National Aeronautics and Space Administration

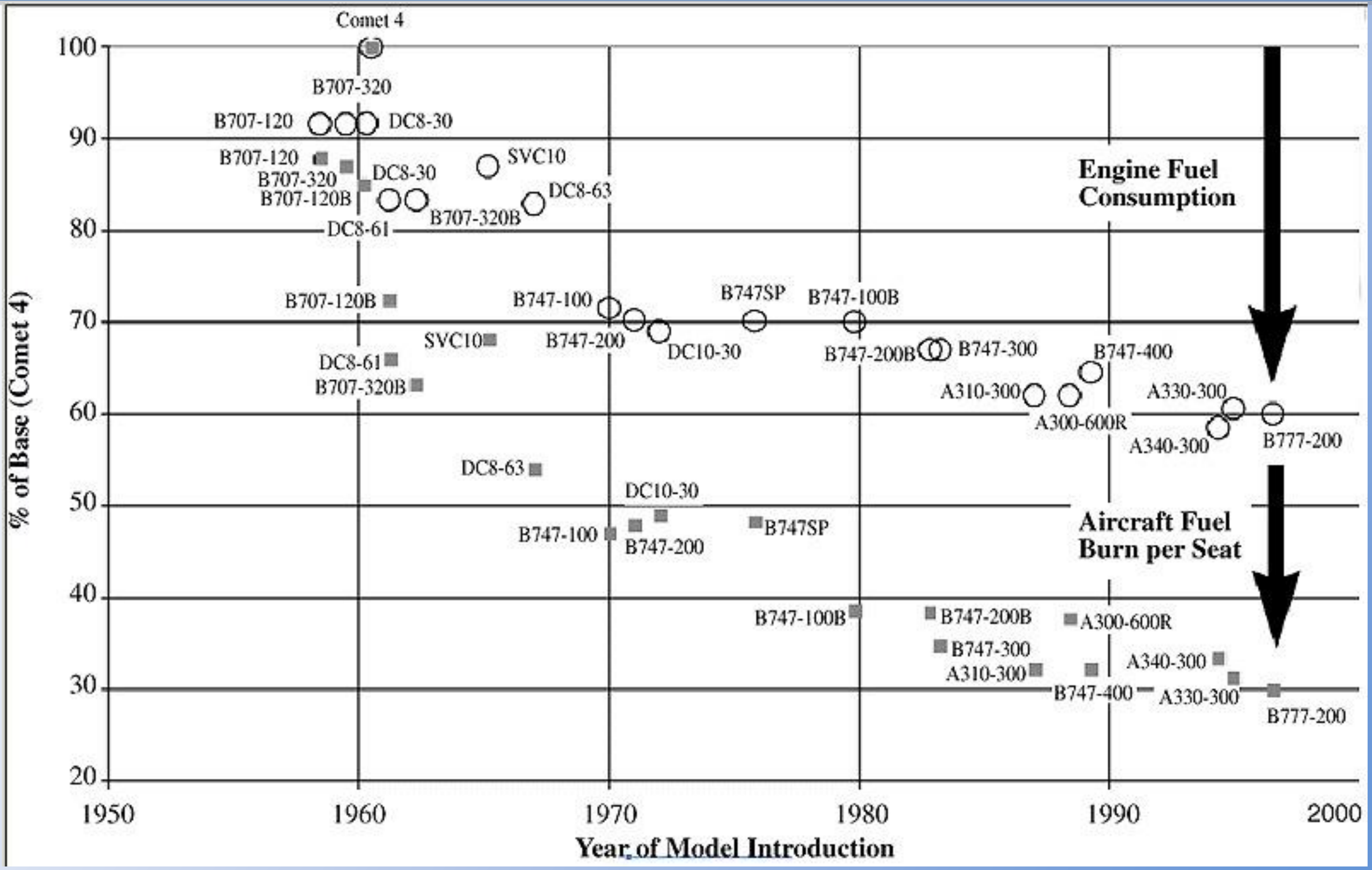
Reducing the Environmental Impact of Aviation: A Data Mining Approach to Instantaneous Estimation of Fuel Consumption

Ashok N. Srivastava, Ph.D., Principal Investigator, NASA Ames Research Center
Stephen Boyd, Ph.D., Stanford University
Sricharan Kumar, Ph.D., Stinger Ghaffarian Technologies Inc.

Introduction

- The environmental impact of aviation is enormous given the fact that in the US alone there are nearly 6 million flights per year of commercial aircraft.
- This situation has driven numerous policy and procedural measures to help develop environmentally friendly technologies which are safe and affordable and reduce the environmental impact of aviation.
- Many technologies require significant capital investment and retrofits which are long and costly enterprises.

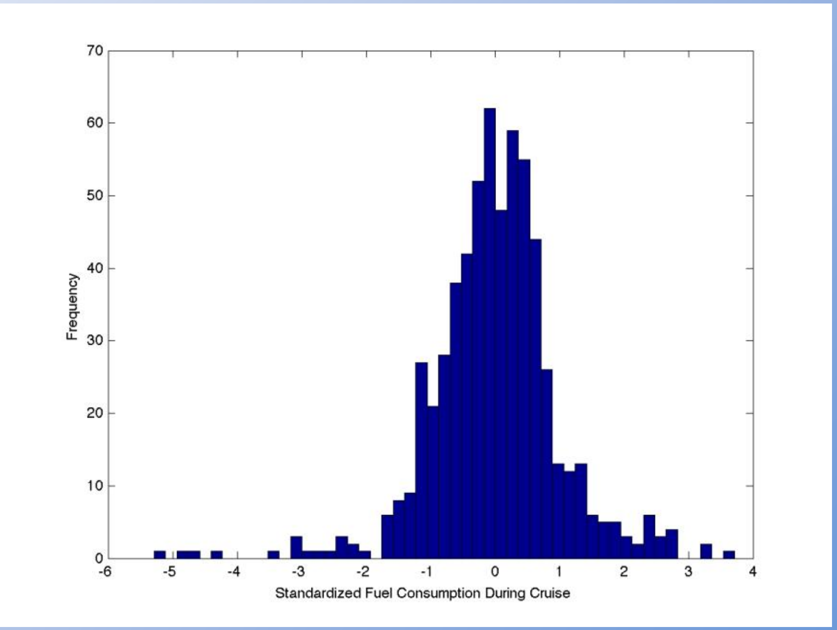
Fuel consumption per seat has declined dramatically



We estimate that a single Boeing 747 can emit as much as 457,000 kg of carbon dioxide into the high atmosphere during a long range flight

