  $P_{1H}$ ,  $P_{2H}$ ,  $P_{3H}$ ,  $P_{4H}$ 는 각각 Point 1, 2, 3, 4의 가로 픽셀 값이다.

### 3.2.2 뿌리 코드(Root chord) 길이

뿌리 코드의 비율과 실제 형상 수치를 알 수 있다면 삼면도 분석을 통해 다음과 같이 뿌리 코드의 길이를 구할 수 있다. 예를 들어 날개 길이(Span)의 실제 길이를 알고 있다면 식 (6)과 같이 구할 수 있다.

$$C_{root} = \frac{P_{2H} - P_{1H}}{P_{6H} - P_{5H}} \times b \quad (6)$$

여기서  $P_{5H}$ ,  $P_{6H}$ 는 각각 Point 5, 6의 가로 픽셀 값이며  $b$ 는 실제 날개 길이이다.

### 3.2.3 후퇴각(Sweepback angle)

후퇴각은 일반적으로 Fig. 4와 같이 뿌리 코드의 1/4지점과 끝단 코드의 1/4지점의 각도로 측정된다. 후퇴각을 구하기 위해서 먼저 1/4 뿌리 및 끝단 코드의 가로 픽셀 길이를 구해 Point 1, 3 가로 픽셀 값에 각각 더해서 1/4지점의 가로 픽셀 값을 알아내야 한다.

$$C_{root1/4} = (P_{2H} - P_{1H}) \times 0.25 \quad (7)$$

$$C_{tip1/4} = (P_{4H} - P_{3H}) \times 0.25 \quad (8)$$

여기서  $C_{root1/4}$ ,  $C_{tip1/4}$ 는 각각 뿌리 및 끝단 코드의 1/4 지점 가로 픽셀 길이이다. 따라서 1/4 뿌리 및 끝단 코드 점의 가로 픽셀 값은 식 (9), (10)과 같이 구할 수 있다.

$$P_{root1/4} = P_{1H} + C_{root1/4} \quad (9)$$

$$P_{tip1/4} = P_{3H} + C_{tip1/4} \quad (10)$$

1/4 뿌리 및 끝단 코드 점의 세로 픽셀 값은 각각 Point 1, 3의 세로 픽셀 값과 동일하다. 따라서 Fig. 4의 1/4 코드 후퇴각은 식 (11)과 같이 구할 수 있다.

$$\Lambda_{1/4} = \tan^{-1}\left(\frac{A}{B}\right) = \tan^{-1}\left(\frac{P_{tip1/4} - P_{root1/4}}{P_{3V} - P_{1V}}\right) \quad (11)$$

여기서  $P_{1V}$ 와  $P_{3V}$ 는 각각 Point 1, 3의 세로 픽셀 값이다. Fig. 5에 실제 사용한 삼면도 분석 데이터 일부를 나타내었고 Fig. 6에는 이를 종합한 최종 Database의 일부를 나타내었다.

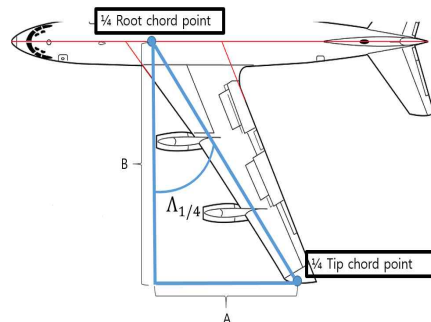


Fig. 4. 1/4 Chord sweepback angle

1	English unit & Pixel	Boeing			
2	Referenece(Roskam)	737-200	727-100	747-100	707-121
3	Point1 (Horizontal Position)	254	607	128	280
4	Point1 (Vertical Position)	181	211	81	484
5	Point2 (Horizontal Position)	180	523	90	189
6	Point2 (Vertical Position)	181	211	81	484
7	Point3 (Horizontal Position)	169	494	70	148
8	Point3 (Vertical Position)	35	56	14	317
9	Point4 (Horizontal Position)	148	468	59	121
10	Point4 (Vertical Position)	35	56	14	317
11	Span Length Pixel	345	363	154	371
12	Root Chord Length Pixel	74	84	38	91
13	Tip Chord Length Pixel	21	26	11	27
14	Root Chord Length x 0.25	18.5	21	9.5	22.75
15	Tip Chord Length x 0.25	5.25	6.5	2.75	6.75

