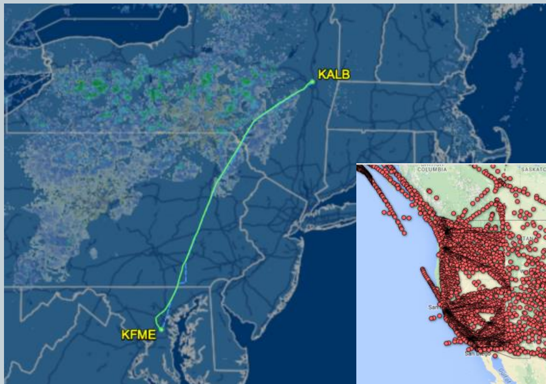
Steering Complex Systems using a Dynamic Data-Driven Modeling Approach

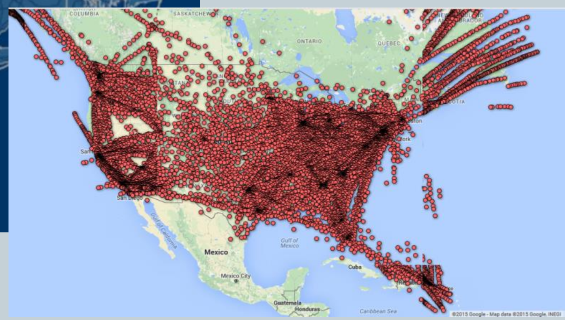
ALESSANDRO GALLI
SHIGERU IMAI
ERIKA MACKIN
NEIL McGLOHON
STACY PATTERSON
CARLOS A. VARELA
WENNAN ZHU

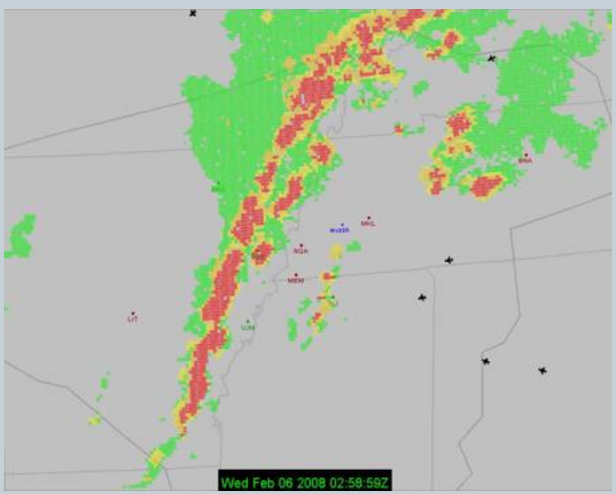
WORLDWIDE COMPUTING LABORATORY & NETWORKED SYSTEMS LABORATORY
DEPARTMENT OF COMPUTER SCIENCE
RENSSELAER POLYTECHNIC INSTITUTE

STREAM 2016 WORKSHOP
Tysons, VA, March 23, 2016

Traveling to STREAM 2016: Albany to DC







https://youtu.be/39eq5lgq9TA?t=1

Preferred Instrument Flight Rules (IFR) routes do not consider weather.

- Weather (clouds and potential icing conditions) initially forecast to be further west from "preferred IFR route". Actual weather was further east intersecting route.

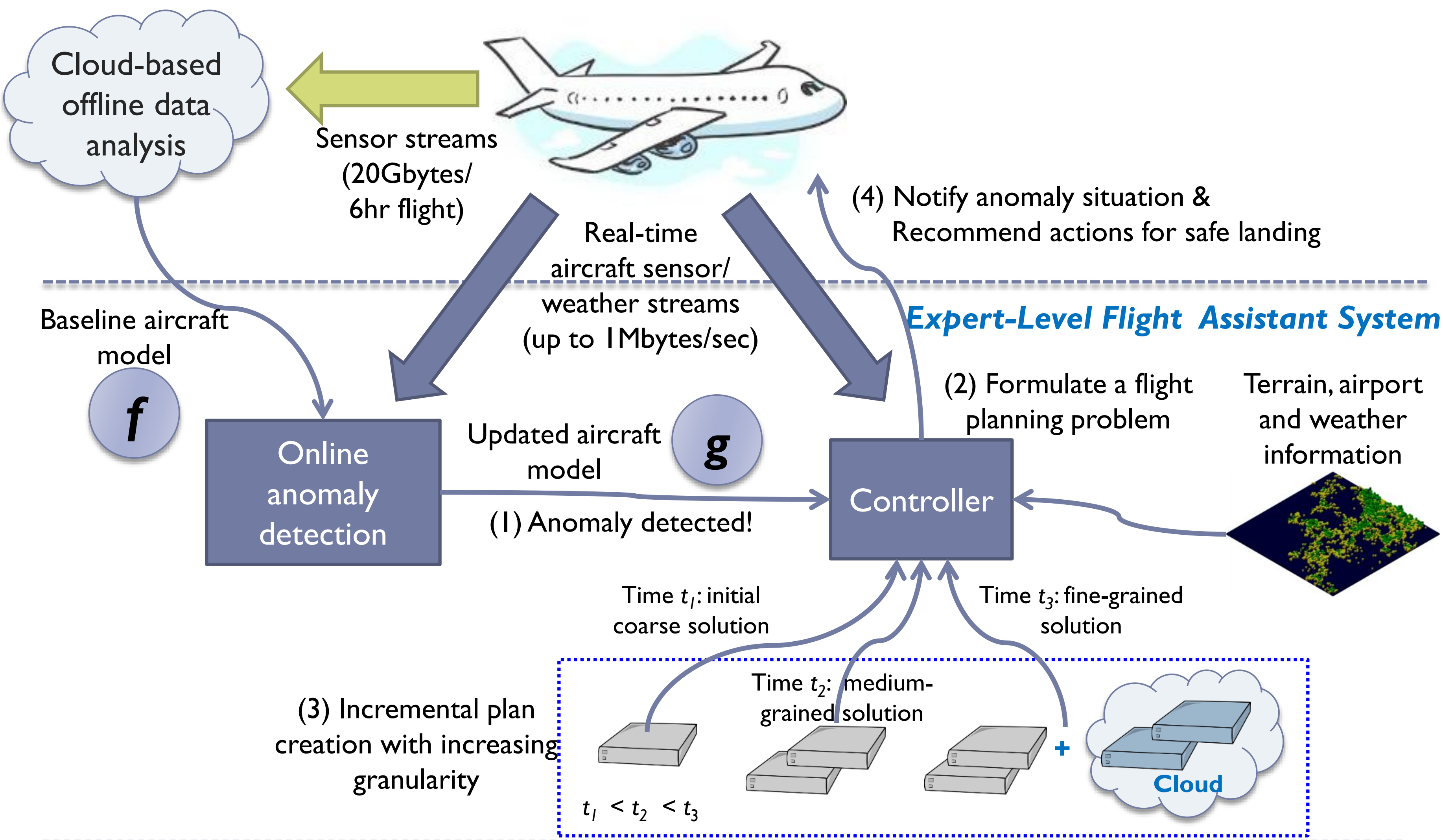
Air Traffic in the U.S.