State of the art

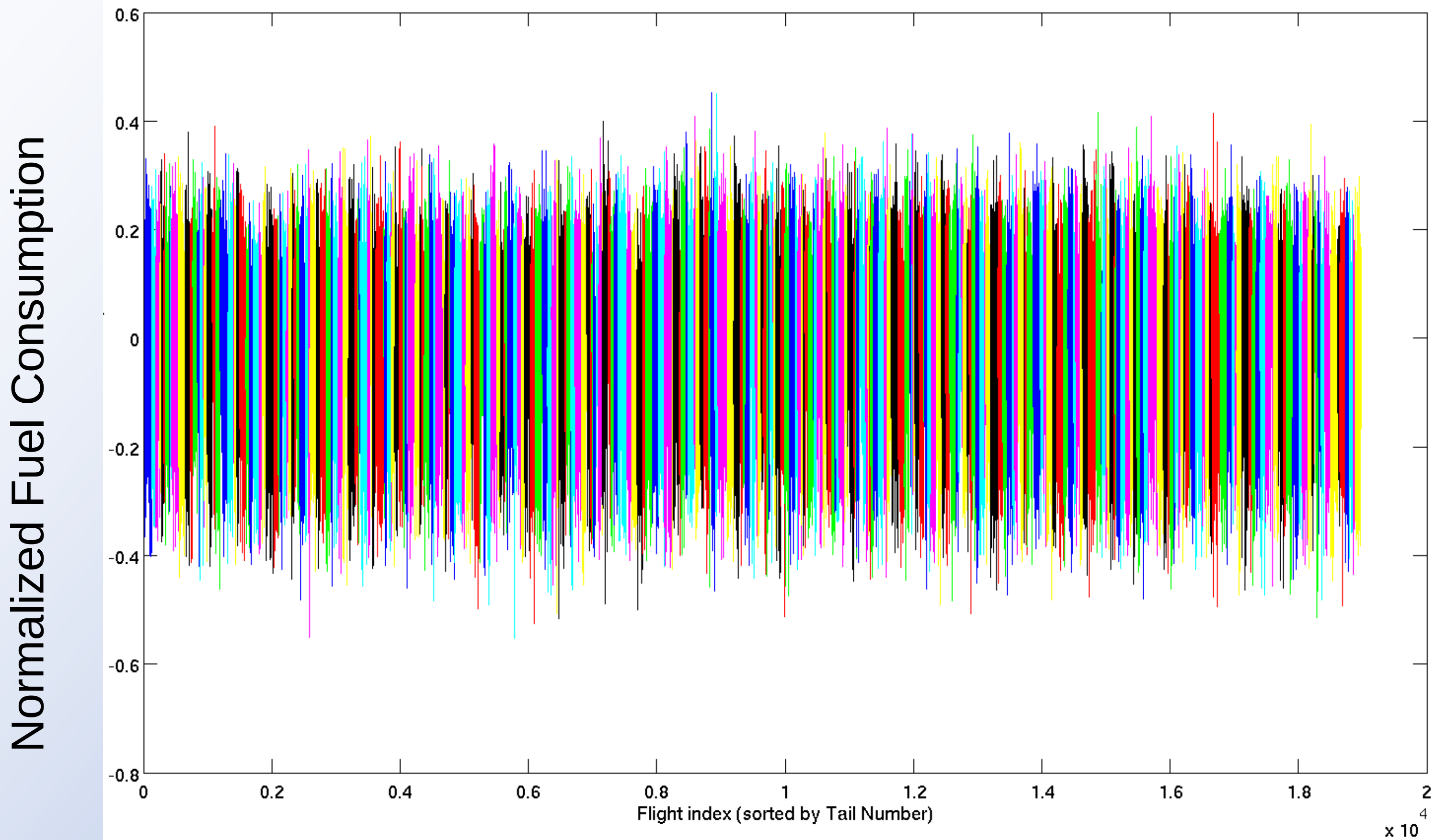
- *Fuel bias number*: compare the total fuel consumed by a flight against an average value based on historical data
- Drawback: Does not account for context of the flight over its course - weather, wind speed, payload etc.
- Subtle performance issues not revealed.



Is a high fuel usage on a specific flight significant or not given the observed data?

Which aircraft burns more fuel?

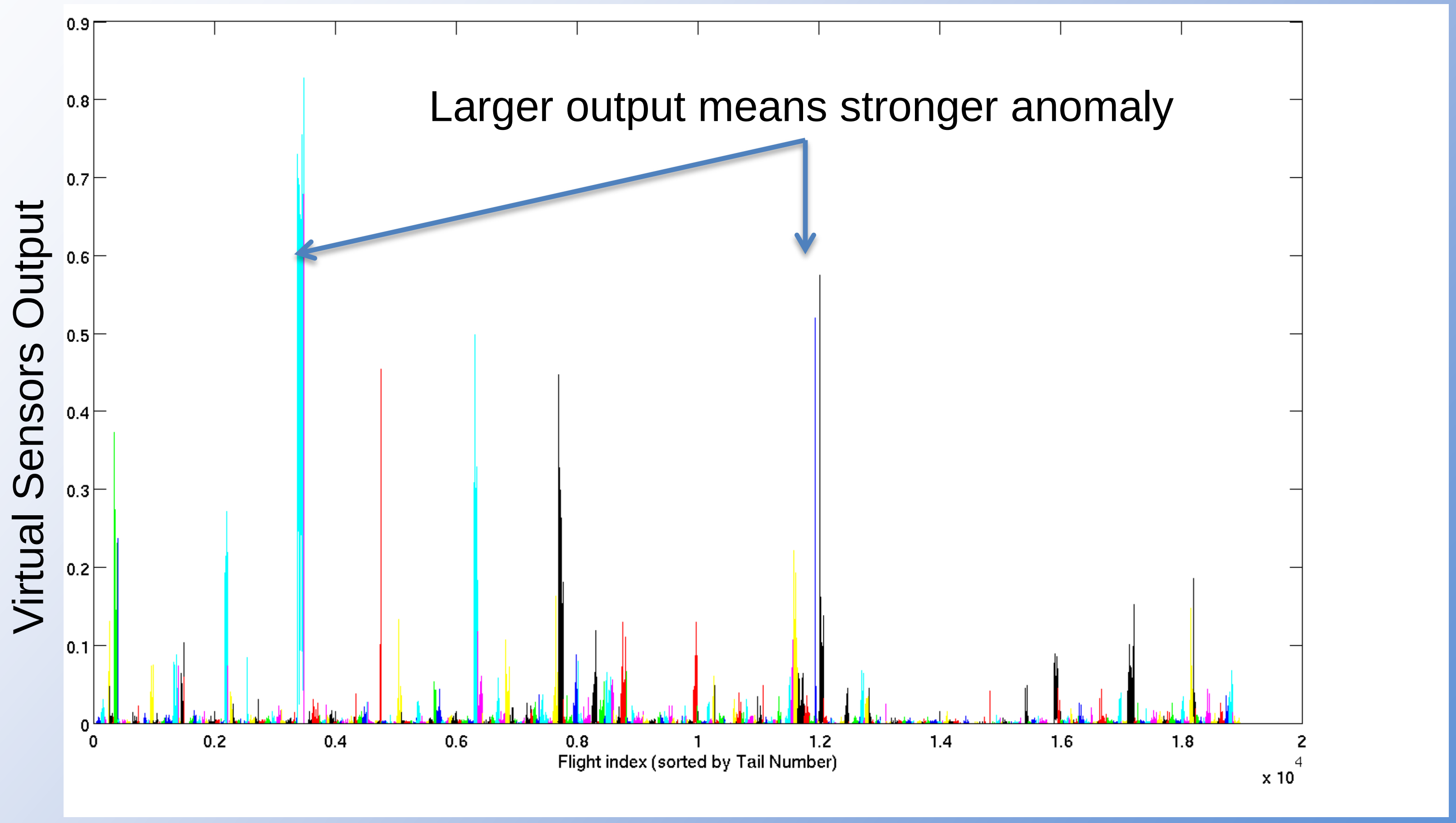
Fuel Consumption as a function of Tail Number for about 20K flights



Each color band is one tail number

Which aircraft burns more fuel?

