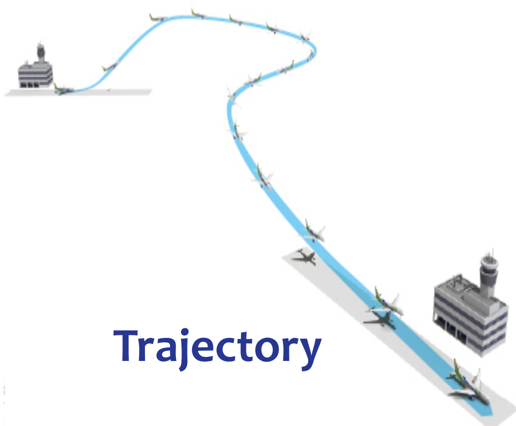
Gaussian Process Regression for Modeling Aircraft Fuel Flow

Hamsa Balakrishnan, Yashovardhan Chati

Motivation: Aircraft fuel burn inventories



Trajectory

Fuel burn

* Need to determine the fuel burn impact of different aircraft trajectories (procedures)
* Must account for variability seen in real world operations

Motivation: Aircraft fuel burn inventories

- Need to determine the fuel burn impact of different aircraft trajectories (procedures)
- Must account for variability seen in real world operations

Current approaches

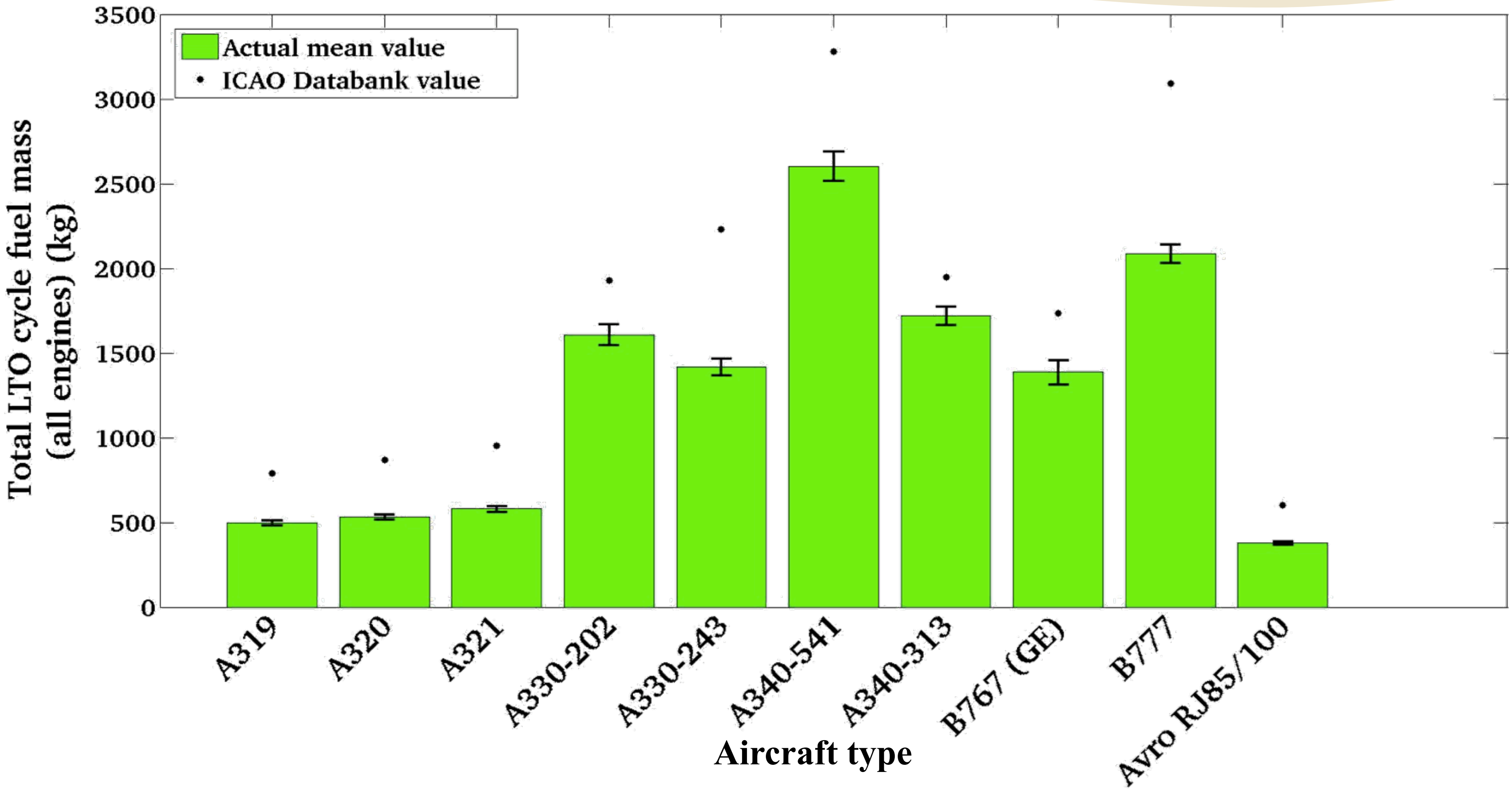
* ICAO Aircraft Emissions Databank
  * Fuel flow rates from ground-based uninstalled engine certification tests
  * Four thrust settings: takeoff (100%), climb out (85%), approach (30%), ground idle (7%)

