

Original Article

<https://doi.org/10.12985/ksaa.2025.33.2.141>
ISSN 1225-9705(print) ISSN 2466-1791(online)

항공기의 연료소모량에 영향을 미치는 요소에 관한 연구: 랜덤포레스트 기법을 이용하여

장효석*

Machine Learning Approach to Uncover Determinants of Aircraft Fuel Consumption using Random Forest

Hyoseok Chang*

ABSTRACT

This study systematically analyzed the key factors affecting aircraft fuel consumption using a random forest regression model based on 1,034 operational samples and 8 predictive variables. The model was validated through 10-fold cross-validation, and hyper parameter tuning identified $mtry = 5$ as the optimal setting. Variable importance analysis revealed that flight distance(41.9%) and payload(24.8%) were the most influential factors, followed by temperature, speed, wind, altitude, speed, and airplane performance monitoring system. The model achieved an RMSE of 501.52, an R^2 of 0.601, and an MAE of 358.09 on the test set, indicating stable performance and demonstrating that distance and payload are the primary determinants of fuel consumption in commercial aviation. This framework supports ICAO's roadmap by transforming raw operational data into actionable insights, empowering airlines to minimize costs and emissions concurrently.

Key Words : Airline(항공사), Flight Plan(비행계획), Fuel Consumption(연료소모량), Fuel Management(연료관리), Random Forest(랜덤포레스트)

1. 서 론

2016년 국제민간항공기구(ICAO)가 국제 항공 운송 부문에서 탄소 배출 감축을 위한 글로벌 시장 기반 조치를 도입한 이후, 항공사는 이 과정에 참여해 왔다. 연료비는 항공사 운영비 중 가장 큰 비중을 차지하며,

2024년에는 전 세계 항공업계 총비용의 약 30%에 이를 것으로 추정된다. 국제항공운송협회(IATA)와 ICAO는 각각 2021년과 2022년에 2050년까지 항공 운송 부문의 이산화탄소 순 배출량을 '제로'로 만들겠다는 목표를 설정하고, 민간과 공공 부문 전반에 걸쳐 이를 추진하고 있다. 항공업계의 핵심 과제는 기존 화석연료를 지속가능항공유(sustainable aviation fuel, SAF) 등 재생할 수 있는 에너지로 전환하거나 연료 효율이 높은 항공기를 도입하고, 연료 관리 전략을 고도화하여 연료 사용의 효율성을 극대화하는 데 있다. IATA는 SAF가 2050년 항공업계 탄소 감축 목표의 약 65%를 달성하는 데 핵심적인 역할을 할 것으로 전망한다.

Received: 29. May. 2025, Revised: 20. Jun. 2025,
Accepted: 27. Jun. 2025

* 한서대학교 항공교통물류학과 교수

연락처 E-mail : daniel.chang@hanseo.ac.kr

연락처 주소 : 충남 태안군 남면 고평로 236-49, 한서대
학교 항공교통물류학과 315호

Institute for Energy Economics and Financial Analysis(IEEFA, 2024)는 현재 SAF는 전체 항공유 소비량의 약 0.1%만을 차지하고 있으며, 그 도입이 늦어지는 가장 큰 이유는 기존 항공유보다 2~5배에 달하는 높은 가격 때문으로 설명한다. Centre for aviation(CAPA)은 코로나19 팬데믹으로 위축되었던 항공업계의 승객 탑승률이 2024년 기준 83.5%로 회복한 것으로 파악한다. 이에 따라 IATA에 의하면 항공사들은 미래 수요 증가에 선제적으로 대응하기 위해 신규 항공기 도입을 추진하고 있으며, 2024년에는 1,254대, 2025년 1,802대의 신규 항공기 도입이 예상된다. SAF 도입이나 신형 항공기 구매와 같은 대규모 투자는 경영진의 전략적 판단이 필요하다. 그러나, 조종사와 운항관리사 등 현장 항공 종사자들이 연료 소모를 분석하고 최적의 연료량을 산정하는 실질적인 연료 관리 과제를 적극 실천하는 행동 변화, 즉 ‘넛지(nudge)’를 통해서도 연료 효율 개선 효과를 얻을 수 있다. 항공사에서 운항 통제, 비행 감시 업무를 담당하는 운항관리사는 항공안전법과 운항기술기준에 따라 규정된 연료량을 탑재해 비행계획을 수립한다. 그러나 Boston Consulting Group에 따르면 조종사와 운항관리사는 기상 악화나 공항 혼잡 등으로 인한 목적지 공항 인근에서 지연을 고려하여, 규정된 연료 외에 추가 연료를 재량으로 탑재하는 경우가 많다고 분석하고 있다. 항공기의 연료소모량 분석은 복잡하며 항공기 중량, 기종, 엔진 상태, 유상하중, 비행 거리, 순항 고도, 속도, 기상 조건 등 다양한 요인의 영향을 받는다. 기존의 전통적인 연료 예측 모델이나 비행계획 시스템(flight plan system, FPS)은 모델의 프레임워크와 가정에 제약을 받거나, 비행계획 관련 요소만을 고려하여 실제 연료소모량 예측에 정확성이 떨어지는 한계가 있다(Wu et al., 2022). 빅데이터 기술의 발전으로 데이터 기반 예측 방법이 주목받고 있으며, 항공 산업에서도 기계학습의 적용이 활발히 이루어지고 있다. 복잡한 비선형 관계를 효과적으로 모형화하려면 다양한 요인을 종합적으로 고려하는 새로운 접근 방식이 필요하다. 따라서 본 논문은 항공기 연료소모량에 영향을 미치는 주요 변수들이 실제 운항에서 연료 사용량에 얼마나 영향을 미치는지 체계적으로 분석하고, 이를 통하여 데이터 기반의 과학적 연료 탑재량 기준을 제시하고자 한다.

II. 연료소모량에 미치는 주요 요인

비행경로 최적화와 AI 기반 궤적 관리 기술은 이륙·상승·순항·하강·착륙 등 각 비행 운영 단계에서 연료 효율을 극대화하여 소비량을 감소시키는 핵심 기술로 입증되었다. 또한 기상 조건(바람·온도)과 같은 환경적 요인과 고효율 엔진 기술 발전도 연료소모량에 직접적인 영향을 미친다. Table 1은 항공기 연료소모량에 미치는 주요 요인을 운영, 환경, 기술 측면에서 문헌 연구를 통해 기술하였다.

2.1 운영 측면

비행시간과 연료 소모는 밀접하게 연관되어 있어, 비행경로와 하강 프로파일을 최적화하면 상당한 연료 절감 효과를 기대할 수 있다. 예를 들어, 동적 프로그래밍을 활용한 궤적 최적화 기법을 적용하면, 최적의 하강 속도와 비행 속도를 선택하여 연료 소비와 비행시간을 동시에 줄일 수 있는 것이 확인되었다(Harada, A et al., 2014). 항공기의 운항 거리는 비행시간이 결정되는 물리적인 기반이며, 거리가 길어질수록 비행시간과 총 연료소모량은 증가하는 경향이 있다. 하지만, 이 관계는 비선형적이며, 특히 이륙 및 상승 단계에서의 높은 연료소모와 장거리 비행 시 추가 연료탑재로 인한 중량 증가의 영향으로 복잡성이 나타난다(Miyosh et al., 2009). 실제로 이륙 단계에서는 순항 단계보다 분당 연료 소비가 2.6~3배 많은 연료 소모가 발생한다(Singh et al., 2019). 그리고, 이륙중량을 구성하는 요소 중 유상하중은 연료소모량과 밀접한 관계가 있다. FAA의 AEDT (2017) 모델은 이륙중량 10% 변화는 연료소모량에 약

Table 1. Key factors influencing flight extra fuel recommendations

Factor	Description	Citation
Operational factors	Flight time, route optimization, holding trajectories, payload and aircraft operations	Harada, A et al., 2014.
Environmental factors	Weather patterns, noise reduction, and population distribution	Kim & Liem, 2022.
Technological factors	Engine design, alternative fuels, AI, and predictive maintenance	Mgbachi, 2024, Singh et al., 2019, Smirti & Hansen, 2009.