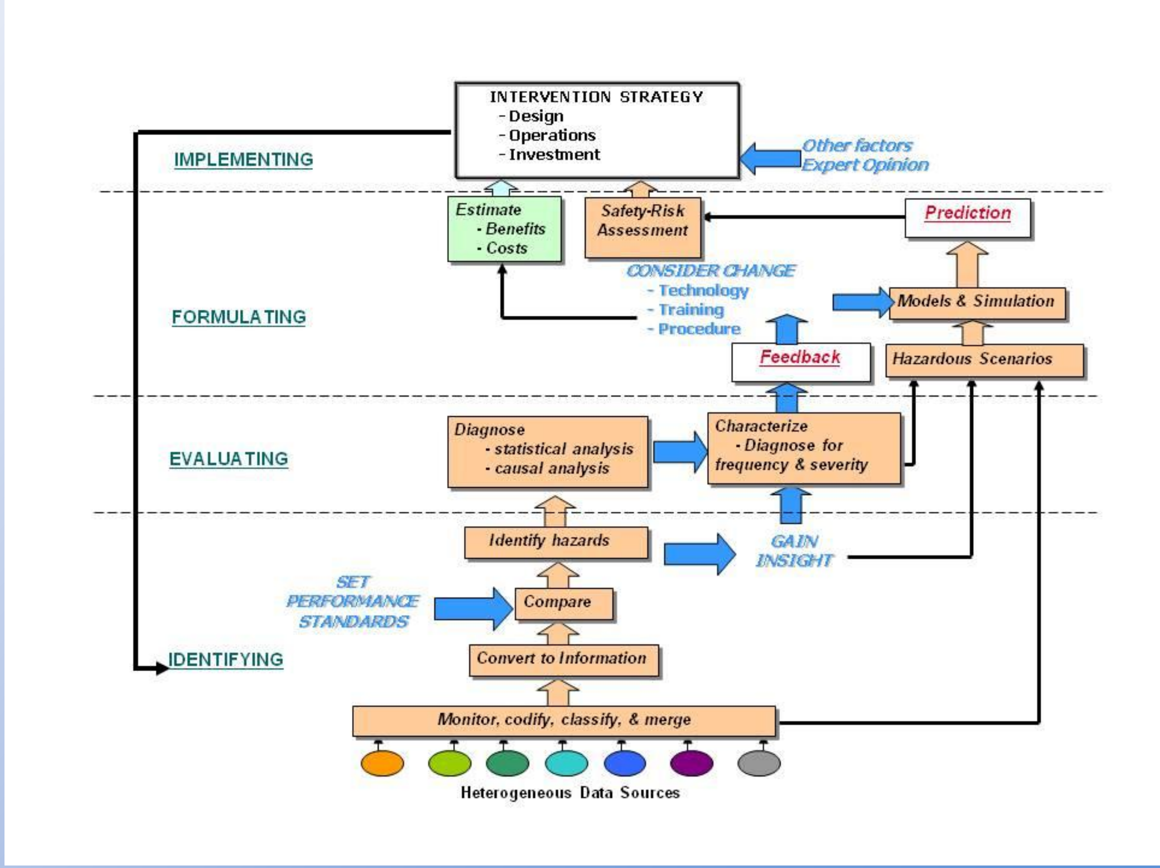
Output of Virtual Sensors as a function of Tail Number for about 20K flights



Each color band is one tail number

Innovation

A novel method based on *Virtual Sensors* to detect an anomaly in the consumption of fuel in modern aircraft using data that is already measured and monitored.

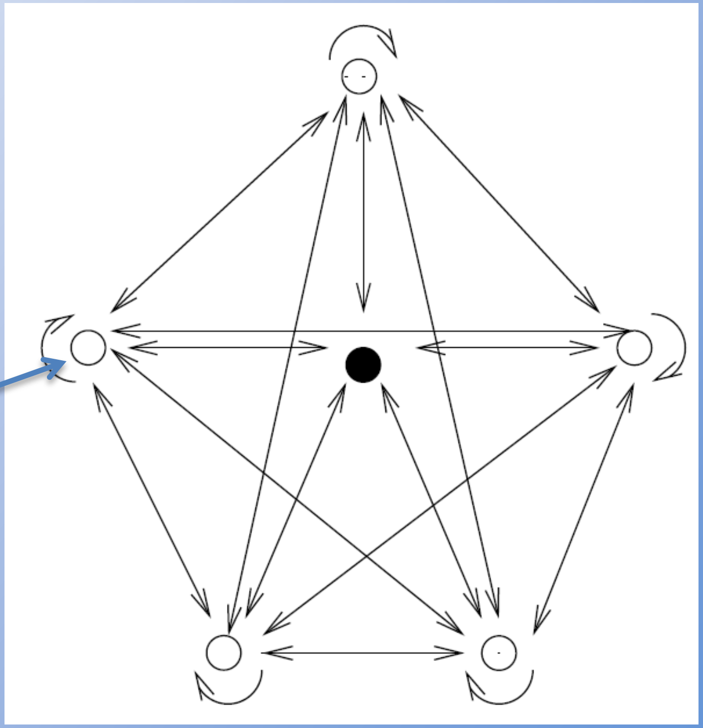


Fuel Consumption Model

$$\mathbf{h}_t = \Gamma(\mathbf{h}^*_{t-1})$$

$$\mathbf{x}_t = \Psi(\mathbf{x}^*_{t-1}, \mathbf{h}^*_t, u_t, \mathbf{c})$$

$$y_t = \Omega(\mathbf{x}_t) + \varepsilon_t$$



- h- hidden state of the aircraft
- x- observed system state from FOQA data
- u- pilot input
- c- context of flight (weather etc.)
- y- fuel burned at time t
- espilon- measurement noise

Virtual Sensors Algorithm

Algorithm 1: Virtual Sensors for Anomaly Detection

Input: $(\mathcal{X}, \mathcal{Y}, \mathcal{C}, \alpha, m, n)$, representing state variables, the target variable, the cost function for minimization, a multiplier on the number of standard deviations to use as the anomaly detection threshold, the number of models, the number of bootstrap samples, respectively.