- 87,000 flights per day (including private and commercial)
- Roughly 5,000 aircraft are flying at any given moment

Can air traffic autonomously avoid bad weather?

- while avoiding collisions, and
- staying within capacity constraints
- e.g., see FedEx Memphis hub operations during MidSouth storms and tornadoes.

Expert-Level Flight Assistant System



Air France Flight 447

- June 1st 2009, Flight 447 from Rio de Janeiro to Paris
- Thunderstorm caused airspeed sensors (*pitot tubes*) to ice and fail
- Autopilot system not able to deal with data failures---disengaged
- Pilots unable to react to erroneous data in a timely manner, eventually stalling the plane into the Atlantic Ocean

Figure 3: Pitot probe (with protection caps)

http://www.bea.aero/en/enquetes/flight.af.447/rapport.final.en.php

http://upload.wikimedia.org/wikipedia/commons/4/4a/Air_France_Flight_447_path.png

Dynamic Data-Driven Avionics

Using a data-driven feedback loop, DDDAS-based avionics continuously analyze spatio-temporal data streams from airplane sensors, identify potential failure modes, and correct erroneous data. Result is new layer of logical redundancy in addition to existing physical redundancy for safer flight systems.

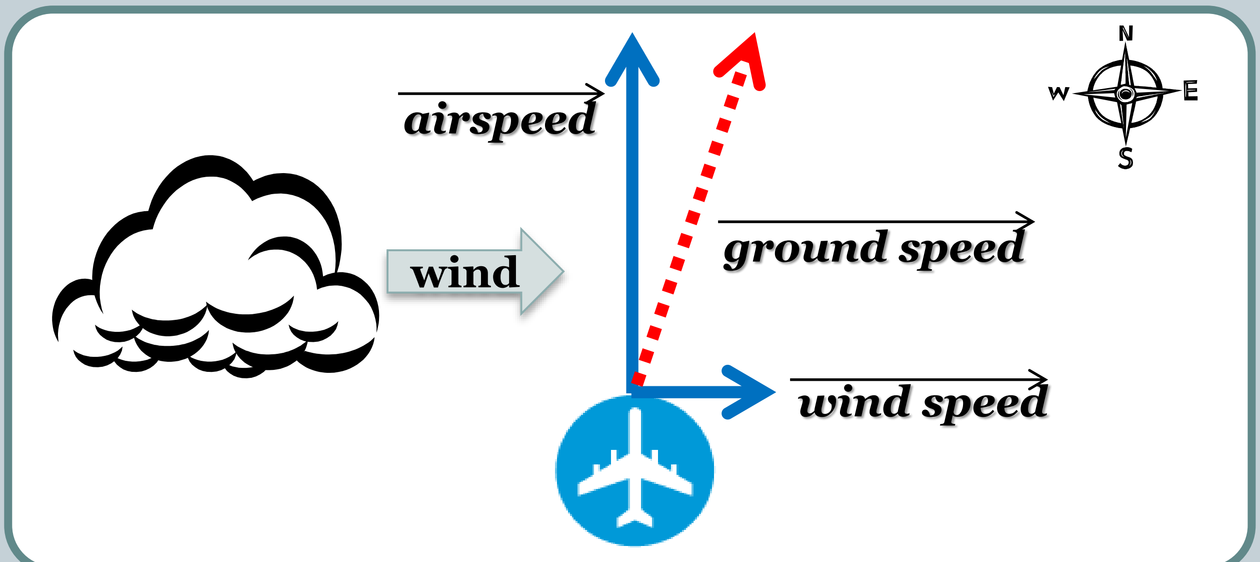
- **New mathematical concepts:**
  - *Error signatures:* Mathematical function patterns with constraints on specific data stream errors/anomalies.
  - *Mode likelihood vectors:* Stochastic selection of DDDAS system operation mode based on well-behaved sets of error signatures.

- **New DDDAS software: PILOTS programming language**
  - Enables declarative (high-level) definition of DDDAS data streaming application models (input-output relationships between data streams), error signatures, and error correction functions.
  - PILOTS software detects specific (e.g., failure-induced) data errors based on signatures and corrects data before processing according to the application model.

- **We have confirmed effectiveness of our approach using data from commercial flight accidents**
  - Air France AF447 accident in June 2009: Airspeed sensor failure of the AF447 flight successfully detected and corrected after 5 seconds from beginning of the failure. Overall error mode detection accuracy reaches 96.31%.
  - Tuninter 1153 accident in August 2005: The underweight condition due to the installation of an incorrect fuel sensor successfully detected with 100% accuracy during the cruise phase of flight.

Data Redundancy

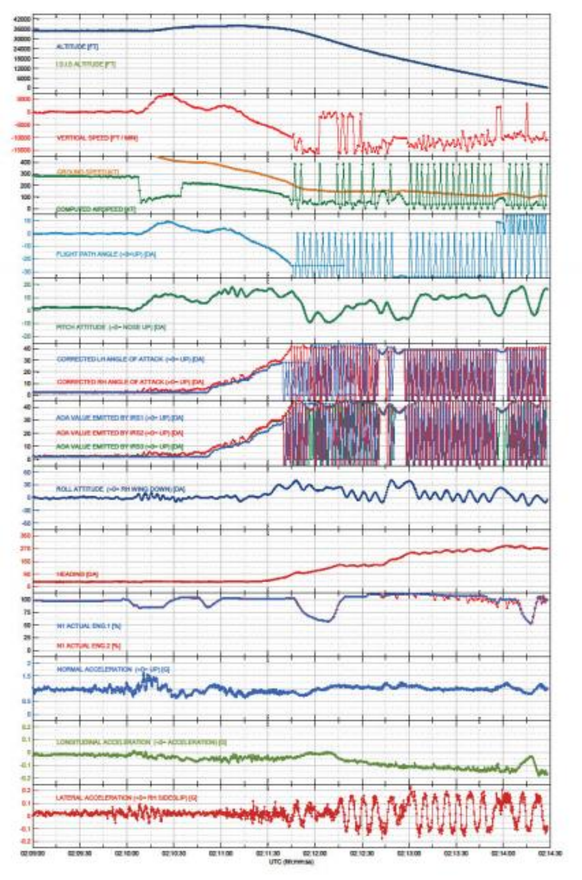
- **Primary cause of the AF447 accident: incorrect airspeed**
- **Airspeed could have been recomputed from ground speed and wind speed**
  - Take advantage of *data redundancy* between independently produced inputs



$$\overrightarrow{\textit{ground speed}} = \overrightarrow{\textit{airspeed}} + \overrightarrow{\textit{wind speed}}$$