* Base of Aircraft Data (BADA)
  * Empirical equations, aircraft type specific equation coefficients in database

* Physics-based simulation software (e.g., engine simulations)
  * Require knowledge of many parameters, which are typically unknown
  * Traditionally intended for design studies

* Disadvantages
  * Rely on information from flight performance manuals and ground tests
  * Deterministic models

Example:
ICAO databank estimates vs. actual fuel burn



[Chati and Balakrishnan, ICRAT 2014]

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

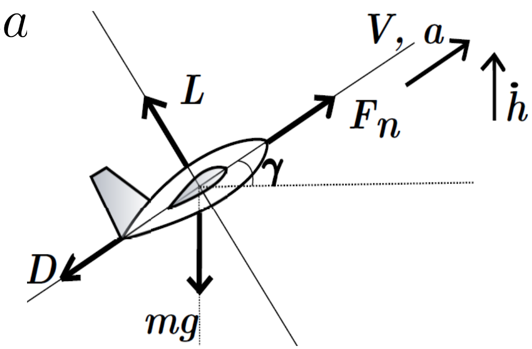
Statistical modeling of aircraft fuel flow

* Use Flight Data Recorder (FDR) records from operational flights to develop statistical models
* Explicitly model uncertainty of estimates using confidence intervals
* Leverage insights from physics (e.g., for feature extraction by considering dependence on various variables)
* Predictive variables restricted to trajectory variables

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Feature selection

$$
\begin{aligned}
F_n &= q\mathcal{S}C_{D_0} + \frac{C_{D_2} m^2 g^2}{q\mathcal{S}} - \frac{C_{D_2} m^2 g^2 \dot{h}^2}{q\mathcal{S}V^2} + mg\frac{\dot{h}}{V} + ma \\
\dot{m}_f &= \frac{TSFC \times F_n}{N_{eng}} \\
\dot{m}_f &\approx \dot{m}_f(q\mathcal{S}, m, \frac{\dot{h}}{V}, a, C_{D_0}, C_{D_2}, TSFC)
\end{aligned}
$$