Output: Sorted list of anomalies *List*

Initialization: Standardize inputs and outputs to have zero mean and unit variance;

begin

  for $k = 1$ *to* m **do**

    Draw bootstrap replicate with n **samples**: $(\mathcal{X}_k, \mathcal{Y}_k)$

    minimize cost function C to obtain estimate: $\hat{\mathcal{Y}}_k = G(\mathcal{X}_k, \theta_k)$;

  Compute mean and standard deviation of the estimates for the m models;

  Compute the percentage of the test data for a given flight that is larger than the mean $+ \alpha$ standard deviations;

  Return rank ordered list *List* of anomalous flights.

A. N. Srivastava, "Greener Aviation through Virtual Sensors: A Case Study, Data Mining and Knowledge Discovery, Data Mining and Knowledge Discovery, Volume 24 Issue 2, March 2012.

Building Virtual Sensors

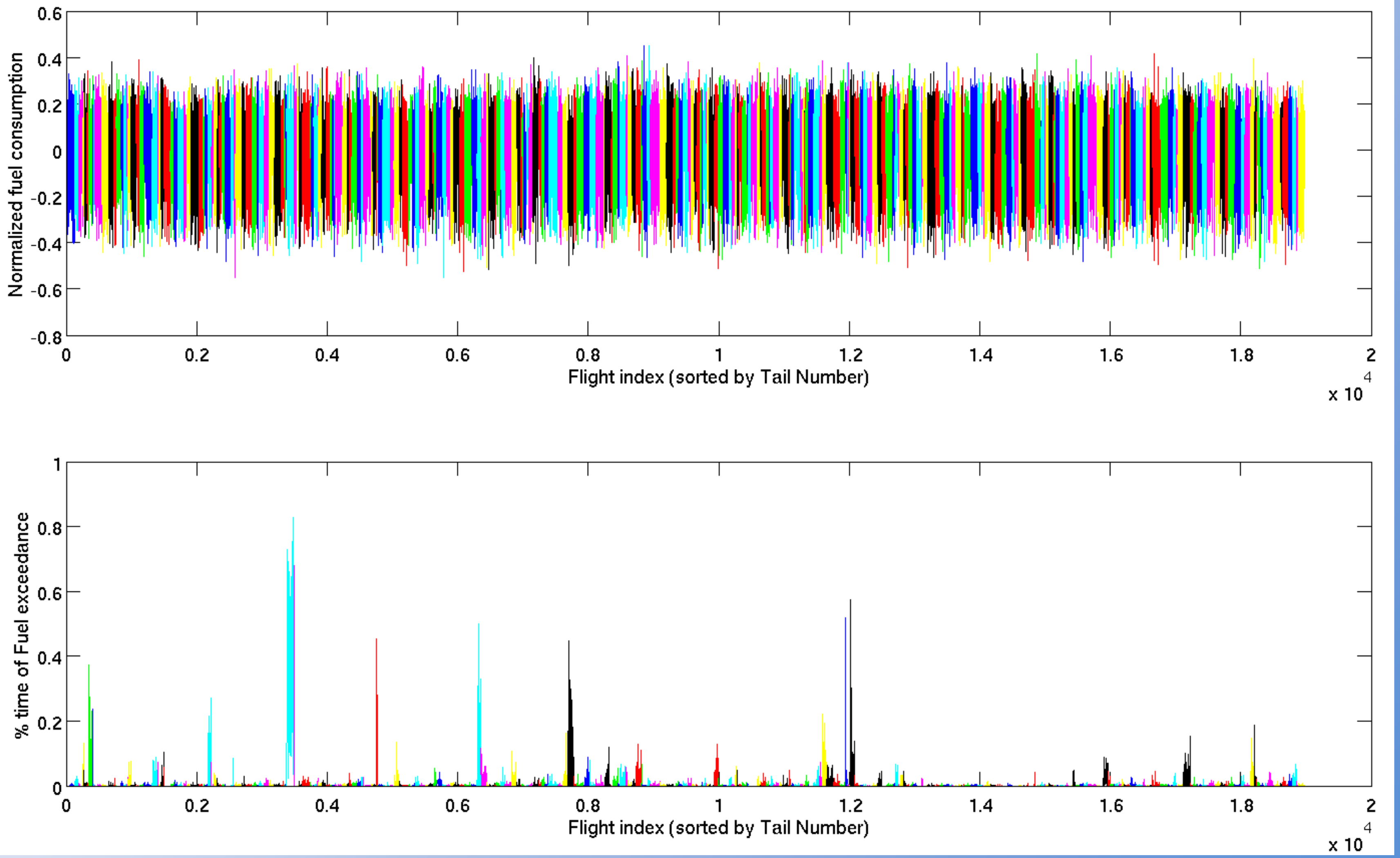
Training Phase

- **Build** nominal fuel consumption model as a function of aircraft state
- Use **Flight Operational Quality Assurance** (FOQA) data
- Use state-of-the-art, robust ensemble regression techniques to **predict** instantaneous fuel consumption.

Testing Phase

- Use model to **predict** instantaneous fuel consumption as a function of aircraft state
- If **true fuel consumption** is much higher than **predicted fuel consumption** we note an anomaly
- Flights with **large number** of such instants classified as anomalous

Virtual Sensors discover anomalies that are not detected by traditional measures

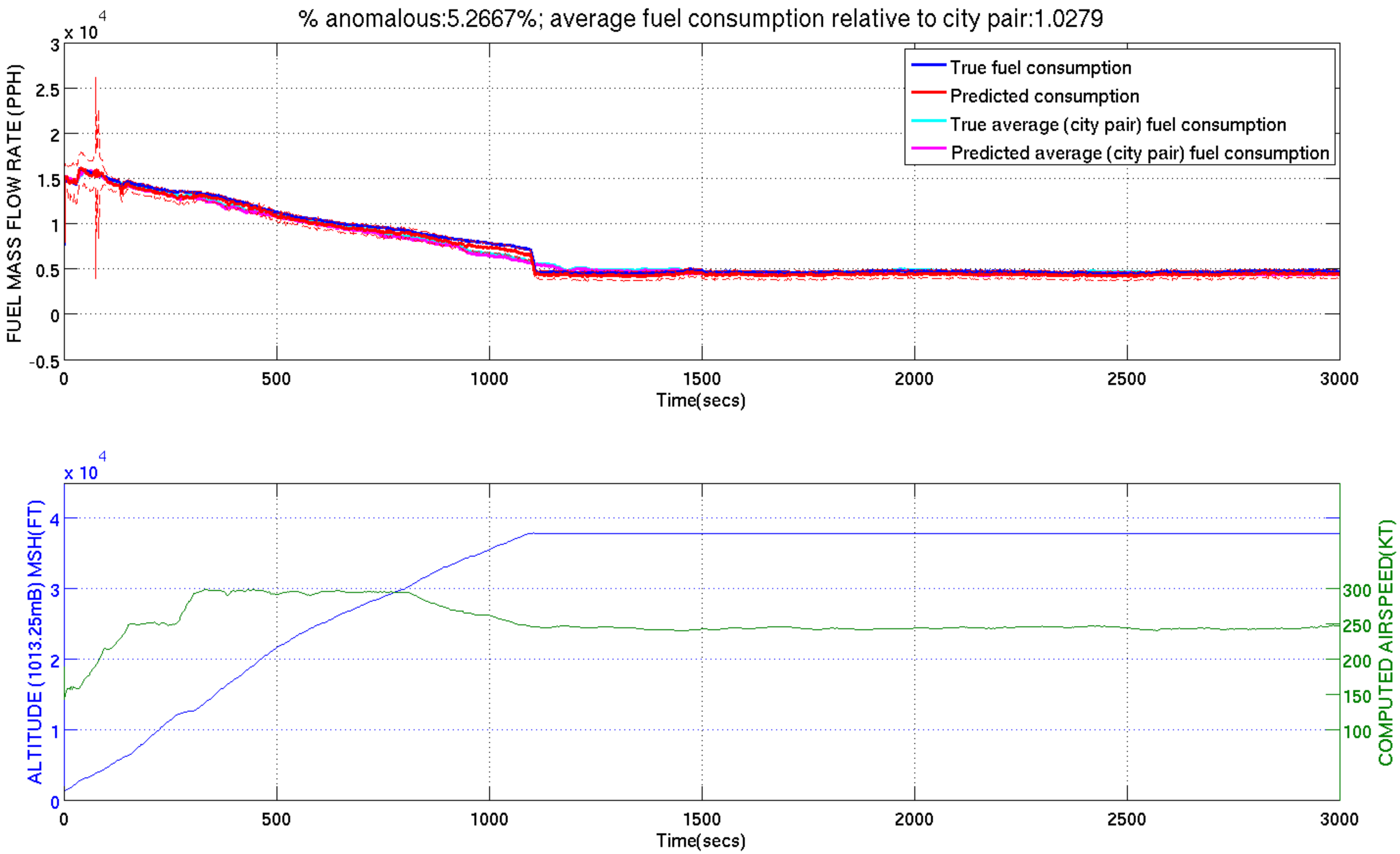


Output of Virtual Sensors Algorithm

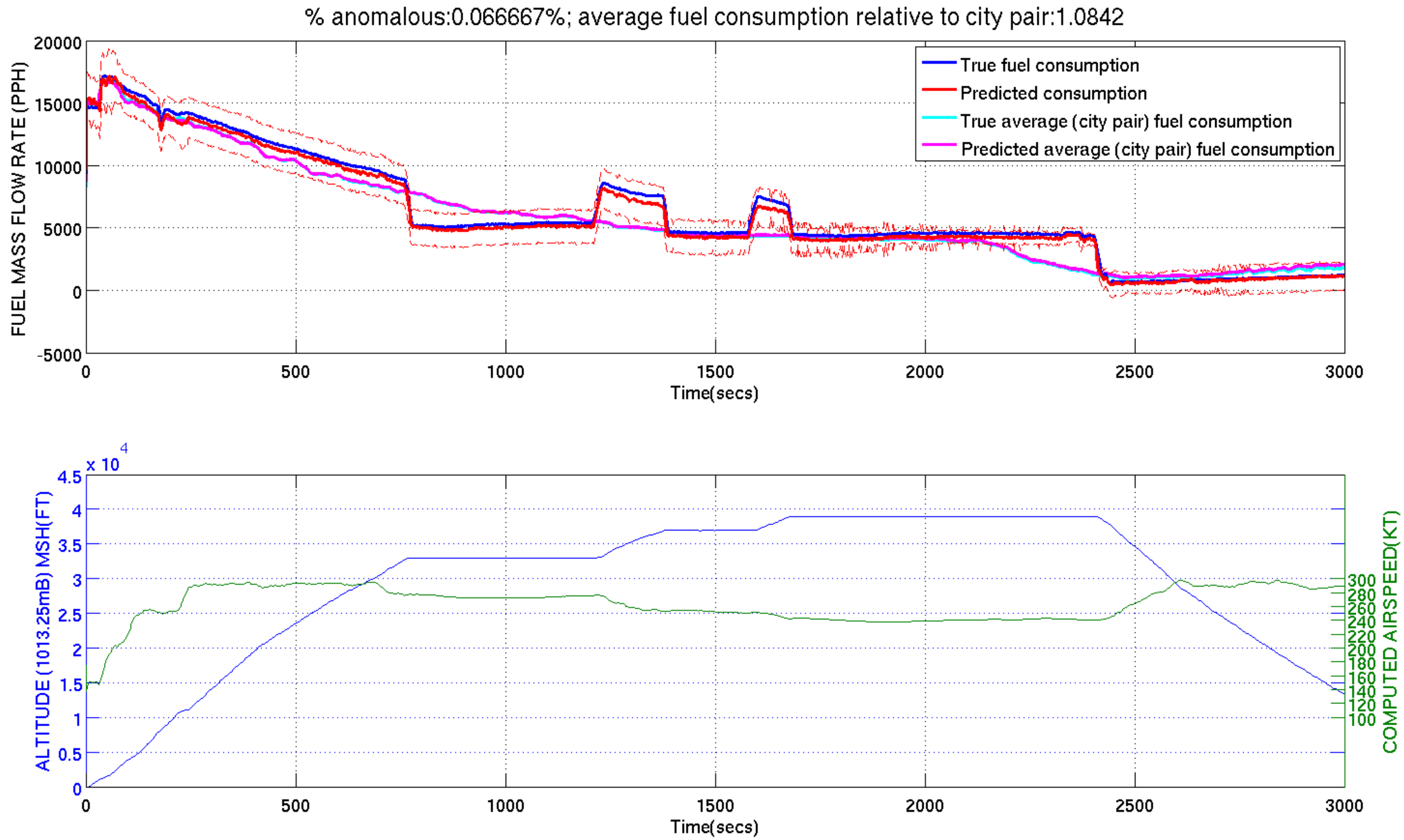
Tail Number*	% High	% Low	% OK	# of Flights
17	39.80	0.00	60.20	111
50	9.98	0.03	89.99	21
2	9.47	0	90.52	61
118	7.60	0.01	92.39	81
49	6.93	0.00	93.07	57
101	4.98	0.00	95.02	65
305	4.57	0.00	95.42	69
802	4.24	0.00	95.76	17
86	3.99	0.00	96.01	65
4	3.19	0.01	96.80	30