Fig. 5. Analysis data of three side view

Aircraft	Wing Group Weight	Wing area	Wing span
DC-9-30	11400	1001	93.5
MD-80	15560	1270	107
DC-10-10	48990	3861	155
DC-10-30	58859	3958	165
DC-8	27556	2773	142.5
DC-9-10	9470	934	89.4
737-200	10613	980	93
727-100	17764	1700	108
747-100	86402	5500	195.8
707-121	24024	2433	130.1
707-320	29762	2892	142.5
707-320C	32255	3050	145.9

Fig. 6. Wing Database

## IV. 모델 평가

### 4.1 모델 평가 방법

회귀모형이 복잡해질수록 모델을 자세히 묘사할 수 있지만 정확하게 예측하기는 어려운 경향이 있다. 이러한 문제를 과적합(Overfitting)이라 한다. 회귀식은 실제 존재하는 모든 자료들 중 일부를 통해 모델링된다. 따라서 실제 자료의 모든 특성을 표현할 수 있는 모델을 만들 수 없으며 때로는 편향된 특성만을 표현하는 모델이 만들어질 수도 있다. 이러한 문제들을 해결하기 위해 데이터베이스 자료를 적절하게 구분 후 모델 평가를 수행해야 한다. 이러한 과적합 문제를 해결하기 위해 모델 평가방법인 Hold-out cross validation 방법과 K-fold cross validation 방법이 사용된다[8].

#### 4.1.1 Hold-out cross validation

Figure 7과 같이 모형 적합을 위해 사용되는 훈련자료(Training data) 50%, 모형 최적화를 위한 검증 자료(Validation data) 25%, 최종 모형 평가를 위한 평가 자료(Test data) 25%로 무작위로 분할하여 각각의 목적에 맞게 사용하는 방식이다. 일반적으로 Hold-out 방식은 충분한 자료를 바탕으로 모형 분석을 할 경우 사용된다.

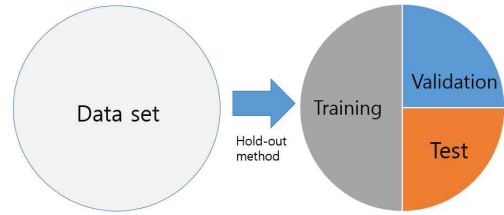


Fig. 7. Hold-out cross validation

#### 4.1.2 K-fold cross validation

주어진 자료의 수가 상대적으로 충분하지 못한 상태에서 Hold-out cross validation과 같은 방식의 자료의 분할은 편향된 특성을 가진 자료들로 분할될 가능성이 높아 과적합 문제를 해결하는데 큰 도움을 주지 못한다. 이러한 경우에 K-fold cross validation 방법을 사용해 주어진 자료를 최대한으로 활용하는 방법이 효율적이다. K-fold cross validation 순서는 Fig. 8에 나타냈으며 아래에 자세히 기술하였다.

- ① K개 그룹으로 데이터베이스 무작위 분할
- ②  $j=1, \dots, k$  일 때, 각  $j$ 에 대하여  $j$ 번째 그룹을 제외한 나머지  $k-1$ 개 그룹의 자료를 바탕으로 회귀식 도출
- ③  $j$ 번째 그룹을 이용하여 ②번에서 도출된 회귀식의 예측오차를 구함
- ④ 각  $j$ 에 대하여 ②, ③번 과정을 반복함
- ⑤  $k$ 개의 예측오차의 평균을 기준으로 최종 모형 평가

$k$ 의 개수는 정확히 정해진 값이 없으며 통상적으로 5 또는 10을 사용한다[8]. 본 연구에서는 민간에 공개된 제트 여객기 자료의 부족으로 인해 발생할 수 있는 모형의 신뢰도 문제를 해결하기 위해 K-fold cross validation 방법을 적용하여 연구를 진행하였다.

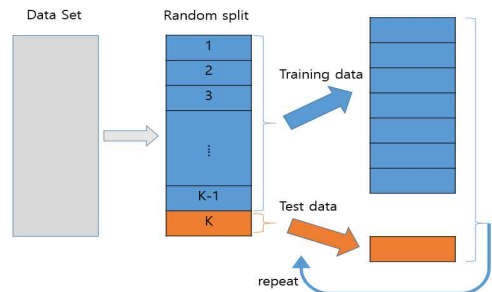


Fig. 8. K-fold cross validation method

## V. 연구 결과

### 5.1 프로그램 구축

본 연구를 위해 3절에 언급한 방법으로 데이터베이스를 확보하였으며 제트 여객기 날개 중량 예측식 도출을 위해 Microsoft Excel의 분석도구 및 VBA 기반으로 2, 4절에 제시된 K-fold cross validation 방법을 적용한 회귀분석 프로그램을 구축하였다.

### 5.2 독립변수 평가 및 최종 모형

본 연구에서 초기 변수는 Table 1과 같이 날개 면적, 극한 하중계수, 날개 길이 등 11가지 독립변수를 설정하여 진행하였다. 또한 회귀식 도출 후 해당 회귀계수의 유의성을 판단하기 위해 t 검정 결과 값(유의수준 5%)을 기준으로 변수 가감법을 적용해 유의한 독립변수를 선택하였다. 본 데이터베이스를 기반으로 하는 중량 예측식의 최종 변수들은 Table 4와 같다. 따라서 최종 회귀 모델은 식 (12)와 같이 설정하였다.

$$W_{wing} = \beta_0 S^{\beta_1} b^{\beta_2} (1 + \lambda)^{\beta_3} W_{to}^{\beta_4} (\cos A_{1/4})^{\beta_5} c^{\beta_6} \quad (12)$$

여기서  $\beta_i$ 는 회귀계수,  $S$ 는 날개 면적,  $b$ 는 날개 길이,  $\lambda$ 는 테이퍼 비,  $W_{to}$ 는 이륙 중량,  $A_{1/4}$ 는 날개 후퇴각,  $c$ 는 뿌리 코드 길이이다.

본 논문에서는 Table 1에서 언급된 변수들 내에서 변수 가감법을 사용해 최적화된 변수들을 찾아 최종 모형을 설정하였지만 이러한 변수들이 항상 날개 중량 회귀모형을 합리적으로 설명하는 것은 아니며 경우에 따라 변할 수 있다.

### 5.3 프로그램 구현

본 논문에서는 Fig. 9에 표현된 것과 같이 K값을 5로 설정하여 5-fold cross validation 방법을 적용한 결과를 제시하였다. 2.1절에서 언급했듯이 최소제곱법으로 회귀계수를 구하기 위해선 가장 먼저 로그스케일로 데이터베이스를 선형화해야 한

Table 4. Independent Variable List

	Independent variable
1	Wing area ( $S$ )
2	Wing span ( $b$ )
3	Taper ratio ( $1 + \lambda$ )
4	Takeoff weight ( $W_{to}$ )
5	1/4 Wing sweep back angle ( $\cos A_{1/4}$ )
6	Root chord length ( $c$ )

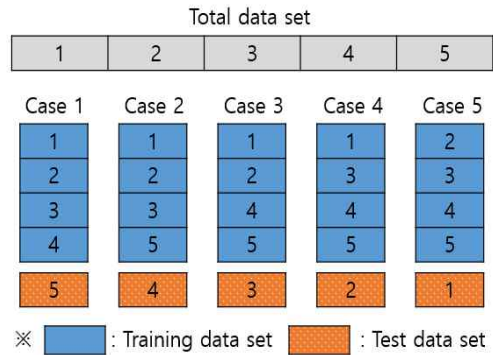


Fig. 9. 5-fold cross validation data set

다. 이를 바탕으로 K-fold cross validation 방법을 적용하여 최종 제트 여객기 날개 중량 예측식을 도출하였다.

#### 5.3.1 선형 데이터 변환

식 (13)과 같이 선형데이터를 얻기 위해 데이터베이스 전체에 자연로그를 취하여 Fig. 10과 같이 로그스케일로 변환하였다.

$$\ln W_{wing} = \ln \beta_0 + \beta_1 \ln S + \beta_2 \ln b + \beta_3 \ln (1 + \lambda) + \beta_4 \ln W_{to} + \beta_5 \ln (\cos A_{1/4}) + \beta_6 \ln c \quad (13)$$

#### 5.3.2 데이터 분배

로그스케일로 변환된 데이터베이스에 K-fold cross validation 방법을 적용하기 위해 본 연구에서는 Fig. 9와 같이 데이터를 분배하는 프로그램도 구현하였다. K값을 5로 설정하였기 때문에 데이터베이스는 무작위로 균등하게 분배된 5개의 그룹으로 자료가 나누어진다. Table 5에 데이터 분배를 반복(Iteration) 후 회귀분석을 진행하였을 때 결과 오차가 가장 작은 경우의 그룹을 나타내었다.

	A	B	E	F
1	ln(DATA)	Aircraft	Wing Group Weight	Wing area
2		DC-9-30	9.341	6.909
3		MD-80	9.652	7.147
4	Mc Donnell Douglas	DC-10-10	10.799	8.259
5		DC-10-30	10.983	8.283
6		DC-8	10.224	7.928
7		DC-9-10	9.156	6.839
8		737-200	9.270	6.888
9	Boeing	727-100	9.785	7.438
10		747-100	11.367	8.613
11		707-121	10.087	7.797
12		707-320	10.301	7.970
13		707-320C	10.381	8.023
14	Airbus	A300-B2	10.695	7.937
15		Fokker 614	8.660	6.535
16	VFW	Fokker F28-1000	8.900	6.712

Fig. 10. Database on logarithmic scale

Table 5. Regression coefficient

	Regression Coefficient (Case 1)	Regression Coefficient (Case 2)	Regression Coefficient (Case 3)	Regression Coefficient (Case 4)	Regression Coefficient (Case 5)
$\ln\beta_0$	-5.42876	-4.79914	-5.48602	-5.24912	-6.10137
$\beta_1$	0.490598	0.208876	0.088676	0.358575	0.12222
$\beta_2$	2.145862	2.655615	3.1364	2.767037	3.232721
$\beta_3$	1.718337	1.11765	2.052912	1.296984	1.579549
$\beta_4$	0.291348	0.103729	0.076039	0.092132	0.09354
$\beta_5$	0.507268	1.125103	-0.31912	0.571751	1.188173
$\beta_6$	-0.75818	-0.26593	-0.51917	-0.63398	-0.51312

### 5.3.3 회귀분석 결과

결과 및 분석은 Case 1~5 모두 동일한 방법으로 진행하였다. 따라서 Case 1을 예로 들어 결과 및 분석을 기술하기로 한다. Case 1 모형 적합을 위한 훈련자료(Training data)로는 Table 5의 1, 2, 3, 4번 데이터 그룹을 사용하였다. 훈련자료로 추정된 회귀계수는 Table 6에 5가지 경우에 대하여 모두 제시하였다. 여기서  $\ln\beta_0$ 는 Y절편의 회귀계수,  $\beta_1$ 은  $\ln S$ 의 회귀계수,  $\beta_2$ 는  $\ln b$ 의 회귀계수,  $\beta_3$ 는  $\ln(1+\lambda)$ 의 회귀계수,  $\beta_4$ 는  $\ln W_{to}$ 의 회귀계수,  $\beta_5$ 는  $\ln(\cos A_{1/4})$ 의 회귀계수,  $\beta_6$ 는  $\ln c$ 의 회귀계수이다.

Table 6의 회귀계수들은 선형 데이터베이스에 관한 계수들이며 식 (13)과 같은 형태로 표현된다. 하지만 최종 회귀식, 즉 날개 중량 예측식의 형태는 식 (12)와 같은 형태로 표현되어야 한다. 따라서 지수함수 형태로 역치환 해주어야 하며 Case 1의 최종 회귀식은 식 (14)와 같다.

$$W_{wing} = 4.39 \times 10^{-3} S^{0.49} b^{2.15} (1+\lambda)^{1.72} \times W_{to}^{0.29} (\cos A_{1/4})^{0.51} c^{-0.76} \quad (14)$$

회귀식의 적합성을 판단하는 기준인 F 통계량의 유의수준은  $5.8 \times 10^{-10}$ 으로 기준인 0.05 보다 작은 값을 나타낸다. 또한 추정된 회귀선이 표본 관측값들을 얼마나 잘 적합시키고 있는지를 나타낼 수 있는 척도로 쓰이는 결정계수( $R^2$ )의 값은 0.995로 최대값 1에 상당히 근접한 값을 갖는다[5]. Case 1의 최종 회귀식에 대한 예측 오차를 구하기 위해서 Table 5의 5번 데이터 그룹(747-100, MD-11, 737-200, DC-8)을 평가자료(Test data)로 사용하였다. 5번 데이터 그룹의 예측값은 각각 91,655 lbs, 60,058 lbs, 9,681 lbs,

Table 6. Total data set list

The number of data set	Aircraft	Wing weight
1	DC-9-30	11400
	747-200B	88741
	MD-83	15839
	727-100	17764
2	707-320C	32255
	DC-10-10	48990
	L-1011	46233
	DC-9-10	9470
3	A300-B2	44131
	Fokker F28-1000	7330
	707-121	24024
4	DC-10-30	58859
	737-320	29762
	MD-80	15560
	B720	23528
5	Fokker 614	5767
	747-100	86402
	MD-11	62985
	737-200	10613
	DC-8	27556

15,426 lbs로 계산되었다. 이에 대한 예측 상대오차는 각각 6%, 4%, 8%, 44%로 나타났으며, 최종 모형 선택 평가 기준인 오차평균은 약 16%로 나타났다. 이와 같은 방법으로 Fig. 9에 제시되어있는 Case 2~5에 대해 회귀분석을 실시하였으며 그 결과를 Table 7에 나타내었다.

### 5.3.4 최종모형 선택

4.1.2절에서 언급했듯이 최종모형 선택은 각각의 Case별 예측오차의 평균을 기준으로 최종모형

Table 7. Regression analysis

Case 1					
Regression equation	$W_{wing} = 4.39 \times 10^{-3} S^{0.49} b^{2.15} (1 + \lambda)^{1.72} W_{to}^{0.29} (\cos A_{1/4})^{0.51} c^{-0.76}$				
Test data	747-100	MD-11	737-200	DC-8	
Estimated wing weight (lbs)	91656	60059	9681	15426	
Relative error (%)	6.08	4.65	8.78	44.02	
Relative error average (%)	15.88	Significance F		5.8E-10	$R^2$ 0.9953
Case 2					
Regression equation	$W_{wing} = 8.24 \times 10^{-3} S^{0.2} b^{2.66} (1 + \lambda)^{1.12} W_{to}^{0.1} (\cos A_{1/4})^{1.13} c^{-0.27}$				
Test data	737-320	MD-80	B720	Fokker 614	
Estimated wing weight (lbs)	34350	16542	25487	5765	
Relative error (%)	15.42	6.31	8.33	0.04	
Relative error average (%)	7.52	Significance F		7.9E-10	$R^2$ 0.9949
Case 3					
Regression equation	$W_{wing} = 4.14 \times 10^{-3} S^{0.09} b^{3.14} (1 + \lambda)^{2.05} W_{to}^{0.08} (\cos A_{1/4})^{-0.32} c^{-0.52}$				
Test data	A300-B2	Fokker F28-1000	707-121	DC-10-30	
Estimated wing weight (lbs)	40539.48025	6703.070842	27352.6	52798	
Relative error (%)	8.14	8.55	13.86	10.30	
Relative error average (%)	10.21	Significance F		2.4E-10	$R^2$ 0.9961
Case 4					
Regression equation	$W_{wing} = 5.25 \times 10^{-3} S^{0.36} b^{2.77} (1 + \lambda)^{1.3} W_{to}^{0.09} (\cos A_{1/4})^{0.57} c^{-0.63}$				
Test data	707-320C	DC-10-10	L-1011	DC-9-10	
Estimated wing weight (lbs)	34236.90506	50610.19358	44082.2	9887.269	
Relative error (%)	6.14	3.31	4.65	4.41	
Relative error average (%)	4.63	Significance F		1.4E-09	$R^2$ 0.9942
Case 5					
Regression equation	$W_{wing} = 2.24 \times 10^{-3} S^{0.12} b^{3.23} (1 + \lambda)^{1.58} W_{to}^{0.09} (\cos A_{1/4})^{1.19} c^{-0.51}$				
Test data	DC-9-30	747-200B	MD-83	727-100	
Estimated wing weight (lbs)	12407.13297	84032.71257	16379.13	15330.04	
Relative error (%)	8.83	5.31	3.41	13.70	
Relative error average (%)	7.81	Significance F		5.5E-10	$R^2$ 0.9953

을 판단하였다. 예측오차의 평균이 가장 작은 회귀식 모형은 Case 4의 4.63%이며 평가자료 각각의 상대오차는 707-320C가 6.14%, DC-10-10이 3.31%, L-1011이 4.65, DC-9-10이 4.41%로 상대오차의 평균뿐만 아닌 각각의 상대오차도 적다는 것을 확인할 수 있다. 또한 F의 유의수준 값은  $1.4 \times 10^{-9}$ , 결정계수( $R^2$ ) 값은 0.9942로 회귀식과 데이터의 적합성을 확인할 수 있다.

### 5.4 최종 회귀식 비교 검증

최종모형의 결과 검증을 위해 문헌에 소개된 Raymer, Kroo, Torenbeek의 날개 중량 예측식과 비교 분석 하였다[1,3,9]. 비교 검증에 사용된 평가 자료는 Case 4에서 사용된 707-320C, DC-10-10, L-1011, DC-9-10이며 결과를 Table 8에 나타내었다. Raymer 예측식과의 상대오차는 각각 12.11%, 8.38%, 1.25%, 26.87%, Kroo 예측식과의 상대오차



Table 8. Relative error of estimated wing weight

Aircraft	707-320C	DC-10-10	L-1011	DC-9-10
Raymer's error (%)	12.11	8.38	1.25	26.8
Kroo's error (%)	33.88	8.62	5.99	8.60
Torenbeek's error (%)	20.92	19.88	16.63	54.15
Case 4 error (%)	6.14	3.31	4.65	4.41

는 각각 33.88%, 8.62%, 5.99%, 8.6%, Torenbeek 예측식과의 상대 오차는 각각 20.92%, 19.88%, 16.63%, 54.15%로 나타났다. Case 4에서 구해진 예측식은 문헌에 소개된 예측식과 비교해 전반적으로 우수한 예측결과를 나타내었다.

## VI. 결 론

본 논문에서는 제트 여객기 주익 중량 예측식 도출 방법과 이를 사용하여 도출된 예측식의 검증 과정에 대해 연구하였다. 이를 위해 우선 제트 여객기 주익 중량 예측모델에 대한 회귀식 도출 방법을 제시하였다. 이후 회귀식 도출을 위한 데이터베이스 구축과정을 기술하였다. 마지막으로 효과적인 회귀모델 평가를 위해 Hold-out cross validation, K-fold cross validation 방법을 제시하였으며 이 중에서 K-fold cross validation을 사용해 Case별 회귀식을 제시하였고 회귀식 모델에 사용되지 않은 새로운 자료를 평가자료로 사용하여 회귀식을 검증함으로써 회귀식의 적합성을 확인하였다. 본 연구에서는 제트 여객기의 날개 중

량 예측식만을 다루었지만 초기 설계 단계에서 이와 같은 방법을 활용해 전투기를 비롯한 여러 종류의 항공기 제원별 중량 예측식을 도출하면 초기 설계단계의 전체적인 중량 흐름 파악에 용이하게 활용될 수 있을 것으로 기대된다.

## 후 기

이 논문은 국토교통부의 재원으로 국토교통과학기술진흥원의 지원을 받아 수행된 연구사업임 (과제번호 : 16CTAP-C114866-01).

## References

- 1) Raymer, D. P., *Aircraft Design: A Conceptual Approach*, AIAA Education Series, 1999.
- 2) Roskam, J., *Airplane Design Part V: Component Weight Estimation*, DARcorporation, 1999.
- 3) Torenbeek, E., *Synthesis of Subsonic Airplane Design*, Delft University Press, 1982.
- 4) Wiegand, B. P., "The Basic Algorithms of Mass Properties Analysis and Control," *Society of Allied Weight Engineers*, No. 2067, 1992.
- 5) Lee, W. R., *Regression Analysis*, 2nd Ed., Tamjin, 2012.
- 6) Sforza, P. M., *Commercial Airplane Design Principles*, Butterworth-Heinemann, 2014.
- 7) <https://www.dfr.nasa.gov/Gallery/Graphics/B-720/>
- 8) Hastie, T., Tibshirani, R., Friedman, J., *The Elements of Statistical Learning Data Mining, Inference, and Prediction*, Springer, 2013.
- 9) Kroo, I., *Aircraft design: Synthesis and Analysis*, Desktop Aeronautics, 2001.