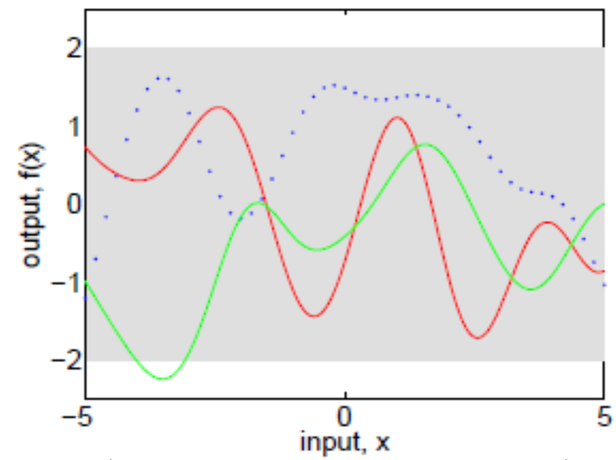
* Simplifying assumptions
  * International Standard Atmosphere
  * Ground speed instead of true airspeed
  * Derivative of ground speed instead of actual acceleration
  * BADA assumptions on Thrust Specific Fuel Consumption (TSFC) and drag coefficients

Gaussian Process Regression for predicting aircraft fuel flow

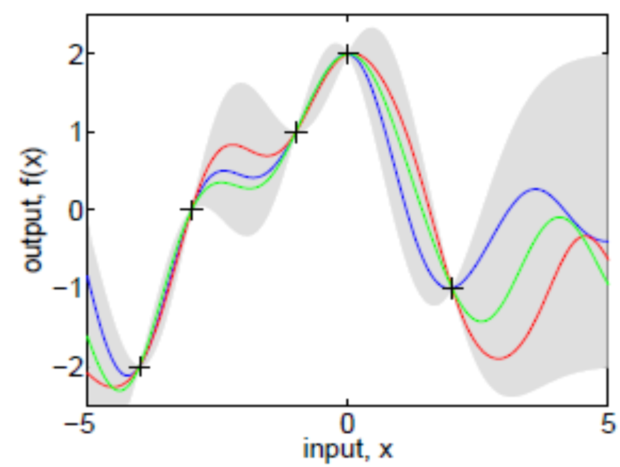
* Features as previously described; output is average fuel flow rate per engine ($\dot{m}_f$) in kg/s
* Gaussian Process Regression
  * Nonparametric, probabilistic method

$$\begin{aligned} y &= f(\mathbf{x}) + \epsilon,\ \epsilon \sim \mathcal{N}(0, \sigma_n^2) \\ f &\sim \mathcal{GP}(0, k(\mathbf{x_p}, \mathbf{x_q})) \end{aligned}$$

  * A function is said to be drawn from a Gaussian Process when any finite set of function values follows a joint Gaussian distribution
* Advantages
  * No need to choose basis functions
  * Fast estimation of predictive distributions
* Disadvantages
  * Computationally expensive due to matrix inversion



(before making observations)



(after making observations)

Rasmussen et al., 2006.

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Application to the A321-100 aircraft

* Divided into training, validation and test sets
  * 65% flights in training set (76 flights)
    * 18261 observations in ascent, 933 in cruise, 34110 in descent
  * 15% flights in validation set (18 flights)
  * 20% flights in test set (23 flights)
* Aircraft takeoff mass used as a surrogate for instantaneous mass
  * Extended to propagate the takeoff mass
* Metrics
  * **Mean error:** Mean relative prediction error on test data
  * **Prediction coverage:** Percentage of observations in test dataset for which 95% prediction intervals include the actual fuel flow rate values

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

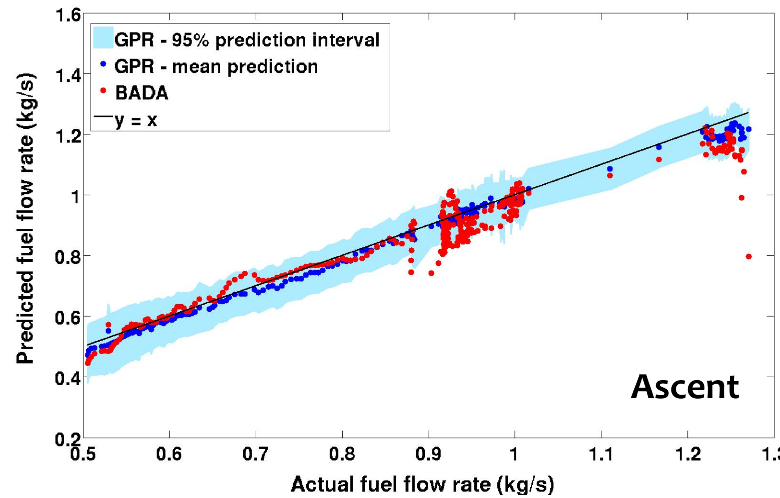
[Chati and Balakrishnan, ICCPS 2017]