*anonymized

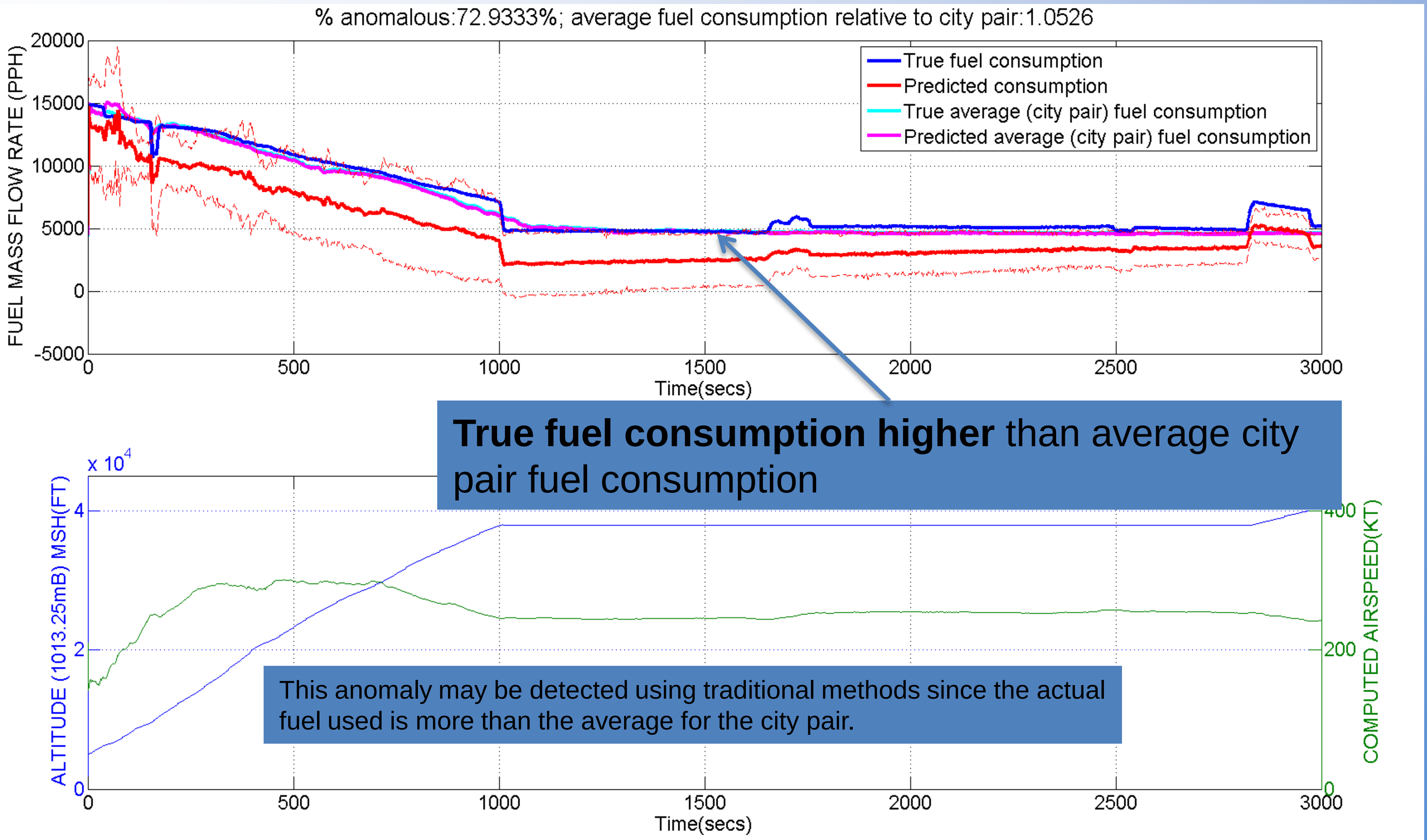
Nominal flight – actual fuel consumption falls within prediction bounds



Nominal flight - fuel consumption falls within prediction bounds



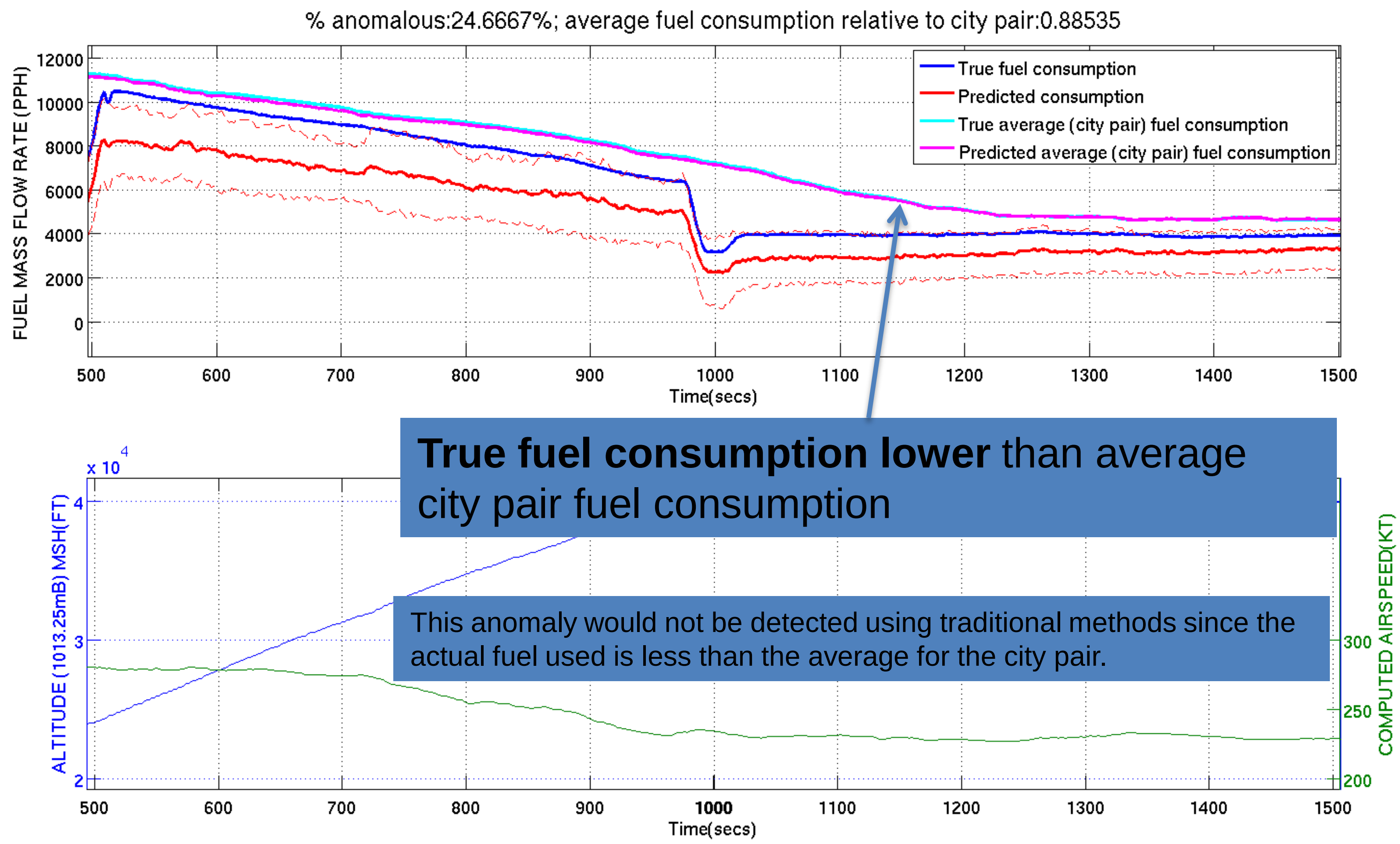
Anomalous flight - fuel consumption falls above prediction bounds.



True fuel consumption higher than average city pair fuel consumption

This anomaly may be detected using traditional methods since the actual fuel used is more than the average for the city pair.

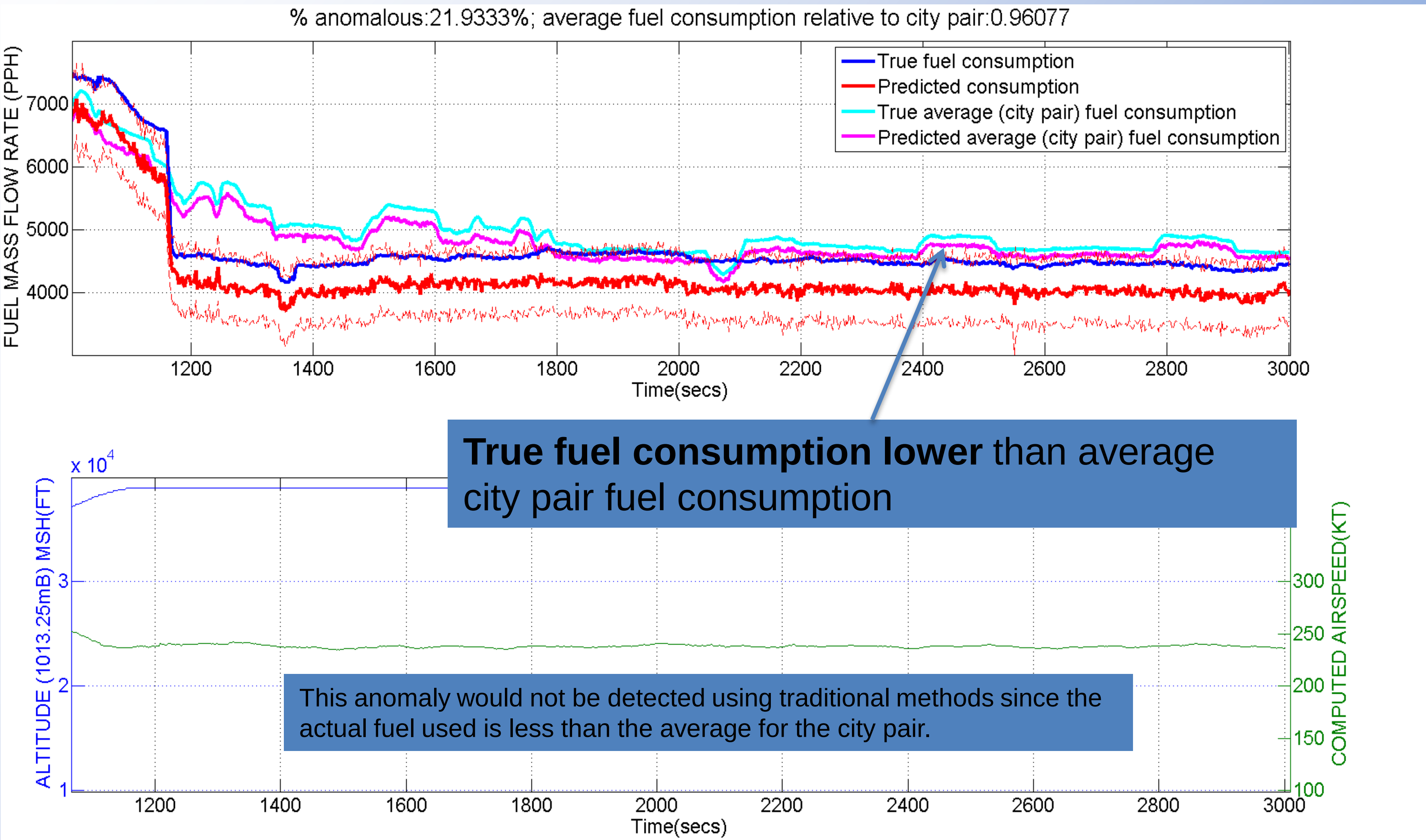
Anomalous flight - fuel consumption falls above prediction bounds



True fuel consumption lower than average city pair fuel consumption

This anomaly would not be detected using traditional methods since the actual fuel used is less than the average for the city pair.

Anomalous flight - fuel consumption falls above prediction bounds

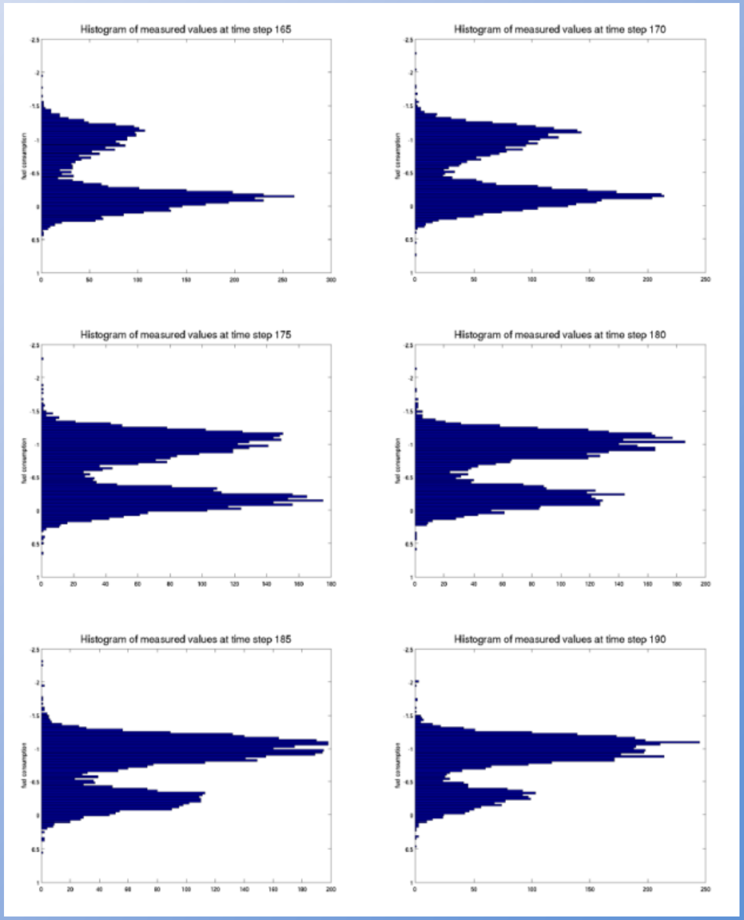


True fuel consumption lower than average city pair fuel consumption

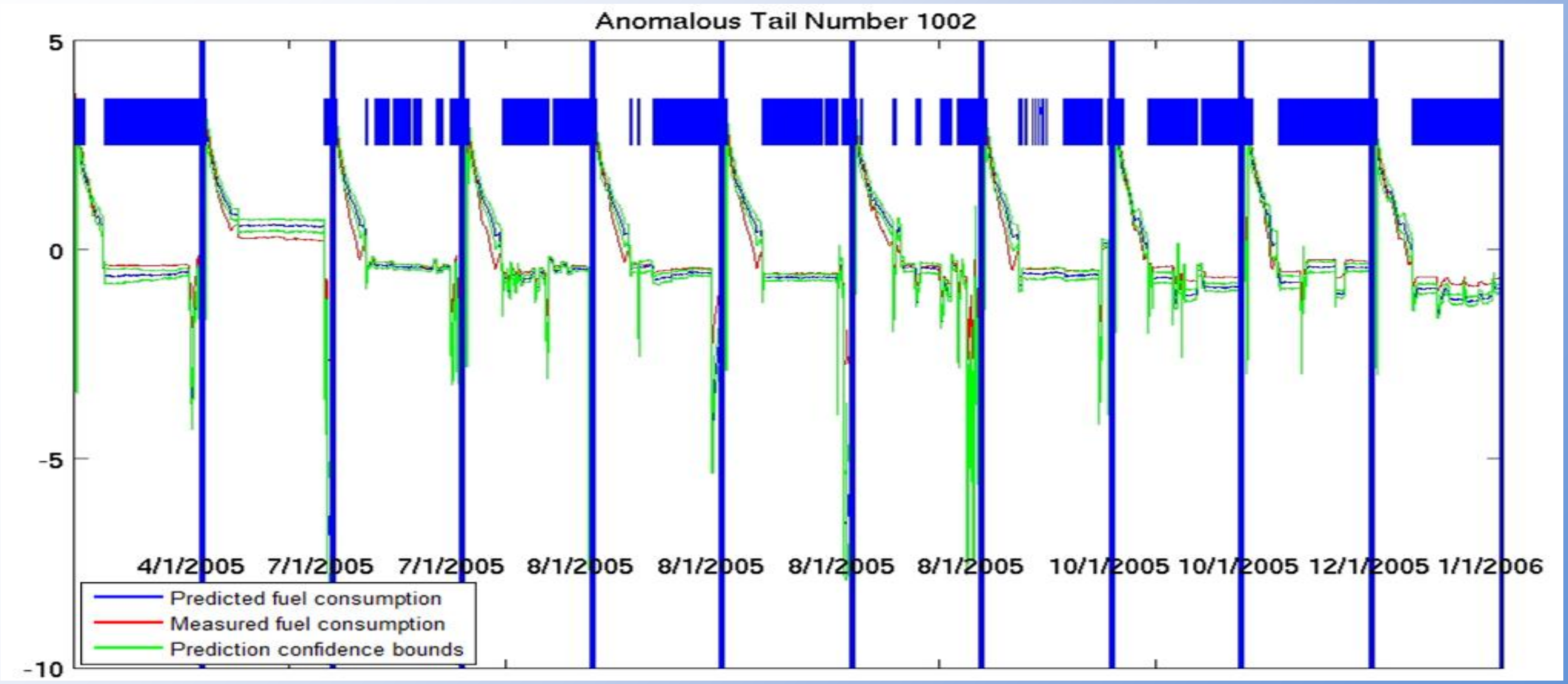
This anomaly would not be detected using traditional methods since the actual fuel used is less than the average for the city pair.

Impact

- Perhaps the first technology to detect fuel consumption anomalies as a function of aircraft state
- Findings may be used to take timely, corrective maintenance measures on fleet
- Under significant testing at Southwest Airlines today with strong support from the carrier.



VS can Discover Trends over Time



Discussion

- Fuel overconsumption may occur even if a flight uses less fuel than expected based on average city pair usage.
- A flight can be classified as **nominal** even though it consumes **more** fuel than average for a city pair.
- A flight can be classified as **anomalous** even though it consumes **less** fuel than average for a city pair.
- Virtual Sensors can explain about 90% of the variance in fuel consumption

Algorithm	stableGP	GLM	nnet	gp	time-based
% high	0.000	0.061	0.008	2.514	**2.004**
% low	0.002	0.064	0.003	1.003	**6.443**
% ok	99.998	99.875	99.989	96.483	**91.553**
NRMSE	0.113	0.160	0.208	0.200	**0.327**

Dissemination

- Virtual Sensors technology is open sourced
- Published in a top journal in the field of data mining:
  - A. N. Srivastava, “Greener Aviation through Virtual Sensors: A Case Study”, Data Mining and Knowledge Discovery, Volume 24 Issue 2, March 2012
- Under review at AIAA Infotech
- In preparation for Conference for Intelligent Data Understanding (in preparation)

Data Min Knowl Disc (2012) 24:443–471
DOI 10.1007/s10618-011-0240-z

Greener aviation with virtual sensors: a case study

Ashok N. Srivastava

Received: 4 June 2010 / Accepted: 21 September 2011 / Published online: 28 October 2011

Abstract The environmental impact of aviation is enormous given the fact that in the US alone there are nearly 6 million flights per year of commercial aircraft. This situation has driven numerous policy and procedural measures to help develop environmentally friendly technologies which are safe and affordable and reduce the environmental impact of aviation. However, many of these technologies require significant initial investment in newer aircraft fleets and modifications to existing regulations which are both long and costly enterprises. We propose to use an anomaly detection method based on Virtual Sensors to help detect overconsumption of fuel in aircraft which relies only on the data recorded during flight of most existing commercial aircraft, thus significantly reducing the cost and complexity of implementing this method. The Virtual Sensors developed here are ensemble-learning regression models for detecting the overconsumption of fuel based on instantaneous measurements of the aircraft state. This approach requires no additional information about standard operating procedures or other encoded domain knowledge. We present experimental results on three data sets and compare five different Virtual Sensors algorithms. The first two data sets are publicly available and consist of a simulated data set from a flight simulator and a real-world turbine disk. We show the ability to detect anomalies with high accuracy on these data sets. These sets contain seeded faults, meaning that they have been deliberately injected into the system. The second data set is from real-world fleet of 84 jet aircraft where we show the ability to detect fuel overconsumption which can have a significant environmental and economic impact. To the best of our knowledge, this is the first study of its kind in the aviation domain.

Responsible editor: Katharina Morik, Kanishka Bhaduri and Hillol Kargupta.

A. N. Srivastava (✉)
Intelligent Systems Division, Intelligent Data Understanding Group, NASA Ames Research Center, Moffett Field, CA 94035, USA
e-mail: ashok.n.srivastava@nasa.gov

Summary and Future Work

Summary

- Fuel consumption depends on state of aircraft(FOQA)
- Anomalous behavior should be determined **relative to** state of the aircraft
- Existing studies do not take this into account
- Proposed fuel study technique a novel effort in this direction
- Promising initial results for determining statistical anomalies

Future Work

- Improve the quality of fuel consumption model by leveraging new results in ensemble learning, robust convex optimization
- Further validation of results with Southwest Airlines.
- Determine whether the anomalies are actually operating in an off-nominal condition.
- Assess actionability and recommend best practices.