5%의 영향을 줄 수 있다고 하였다. IATA는 단위 중량 변화 비용 요소(cost of weight factor)를 통해 중량 변화가 추가 연료 소모에 미치는 영향을 측정하며, 기종, 운항 거리, 고도, 속도에 따라 2.5~5.0% 정도의 추가 연료 소모 분석 결과를 제시했다.

2.2 환경 측면

바람, 기온 등의 기상 조건은 연료소모량에 영향을 미칠 수 있다. 예를 들어, 정풍은 연료 연소를 증가시키고, 배풍은 연료 연소를 감소시킬 수 있다. 그러나 현재 모델은 실시간 기상 데이터를 고려하지 않는 경우가 많으므로 정확도가 제한될 수 있다(Kim & Liem, 2022). 그리고, Ren and Leslie(2019)는 대기 온도가 상승하면 항공기 엔진의 열효율이 감소하고, 공기 점성이 증가해 항력이 커지며, 결과적으로 연료 효율이 저하된다고 분석하였다.

2.3 기술 측면

K. Seymour et al.(2020)은 항공기 성능을 상세하게 반영한 정밀 비행 프로파일 시뮬레이터와, 출발지·목적지 공항과 항공기 기종만을 입력 변수로 하는 간소화된 연료 소비 모델을 개발하였다. 이를 통해 다양한 운항 조건과 항공기 특성에 따라 연료 소비를 정밀하게 예측하고, 효율적인 운항 전략 수립이 가능함을 보여주었다. 엔진 기술의 발전은 연료 소비를 크게 줄일 수 있다. 이러한 기술의 발전은 연비를 개선하고 환경에 미치는 영향을 줄이는 데 매우 중요하다(Singh et al., 2019). 그리고, 인공지능(AI)은 예측 유지보수 및 경로 최적화를 통해 연료 소비를 최적화하는 데 중요한 역할을 한다. AI 기반 시스템은 다양한 사례에서 연료 소비를 9~14%까지 절감하는 것으로 나타났다(Mgbachi, 2024).

III. 연구 방법

3.1 연구 설계 및 데이터 수집

본 연구는 항공기 운항 데이터를 활용하여 연료소모량을 예측하는 회귀 모델을 개발하는 데 있다. 2017년 한 해 동안 국내 대형 항공사의 연료 관리시스템을 통해 수집된 실제 운항 데이터와 연료소모량 데이터를

활용하였다. 이러한 정보를 기반으로, 항공기 연료소모량 예측을 위한 기계학습 모델을 구축하였다. 분석 대상 기종으로는 Airbus 330-200을 선정하였으며, 이 기종은 단거리부터 장거리까지 운항 기록이 축적되어 있어 예측 모델의 신뢰성과 일반화 가능성을 높일 수 있다. 데이터는 총 1,034개의 관측치와 8개의 변수로 구성되어 있다.

3.2 변수 정의 및 전처리

각 변수의 정의는 다음과 같다. Altitude는 항공기의 순항 고도를, Speed는 순항속도를 의미한다. Distance는 비행계획 시 출발 공항에서 목적지 공항까지 거리로 정의하며, 이때 거리는 바람의 영향을 반영하지 않은 지상거리(ground distance)를 사용하였다. 이는 바람의 영향을 다른 변수로 사용하였기 때문이다. Payload는 유상하중으로 승객, 화물, 수하물의 무게로 정의하고, WB는 항공기의 무게중심(center of gravity, CG)으로 평균공력시위(mean aerodynamic chord, MAC) 대비 %로 표시하여 나타냈다. 연료소모량(fuelburn_hr)은 각 항공편의 실제 연료소모량을 비행 시간으로 나누어 시간당 소모된 연료량으로 나타낸 연속형 수치 데이터로, 본 연구의 목표 변수(target variable)이다. Wind는 비행 중 측정된 평균 풍속을 나타내며, 정풍과 배풍을 반영한 정수형 변수이다. Temperature는 비행 고도에서의 외부 평균 기온을 표시하였다. APMS(Airplane Performance Monitoring System)는 항공기 운용 시 실제 연비가 제작사에서 제공하는 성능 지침과 다르게 나타나는 경우가 많이 발생하는데, 차이점을 분석해서 비행계획에 반영하는 프로그램으로 성능관리 프로그램이라고 한다(유재림, 2023). 항공기 엔진 운영에 따라 100% 성능을 기준으로 추가로 연료가 소모되는 양을 고려한 지표로, 백분율로 나타낸다. 데이터 전처리 과정에서 결측치가 없어 모든 데이터가 그대로 유지되었고, 전체 데이터를 훈련 세트와 테스트 세트 8:2의 비율로 나누어, 훈련 세트는 모델 학습에, 테스트 세트는 모델 성능 평가에 사용하였다.

3.3 랜덤포레스트 모델 구축

본 연구는 랜덤포레스트 회귀(random forest regression) 알고리즘을 사용하였고, 모델의 성능은 RMSE(root mean squared error), R-squared(결정 계수),

MAE(mean absolute error) 등 대표적인 회귀 평가 지표를 통해 검증하였다. 랜덤포레스트(RF) 알고리즘은 정확성과 계산 효율성 사이에서 균형을 유지하면서 비행 추가 연료 소요량을 예측하는 도구로 연구에 이용하고 있다. Wang et al.(2020)은 비행계획, 항공기 운항 및 성능 데이터를 분석하여, 과거 운항 데이터를 학습해 비행 연료 소비량을 예측하는 랜덤포레스트 모델을 구축했다. 안전에 대한 여유분 조정과 실험적 검증을 결합함으로써, 추가 연료 예측의 신뢰성과 실효성을 높였으며, 이 모델은 기존 방법 대비 연료 소비 예측의 정확도를 향상해 항공사의 경제적 비용 절감 효과가 있음을 증명하였다. 랜덤포레스트는 앙상블(ensemble) 학습 방법의 일종으로 여러 개의 의사결정트리를 생성하고 그 예측을 종합하여 최종 결과를 도출하는 기법이다. 앙상블 예측값은 아래의 식 (1)에서 최종 예측값 $\hat{y}(x)$ 는 M 개의 의사결정트리 예측에 대한 평균으로 계산한다. 여기서 T 는 j 번째 트리의 예측 함수, Θ_j 는 해당 트리의 분할 파라미터를 나타낸다.

$$\hat{y}(x) = \frac{1}{M} \sum_{j=1}^M T_j(x; \Theta_j) \quad (1)$$

$$\text{where } T_j(x; \Theta_j) = \frac{1}{|A_{\text{invert}} \mathbf{x}_i \in A_j(\mathbf{x}, \Theta_j)|} \sum Y_i$$

개별 트리의 예측은 훈련 표본 X 가 포함된 말단노드 $A(x, \Theta)$ 내 관측값의 평균이다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

RMSE는 실제값과 예측값 간의 제곱 오차의 평균에 루트를 취한 값으로, 큰 오차에 더 민감하게 반응하고, 여기서 \bar{y} 는 관측값의 평균이다. MAE는 실제값과 예측값 간의 절대오차의 평균으로, 모든 오차를 동등하게 취급하여 이상치에 상대적으로 덜 민감하다.

랜덤포레스트의 주요 장점 중 하나는 변수 중요도(variable importance)를 계산하여 어떤 입력 변수가 예측에 가장 큰 영향을 미치는지를 정량적으로 파악할 수 있다(Random Forests for Big Data, 2017).

$$CV_{RMSE} = \frac{1}{10} \sum_{k=1}^{10} RMSE(D_k^{\text{test}}) \quad (5)$$

모델의 일반화 능력을 높이기 위해 10-겹 교차 검증을 적용하였다. 10-겹 교차 검증(10-fold cross validation)은 전체 데이터를 크기가 거의 같은 10개의 부분집합(폴드)으로 나눈 뒤, 각 폴드가 한 번씩 테스트 데이터가 되고 나머지 9개 폴드는 학습 데이터로 사용하여 총 10번 모델을 학습·평가하는 방법이다. 이렇게 얻은 10개의 평가 지표(예: 정확도, RMSE 등)를 평균값으로 통합하여 모델의 최종 성능을 산출함으로써, 데이터 분할 방식에 따른 편향을 줄이고 모델의 일반화 성능을 보다 신뢰성 있게 평가할 수 있다.