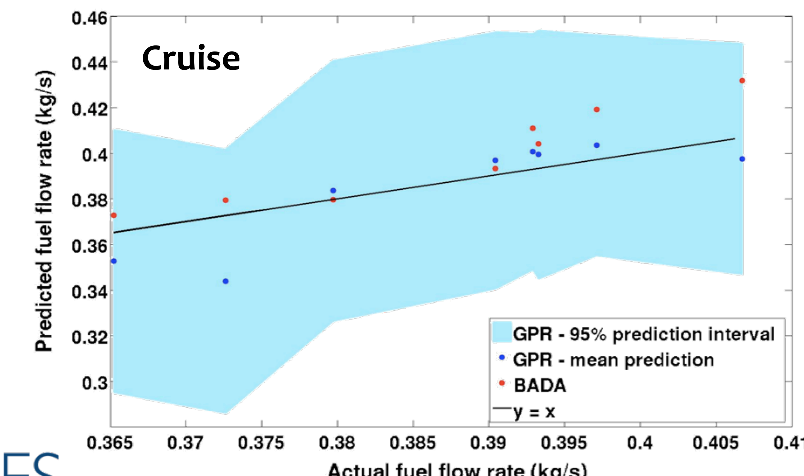
Model performance

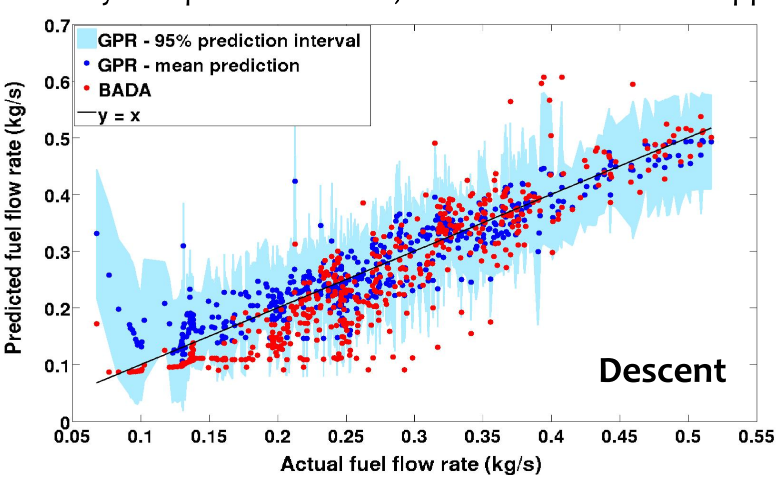
Performance on test data

Phase	Metric	GPR	BADA*
Ascent	Mean Error (%)	2.87	5.49
	Prediction Coverage (%)	94.71	n/a
Cruise	ME (%)	6.73	8.20
	PC (%)	96.22	n/a
Descent	ME (%)	15.52	23.56
	PC (%)	92.66	n/a

* Base of Aircraft Data. These values use information from the FDR, and thus represent an upper bound on the performance of the BADA estimates. BADA yields point estimates, so the PC metric is not applicable.







FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

[Chati and Balakrishnan, ICCPS 2017]

Next Steps

* Gaussian Process Regression appears to significantly improve the estimation of operational fuel burn
  * Extensions to other engine performance variables, such as thrust
  * **Transition to practice:** Enhancing the surface models of the FAA's Aircraft Environmental Design Tool (AEDT)

* **Scaling to streaming/larger data sets:** Computationally expensive
  * Projection-based matrix factorization presents a promising approach (Bopardikar et al., IEEE Big Data 2016)
  * Pursuing collaboration with UTRC on this topic

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS