Tail 687 Cruise Efficiency — Problem Statement

DATA 5100 — Fall 2025 — Group Project
Team: Duy Nguyen · Hemant Kumaar Aruljothi · Prithika Kandasamy

Executive Summary

We analyze NASA DASHlink flight-recorder data for Aircraft **Tail 687** to identify **actionable** levers that reduce **cruise fuel burn**. Focusing on altitude–speed choices under real winds and weight, we will produce interpretable, operations-ready recommendations (e.g., efficient altitude–Mach bands by weight and along-track wind) rather than generic correlations.

Research Question & Objectives

Primary question.
How do **altitude** and **airspeed** (Mach/TAS) combinations affect fuel efficiency during **steady-state cruise** for Tail 687, and what **cruise profiles** are optimal under varying **wind** and **aircraft weight**?

Objectives.

1. Quantify marginal effects on total fuel flow (FF_total) from altitude, Mach, along-track wind, angle of attack, and weight.

2. Map efficient operating bands (Altitude×Mach) across wind/weight regimes.

3. Screen engine-level dispersion (FF vs. N1/EGT) for potential maintenance/efficiency issues.

Data & Analytical Sample

- **Source.** NASA DASHlink (Tail 687), 2012: **652** flights, **186** parameters, mixed sampling rates.

- **Cruise-capable flights. 312** flights reached > **25,000 ft**; we include all 312 to leverage maximum statistical power.

- **Sampling strategy.** Restrict to **4 Hz** signals to avoid interpolation and ensure perfect time alignment.

- **Cruise definition.** Altitude > 25,000 ft; |altitude rate| ≤ 500 ft/min; persistent windows (exclude transient level-offs — meaning we **remove short periods where the aircraft briefly levels off** (stops climbing or descending for a moment) but does not remain in steady cruise).
  - We only include segments where the aircraft is consistently at cruise altitude and stable, not just passing through or pausing briefly before changing altitude again.
- **Final analytical set.** ~**1,882,573** cruise records (~**130 hr** at 4 Hz).
- **Outliers checks.** Alt 25–35 kft; Mach 0.512–0.748 (typ. 0.69–0.72); mean N1 ≈ 91% (σ≈3%); EGT ≈ 557 °C (σ≈22 °C).
- **Cleaning.** Removed ~**0.4%** boundary points (FF outliers consistent with climb/descent bleed-through). **No missingness** in selected 4 Hz variables.

Variables & Feature Engineering

Target. FF_total (lbs/hr) = $FF_1+FF_2+FF_3+FF_4$. (We will check engine-level dispersion for asymmetries.)

Predictors (initial set).

- **Engine performance:** N1 (avg), N2 (avg), EGT (avg).
- **Flight conditions & controls:** Pressure altitude, Mach, TAS, corrected angle of attack (AOAC), $PLA_{1\ldots4}$, N1T/N1C.
- **Environment:** Decompose winds into **along-track** (head/tail) and **cross-track** components.
- **Aircraft state (derived): Weight(t)** via initial fuel quantity minus integrated FF_total (synchronizes 1 Hz to 4 Hz).

Collinearity & interactions.
We will investigate to diagnose with correlations/VIF; prefer parsimonious sets (e.g., Mach over TAS if redundant). Include **Altitude×Mach**, **Weight×Mach**, **Along-wind×Mach** to capture regime dependence.

Methods & Inference Plan

EDA.

- Heatmaps/contours of FF_total vs. (Altitude, Mach).

- Partial-residual diagnostics for AOAC, along-wind, weight.
- Per-engine FF vs. N1/EGT to flag asymmetric loads.

Modeling (inference-first).

- Multiple linear regression with interpretable coefficients (e.g., +0.01 Mach → +ΔFF_total at fixed Altitude/Wind/Weight).
- Sensitivity across alternative specifications (N1 vs. PLA vs. N2).
- Robust SEs for time correlation (cluster-robust/HAC); optional altitude-band fixed effects.

Interpretation.
Report **practical effect sizes** (lbs/hr, % change) and **decision charts** (efficient Altitude×Mach bands by wind/weight) with uncertainty.

Scope note. We are aware that observational data are limited to causal claims (ATC constraints, comfort, fuel policies unobserved). We aim for physics-consistent associations and operations-testable guidance.

Expected Deliverables

1. **Cruise efficiency maps:** Altitude×Mach grids predicting FF_total, stratified by **weight** and **along-track wind**.
2. **Marginal-effects table** (95% CIs) for key levers (Altitude, Mach, AOAC, along-wind, Weight).
3. **Engine-health snapshot:** per-engine FF vs. N1/EGT dispersion.
4. **Operational playbook:** concrete targets (e.g., “At 29–31 kft and weight W, fly Mach 0.70–0.71 when headwind ≥ X kt”).

Risks & Limitations

- **Endogeneity/confounding:** route/ATC/payload; partially mitigated by conditioning on weight/wind and steady-state cruise filter.
- **Generalizability:** Tail- and era-specific.
- **Sensors/rates:** 4 Hz restriction removes some high-rate dynamics but preserves core cruise physics.

- **Actionability:** Recommendations require operational validation.

Project Plan & Roles

Two meetings/week: (1) planning/strategy; (2) coding/QA. Minutes shared with instructor (**fischer9@seattleu.edu**). Day-to-day via GitHub (issues/PRs).

Initial sprint (3 tasks).

- **Task 1 — .mat data dictionary (Hemant & Duy):** enumerate fields (data, Rate, Units, Description); export CSV dictionary.
- **Task 2 — 4 Hz master table (Prithika & Duy):** finalize variable list; extract aligned 4 Hz signals per flight; stack; add flight_id.
- **Task 3 — Cruise filter & QA (Collaborative; domain review by Prithika):** apply filters; drop non-operational segments (e.g., N1<15%); fix impossibles; produce clean EDA outputs.

Tools & Reproducibility

- **Python:** pandas/NumPy (ETL), SciPy (utils), statsmodels & scikit-learn (models), matplotlib & seaborn (viz).
- **Repro:** requirements.txt, data dictionary CSV, deterministic ETL script; GitHub for code/reviews.
- **Compute:** In-memory pandas is sufficient for post-filter sizes.

References

- Boeing Commercial Airplanes. (2024). The Boeing ecoDemonstrator Program [Backgrounder]. Boeing.
- Boeing. (2024). ecoDemonstrator Program. Boeing Sustainability. [Link]
- Matthews, B. (2012). Flight Data for Tail 687 [Dataset]. NASA DASHlink (C3).

Appendix: Key Quantities

- 652 total flights; **312** with cruise > 25 kft.
- ~**1.88 M** 4 Hz cruise records (~**130 hr**).

- Cruise ranges: Alt 25–35 kft; Mach 0.512–0.748; mean N1 ≈ 91%; mean EGT ≈ 557 °C.
- Outlier removal: ~**0.45%** boundary points.