Air France Flight 447

- **Data extracted from the final report of Air France Flight 447**
  - ***airspeed, air angle***: extracted from the graphs
    - **Real pitot tube failure is recorded**
  - ***ground speed, ground angle***: extracted from the graphs
  - ***wind speed, wind angle***: "the wind and temperature charts show that the average effective wind along the route can be estimated at approximately ***ten knots tail-wind."***
    - *wind speed* ← 10 knots
    - *wind angle* ← *air angle*

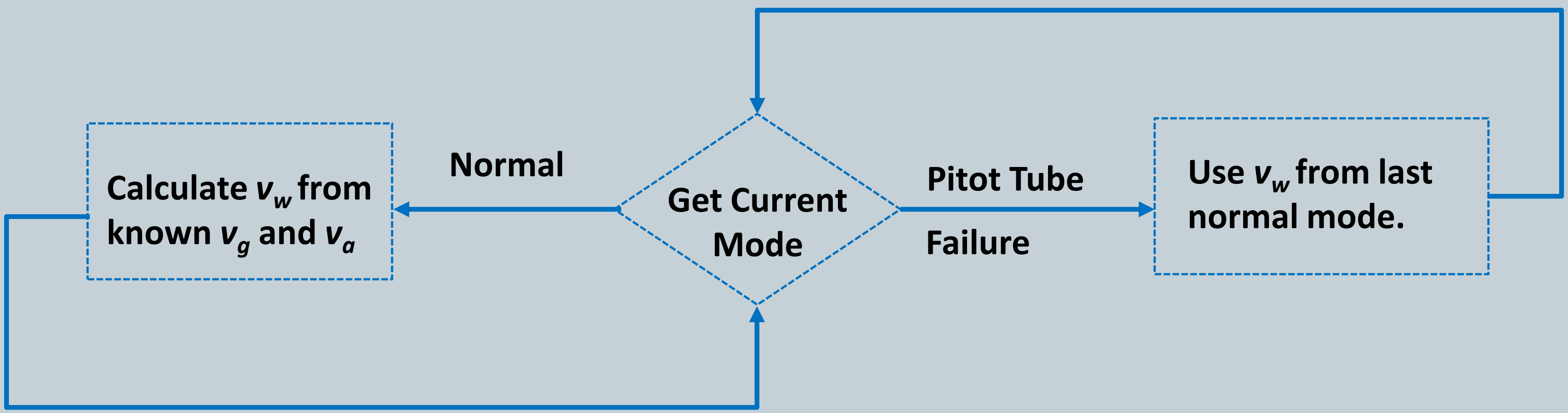


http://www.bea.aero/docspa/2009/f-cp090601.en/pdf/annexe.03.en.pdf

Air France AF447 PILOTS Demo

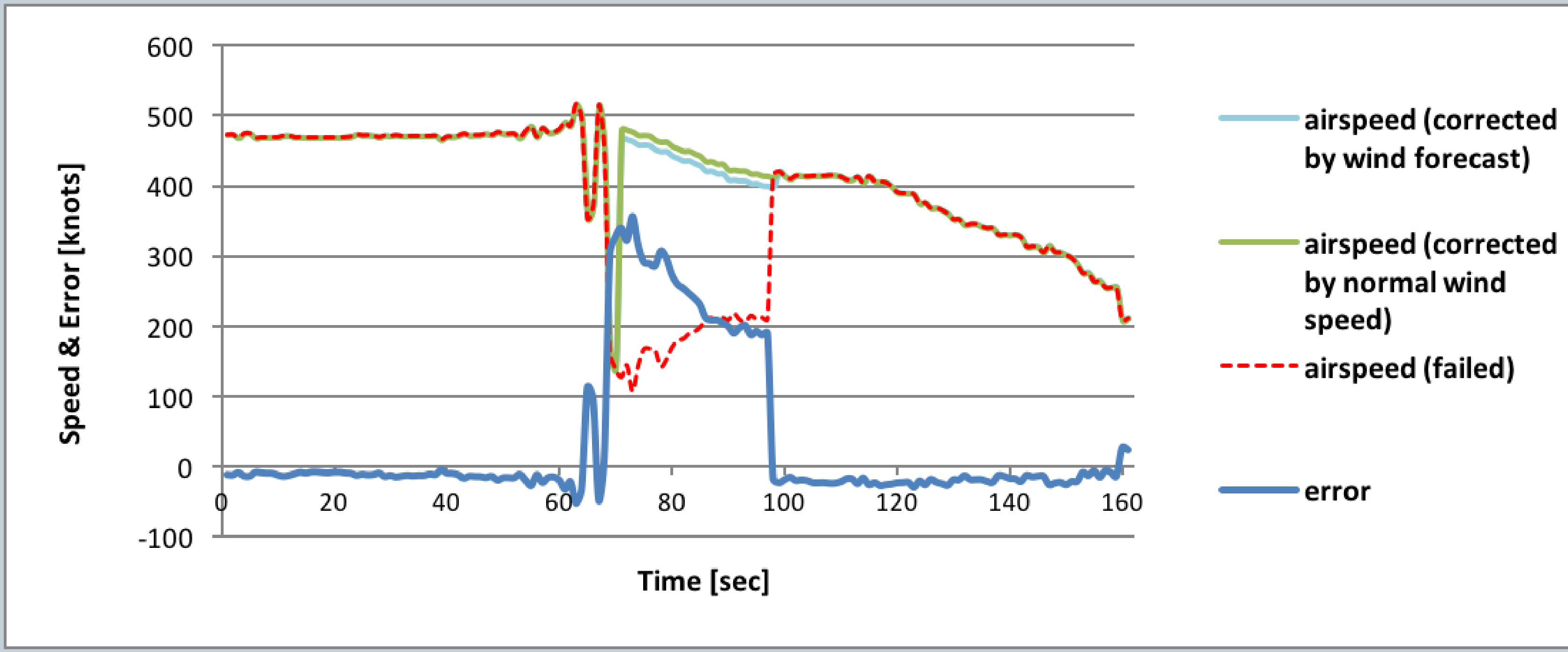
Wind Speed Estimation

- Calculate wind speed from ground speed and air speed in normal mode.
- When pitot tube fails, use wind speed from last normal mode calculation to correct air speed.

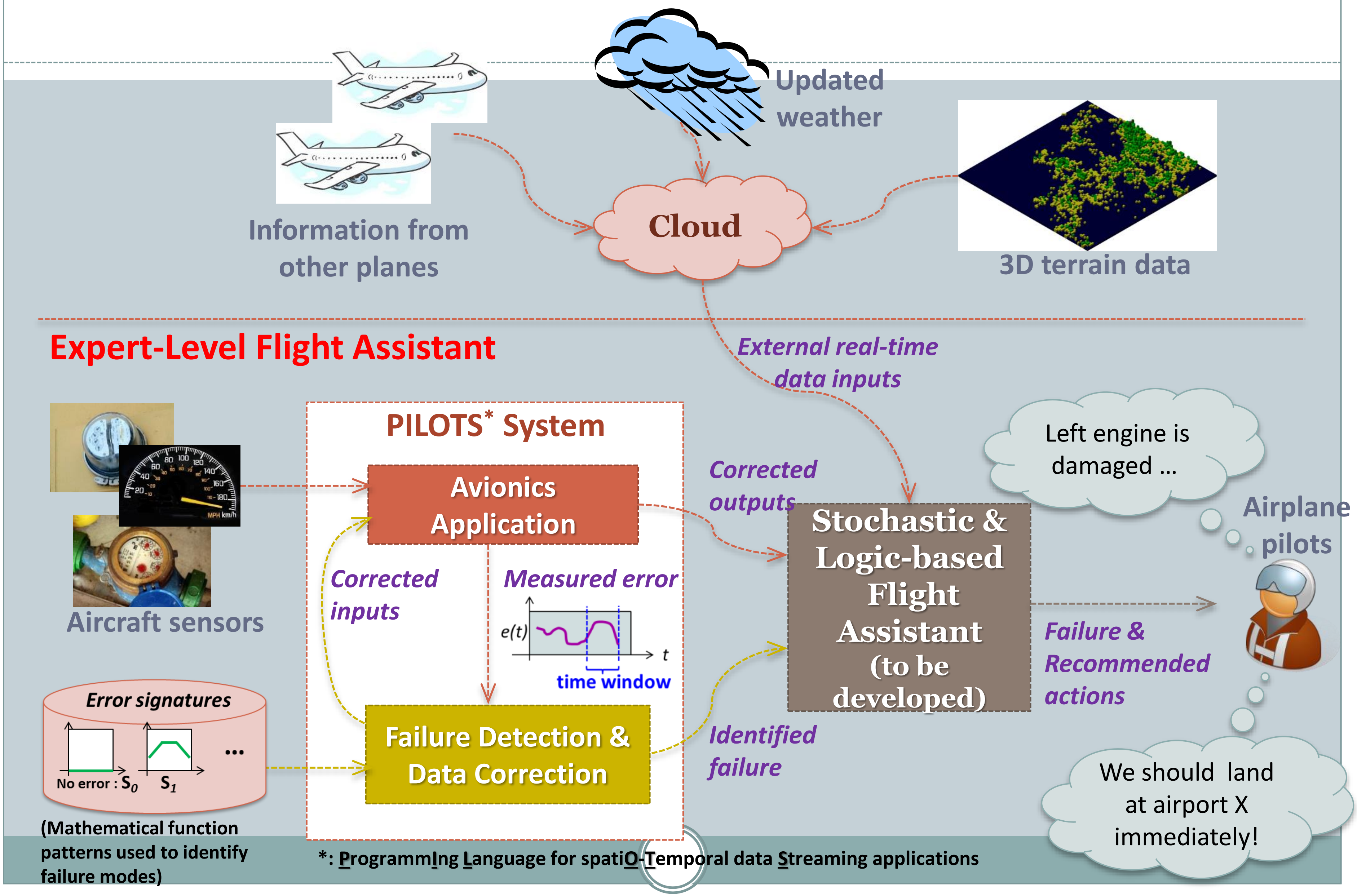


Wind Speed Estimation

- Air Speed Corrected by wind speed from weather forecast / the last normal mode.