IV. 분석 결과

랜덤포레스트 모델 분석은 1,034개 표본과 8개의 예측 변수를 사용했으며, 10-겹 교차 검증으로 모델을 검증했다. 하이퍼파라미터인 mtry 선택 결과, mtry=5 일 때 RMSE가 가장 낮아 최적의 변수 선택 수로 결정되었다. 변수 중요도 분석에서 distance (41.9%)가 가장 큰 영향력을 보였고, payload(24.8%), temperature(10.9%), speed(8.4%), wind(6.2%), altitude(5.7%), APMS(2.1%)로 분석되었다(Fig. 1). WB 변수 중요도는 0%로 나타났지만, 이는 해당 변수가 전혀 영향을 미치지 않는다는 의미가 아니라, 다른 변수들에 비해 상대적으로 영향력이 낮음을 의미한다. 이는 상업용 항공기가 출발 전 무게중심을 최적으로 조정하여 운항하기 때문에, 실제 연료 소모에 있어 무게중심의 영향이 다른 요인들과 비교하면 제한적으로 작용한 것으로 해석할 수 있다. 테스트 세트 성능 평가 결과, RMSE는 501.52, 결정 계수인 R^2 는 0.601, MAE는 358.09로 나타났다. 교차 검증 및 테스트 성능이 유사하여 모델이 안정적이며, 약 60.1%의 설명력으로 distance와 payload가 결과에 가장 큰 영향을 미치는 주요 요인임을 확인했다.

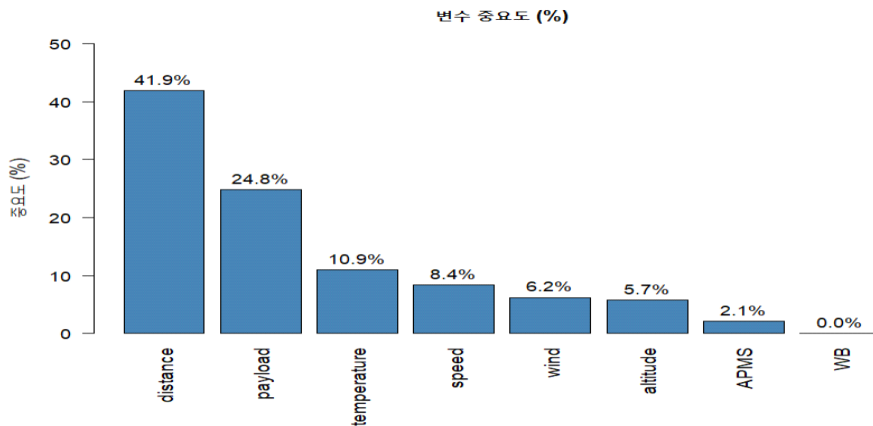


Fig. 1. Variable importance

V. 결 론

본 연구는 랜덤포레스트 회귀 모델을 사용하여 항공기 연료소모량에 영향을 미치는 주요 요인을 체계적으로 분석하였다. 분석 결과, 운항 거리(distance)와 유상하중(payload)이 연료 소모에 가장 큰 영향을 미치는 변수로 확인되었으며, 그 외에도 온도, 바람, 고도, 속도 등이 연료 사용에 영향을 주는 것으로 나타났다. 무게중심은 상대적으로 영향력이 낮았는데, 이는 상업용 항공기가 출발 전 무게중심을 최적으로 조정하여 운항하기 때문으로 해석된다. 모델의 예측 성능(RMSE, R^2 , MAE 등)도 교차 검증과 테스트 세트에서 유사하게 나타나, 모델이 안정적이고 신뢰할만함을 보여준다. 하지만 연료소모량에 영향을 미칠 수 있는 갑작스러운 기상 악화, 공항 혼잡, 비상 상황과 같은 예외적 상황이나 정책 변화 등은 본 연구의 모델에 충분히 반영할 수 없었다는 점이 한계라고 할 수 있다. 향후 연구에서는 이러한 예외적 요인들을 반영하기 위해, 탑재연료량 대비 실제 연료소모량이 과도하게 많은 사례를 더미 변수 등으로 설정하여 분석을 보완할 계획이다. 이러한 결과는 데이터 기반의 과학적 연료탑재 기준 마련과 연료 효율성 제고, 나아가 항공 산업의 탄소 감축 목표 달성에 기여할 수 있음을 시사한다.

References

1. Boston Consulting Group, "Why airlines need to nudge behavior, not just hedge fuel co-

- sts", 2022, Available from: <https://www.bcg.com/publications/2022/why-airlines-need-to-do-more-than-just-hedge-fuel-costs-to-help-reduce-fuel-consumption>
2. CAPA, "Airline industry completes full recovery with exceptional operational performance in 2024", 2025, Available from: <https://centreforaviation.com/news/insights/airline-industry-completes-full-recovery-with-exceptional-operational-performance-in-2024-709251>
3. Genuer, R., Poggi, J.-M., Tuleau-Malot, C., and Villa-Vialaneix, N., "Random forests for big data," *Big Data Research*, 9, 2017, pp. 28-46.
4. Harada, et al., "Analysis of air traffic efficiency using dynamic programming trajectory optimization", 29th Congress of the International Council of the Aeronautical Sciences, ICAS, St. Petersburg, 2014.
5. IATA, "Airline Profitability Outlook Improves for 2024", 2024, Available from: <https://www.iata.org/en/pressroom/2024-releases/2024-06-03-01/>
6. IATA, "Supply Chain Issues Continue to negatively Impact Airline Performance into 2025", 2024, Available from: <https://www.iata.org/en/pressroom/2024-releases/2024-12-10-02/>

7. IEEFA, "Can South Korea's Aviation Industry Pivot to Green Skies?", 2024, Available from: <https://ieefa.org/resources/can-south-koreas-aviation-industry-pivot-green-skies>
8. Kim, D., Seth, A., and Liem, R. P., "Data-enhanced dynamic flight simulations for flight performance analysis", *Aerospace Science and Technology*, 121, 2022.
9. Mgbachi, C., "AI in business aviation route optimization: reducing fuel consumption and environmental impact", *Journal of Business and Strategic Management*, CARI Journals Limited, 9(5), 2024, pp.48-82.
10. Miyoshi, C. and Mason, K. J., "Airline business models and their respective carbon footprint", *Omega*, 2009.
11. Ren, D., and Leslie, L. M., "Impacts of climate warming on aviation fuel consumption." *Journal of Applied Meteorology and Climatology*, 58(7), 2019, pp.1473-1485.
12. Seymour, K., Smith, T., and Lee, J., "Development of an integrated framework for high-fidelity flight profile simulations and simplified fuel burn prediction models", *Aerospace Science and Technology*, 95, 2020.
13. Singh, R., Sharma, S., and Agarwal, A., "Analysis of fuel consumption for commercial aircraft: A case study", *International Journal of Engineering and Advanced Technology (IJEAT)*, 8(6S3), 2019, pp.1096-1102.
14. Yoo, J., "Aircraft fuel efficiency improvement and effect through APMS", *Journal of the Korean Society for Aviation and Aeronautics*, 31(2), 2023, pp.81-88.
15. Wang, Y., Luo, L., Zhou, Q., Lin, J., and Wu, Z., "A Recommendation for flight extra fuel based on random forest", *Proceedings of the 2020 9th International Conference on Computing and Pattern Recognition (ICCPR '20)*, Association for Computing Machinery, New York, NY, USA, 2021.
16. Wu, Z., Zhang, N., Gao, K., and Peng, R., "Flight trip fuel volume prediction based on random forest with adjustment to risk preference", *Acta Aeronautica et Astronautica Sinica*, 43(2), 2022.