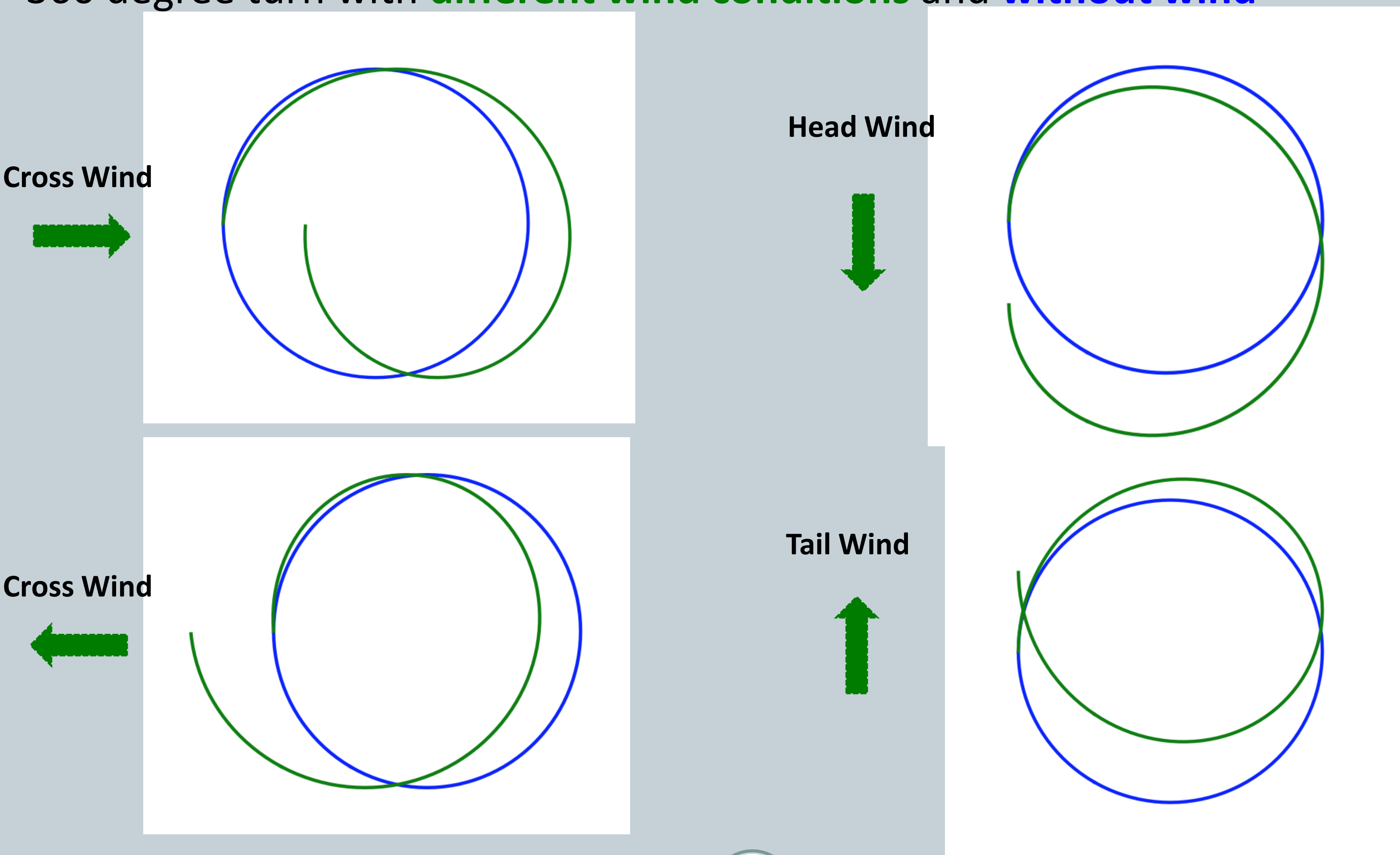


Multi-Aircraft Collaborative Flight Assistant System

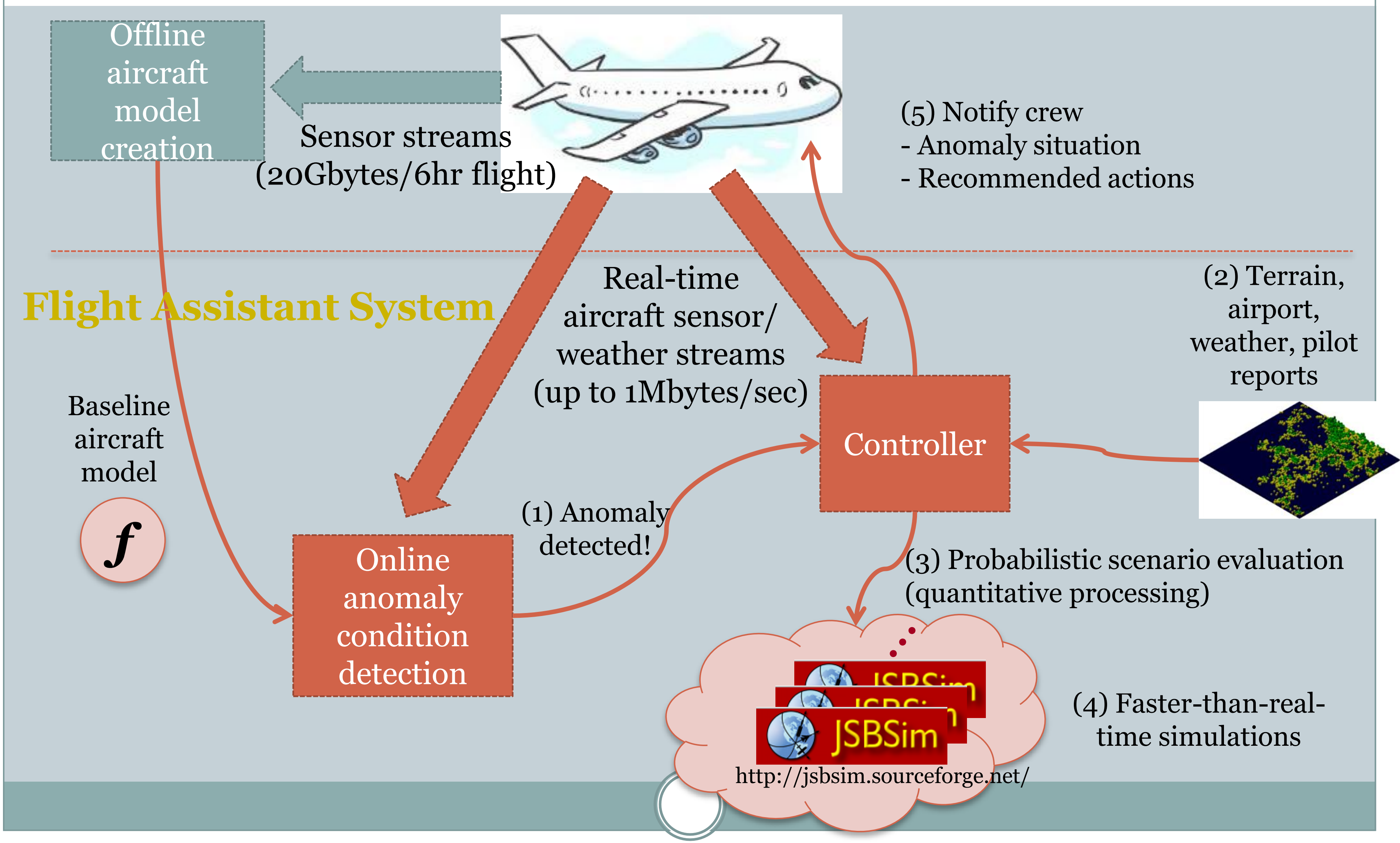


Steering aircraft to estimate wind speed

360 degree turn with **different wind conditions** and **without wind**



Aircraft Sensor Stream Processing for Expert-Level Flight Assistant System



Research Challenges (1)

- **Each participant has (spatial and temporal) quantitative model of system environment**
  - Some components computed offline, some online.
  - May be multiple contradictory models (e.g., weather models)
  - Should be able to create and modify plans based on logical inferences (rules for behaviors)

If…	then…
New pilot report: icing en route	New route
New winds aloft	New altitude
New surface winds at destination	New airport
Imminent engine failure	Nearest airport

Dynamic Data-Driven Flight Plan Adaptation Examples

Research Challenges (2)

- Each participant has a "view" of the ground truth
  - How to reconcile these multiple views efficiently?
  - Will have communication delays and failures
  - Bandwidth is limited
  - Example application: Next Generation Transportation system (ADS-B)
- There is uncertainty in these models
  - How can a participant quantify uncertainty?
  - How to use information propagation to reduce "cone of uncertainty"?
- How use steering to optimize a goal?
  - E.g., Information gathering to reduce uncertainty or gain knowledge
  - Example: determining wind speed with maneuvers

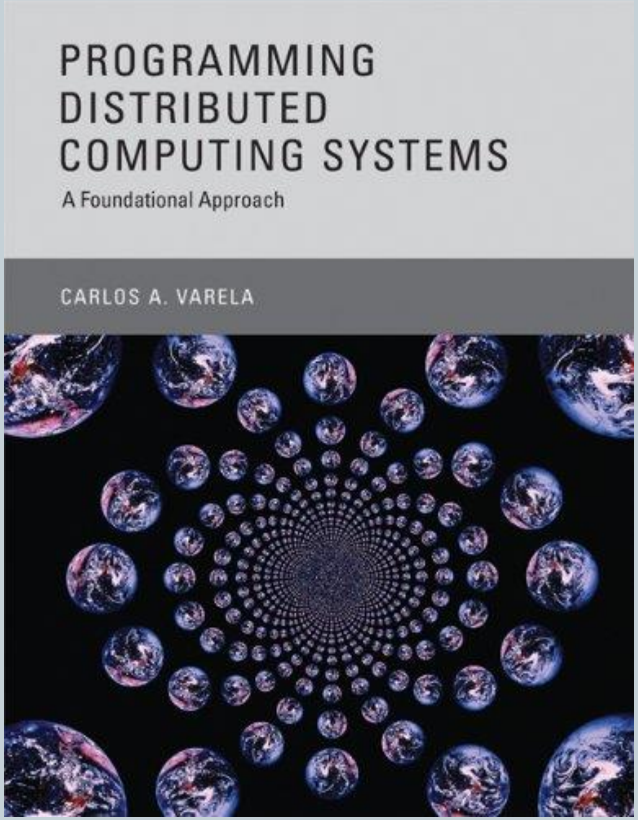
Research Challenges (3)

- **Need domain-specific languages and frameworks for data analytics**
  - **Easier data analyses, information generation, decision support.**
  - **Separation of concerns**
  - **Enables compiler (static) and middleware (dynamic) optimizations**
  - **First steps:**
    - PILOTS: ProgrammIng Language for SpatiO-Temporal data Streaming apps
    - Distill: A framework for distributed data analytics in the IoT

Questions?

- **Download open-source PILOTS 0.2.4 at:**
  http://wcl.cs.rpi.edu/pilots
- **Distill framework information at:**
  http://nsl.cs.rpi.edu/
- **Partial support from:**
  Air Force Office of Scientific Research
  DDDAS Program
  Dr. Frederica Darema
  (AFOSR Grant No. FA9550-15-1-0214),
  National Science Foundation
  CAREER Grant No. 1553340; EAGER/Dynamic Data Program Grant No. ECCS 1462342
  Yamada Corporation Fellowship

Consider textbook:



MIT Press, June 2013

Extra Slides

Dynamic Data-Driven Avionics Systems

- **To facilitate development of smarter (flight) data streaming systems, we investigate:**
  1. **Programming technology that can model spatio-temporal data streaming applications easily**
     - ***PILOTS*** (**P**rogramm**I**ng **L**anguage for spati**O**-**T**emporal data **S**treaming apps)
  2. **Error detection using *error signatures* and error correction based on *data redundancy***

Error Signatures

- **An *error signature* is a constrained mathematical function pattern defined as follows:**
  - $S(\bar{K}, f(t), \bar{P}(\bar{K})) = \{f(t) | p_1(\bar{K}) \wedge \cdots \wedge p_l(\bar{K})\}$ **where,**
    - f : a function of time
    - $\bar{K} = \langle k_1, \ldots, k_m \rangle$: a vector of constants
    - $\bar{P} = \{p_1(\bar{K}), \ldots, p_l(\bar{K})\}$: a set of constraint predicates

- **An *error signature sample* is a particular function in an error signature**
  - $s(t, \bar{K}) = f(t) \text{ s.t. } s(t, \bar{K}) \in S(\bar{K}, f(t), \bar{P}(\bar{K}))$

Mode Likelihood Vectors

- **Calculate the distance between measured error *e* and a signature S_i**

$$\delta_i(t) = \min_{\bar{K}} \int_{t-\omega}^{t} |e(t) - s_i(t, \bar{K})| dt.$$

- **Calculate the mode likelihood vector**

$L(t) =< l_0(t), l_1(t), \ldots, l_n(t) >$ where each $l_i(t)$ is defined as:

$$l_i(t) = \begin{cases} 1, & \text{if } \delta_i(t) = 0 \\ \frac{\min\{\delta_0(t), \ldots, \delta_n(t)\}}{\delta_i(t)}, & \text{otherwise.} \end{cases}$$

If 2nd greatest element of *L* is greater than significance threshold τ, error is *unknown*, else greatest element of *L* determines current error mode.

$\tau = 0.70$

$L = <0.3, 0.75, 1.0, 0.05>$

error mode = unknown

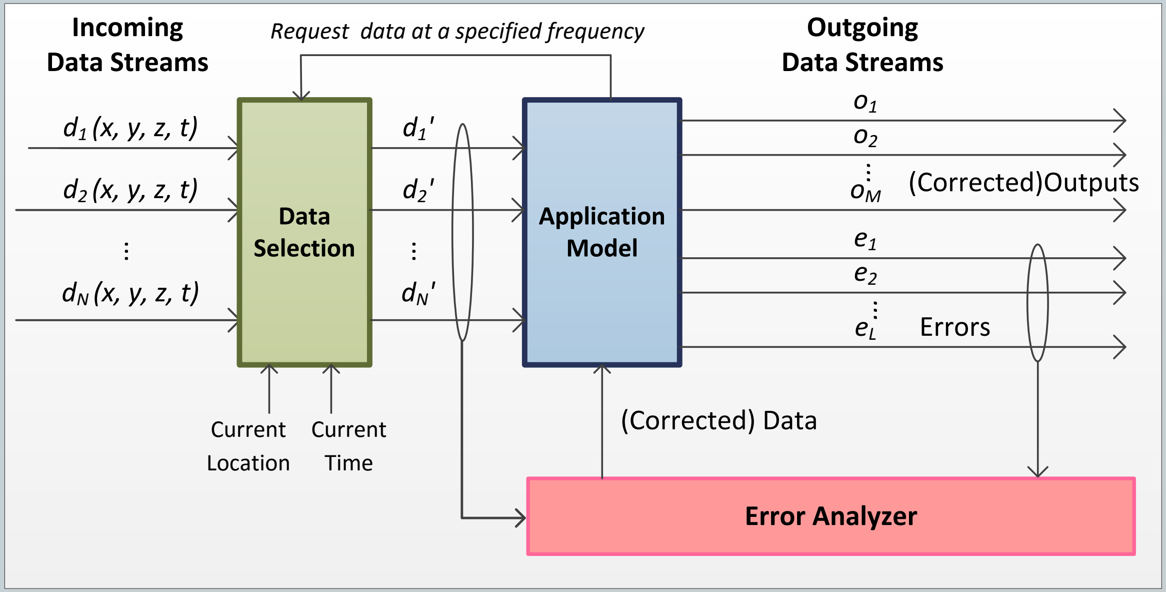
$\tau = 0.80$

$L = <0.3, 0.75, 1.0, 0.05>$

error mode = 2

PILOTS: System Architecture

- *Application Model*
  - Compute outputs and errors repeatedly
- *Data Selection*: from heterogeneous to homogeneous data
  - Selection operations to approximate data as a contiguous space
- *Error Analyzer*: error detection and correction



Tuninter 1153 Flight Accident

- Flight from Bari, Italy to Djerba, Tunisia on August 6th, 2005
- ATR-72 ditched into the Mediterranean sea
  - 16 of 39 people on board died

Figure 10: Chart showing the route of the TS-LBB as obtained from FDR data

"Final Accident Report for TS-LBB"
http://www.ansv.it/cgi-bin/eng/FINAL%20REPORT%20ATR%2072.pdf

http://www.airdisaster.com/photos/ts-lbb/5.shtml

["Mayday" TV Series on Tuninter 1153](https://youtu.be/aCrZwctnNWo?t=1904)
https://youtu.be/aCrZwctnNWo?t=1904

Initial Cause of the Accident

- **Incorrect fuel quantity indicator (FQI) installment**
  - **FQI for ATR-72 was not working properly (LED failure)**
  - **Technicians replaced the FQI with one designed for ATR-42**
    - FQI showed **2,700 kg** of fuel, but fuel actually weighed **550 kg**
    - Pilots did not realize data error eventually leading to fuel exhaustion

Photo 1: The FQI installed on TS-LBB before replacement.

"Final Accident Report for TS-LBB" http://www.ansv.it/cgi-bin/eng/FINAL%20REPORT%20ATR%2072.pdf

PILOTS Program

```
program WeightCheck;
/* v_a : airspeed , w: weight , h: altitude */
inputs
    v_a , w, h(t) using closest (t);
outputs
    corrected_w : w at every 10 sec;
errors
    e: v_a - (6.4869E+01 +
              1.4316E-02 * w +
              6.6730E-03 * h +
              (-3.7716E-07) * w * h +
              (-2.4208E-07) * w * w +
              (-1.1730E-07) * h * h) + 2.59;
signatures
    S0(K): e = K, -2 < K, K < 2     "Normal";
    S1(K): e = K, 4.69 < K          "Underweight";
correct
    S1: w = 3.34523E-12 *
            (sqrt(1.09278E+22 * h * h + (-1.65342E+27) * h +
                  (-3.69137E+29) * v_a + 1.01119E+32) -
             2.32868E+11 * h + 8.83906E+15);
end
```

$$e = v_a - \widehat{v_a} - K_{calib},$$

where $\widehat{v_a}$ is:

$$\hat{v_a}(t_\Delta = 10) = \mathbf{z}^{\mathrm{T}} \cdot \mathbf{k} = \begin{bmatrix} 1 \\ w \\ h \\ w \cdot h \\ w^2 \\ h^2 \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} 6.4869\mathrm{E}+01 \\ 1.4316\mathrm{E}-02 \\ 6.6730\mathrm{E}-03 \\ -3.7716\mathrm{E}-07 \\ -2.4208\mathrm{E}-07 \\ -1.1730\mathrm{E}-07 \end{bmatrix}$$

4.69 corresponds to **10%** discrepancy in weight

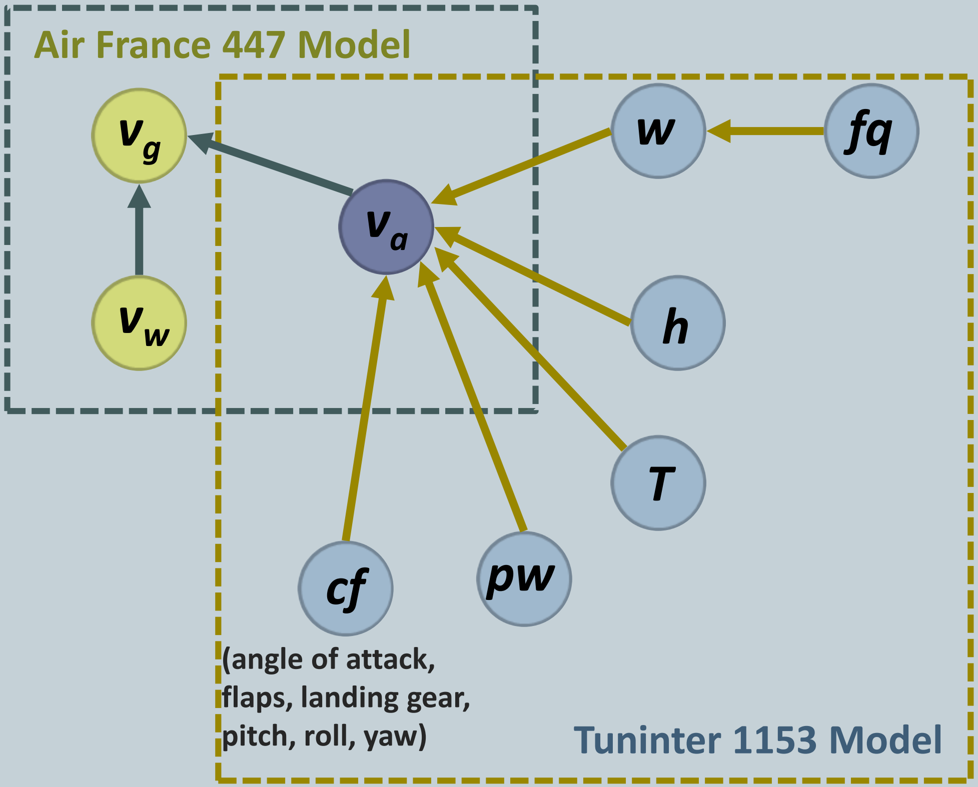
$$\hat{w} = \frac{-b - \sqrt{b^2 - 4ac}}{2a},$$

where $a = \mathbf{k}(5)$, $b = \mathbf{k}(2) + h \cdot \mathbf{k}(4)$,

$c = \mathbf{k}(1) + h \cdot \mathbf{k}(3) + h^2 \cdot \mathbf{k}(6) + K_{calib} - v_a$.

Complex Dependencies Between Data Streams

v_g : ground speed
v_w : wind speed
v_a : airspeed



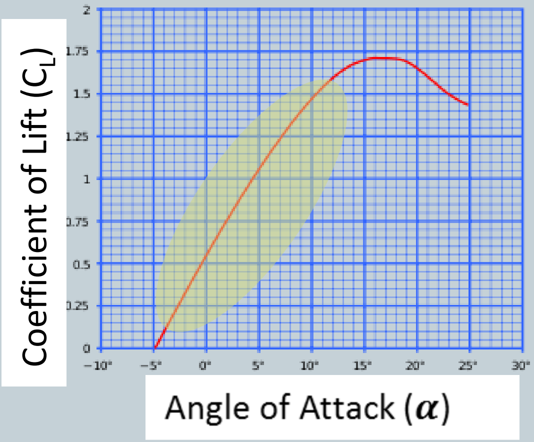
fq : fuel quantity
w : aircraft weight
h : altitude
T : temperature
pw : engine power
cf : aircraft configuration

Physics-based Models Parameter Learning

- **Model improvement for the Tuninter accident**
  - Revisit aerodynamics theory

$$w = K_1 \cdot \left(1 - \frac{\gamma}{T_0} \cdot c_{ft \to m} \cdot h\right)^{\frac{gM}{RL}} \cdot \alpha \cdot {v_a}^2 + K_2$$

    - Known constants : $\gamma, T_0, c_{ft \to m}, g, M, R$
    - From data : h, α, v_a, w
    - From linear regression : K_1, K_2
  - Assuming cruise flight
    - $w = L$
    - $\alpha \propto C_L$



http://upload.wikimedia.org/wikipedia/commons/thumb/d/d1/Lift_curve.svg/300px-Lift_curve.svg.png

Towards a Data-Driven Failure Model Learning Toolkit

- **Expand PILOTS language into a DDDAS Model Learning Toolkit to include:**
  - Montecarlo simulation to learn model parameters from data.
  - Kalman filters to reduce the impact of noise in data and enable more robust models.
  - Probabilistic (Bayesian) approach to continuously tune model to data.

Analysis of Flight Accidents and Possible Precautionary Measures

Flight	Date	Description	Precautionary Measures
Trans Asia Flight 235	February 4th 2015	2 minutes after takeoff, pilots report engine flameout. Right engine failure alert, warning sounds for 3 sec. Crew reduces and then cuts the left engine.	Decision support system to not turn off the left engine.
Asiana Airlines Flight 214	July 6th 2013	Descent below visual glide path and impact with seawall. 82 seconds before impact at 1,600 ft, autopilot was turned off and throttles set to idle. Final approach speed was 34 knots below the target approach speed of 137 knots. Pilots unaware that the auto-throttle was failing to maintain that speed.	Internal glide-path assistance. Airspeed crosscheck.
Turkish Airlines Flight 522	February 25th 2009	Aircraft had an automated reaction which was triggered by a faulty radio altimeter. Auto-throttle decreased the engine power to idle during approach. Crew noticed too late. Although the pilots did try to hold the glide slope after increasing the throttle, the auto-throttle decreased it to idle again.	Sensing the altimeter error using crosschecks.

Imai, Blasch, Galli, Lee, Varela, "Airplane Flight Safety using Error-Tolerant Data Stream Processing", IEEE AESM, in revision after initial review.

Analysis of Flight Accidents and Possible Precautionary Measures (cont.)

Flight	Date	Description	Precautionary Measures
British Airways Flight 38	Jan. 17, 2008	Although aware of the outside temperature conditions being -65C to -74C, the crew simply did not monitor the temperature of the fuel, which was well below freezing point. A small quantity of water within the fuel did freeze, causing ice on the fuel lines, ultimately leading to fuel starvation near the final stages of approach.	Check for fuel temperature when outside air temperature outside normal range.
Azerbaijan Airlines Flight 217	Dec. 23, 2005	After climbing to 6,900 ft entered a descending spiral tightening from 500 m to 100 m. Absence of all three gyroscopes during the climb. Lack of pitch, roll, and heading performance.	Attitude indicator crosscheck. Re-create a virtual artificial horizon from non-gyroscopic data.
Air Midwest Flight 5481	January 8th 2003	Elevator range of motion cut to only 7 degrees out of the full 14. Stalled after take-off due to overloading and maintenance error.	Weight and systems check from sensors onboard before departure.
Austral Lineas Aereas Flight 2553	October 10th 1997	Pitot tube icing caused faulty airspeed readings. Pilots interpreted as a loss of engine power and added power. No improvement to airspeed, so they descended and increased the speed. Wing slats were torn off one wing and the plane became uncontrollable.	Airspeed crosscheck.

Data Generation for Different Failure Modes

- **Data generation from Precision Flight Control's CAT III Flight Simulator at RPI's Worldwide Computing Laboratory:**

US Airways Flight 1549

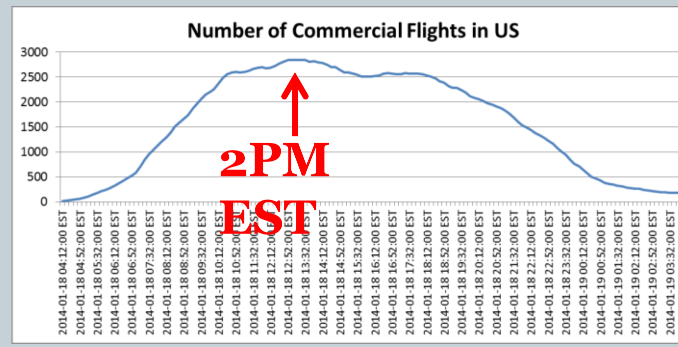
- **On January 15, 2009, US Airways Flight 1549 was struck by birds and lost thrust from both engines**
- **Captain Sullenberger successfully ditched the aircraft over the Hudson river without causing any loss of life**

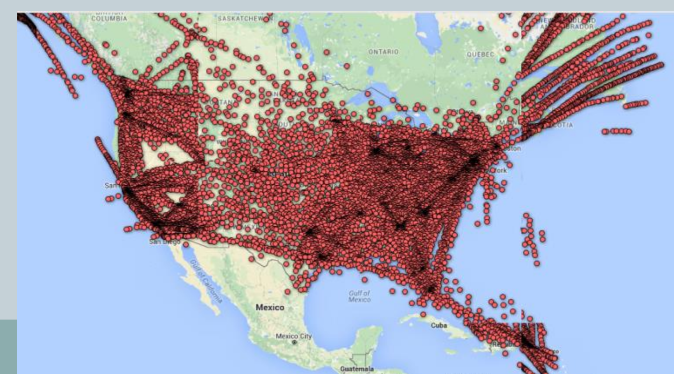
Map and picture are from Wikipedia (https://en.wikipedia.org/wiki/US_Airways_Flight_1549)

Aircraft Position Stream Processing for Efficient Air Traffic Management

- **Air Traffic in the U.S.**
  - 87,000 flights per day (including private and commercial)
  - Roughly 5,000 aircraft are flying at any given moment
  - Data rate for aircraft position and speed data streams:
    120 [bits/msg] * 1 [msg/sec] * 5,000 = 73 [KB/sec]

- **Air Traffic Management Problem**
  - *Objective*: minimize the total delay
  - Computationally expensive due to exponential number of combinations
  - Fluctuating computational demand
  - *Challenge*: How to timely finish the computation while keeping the monetary cost as low as possible?
    → ***Elastic stream processing*** in the cloud
  - *Imai, Patterson, and Varela,* "Elastic Virtual Machine Scheduling for Continuous Air Traffic Optimization," CCGrid, May 2016.

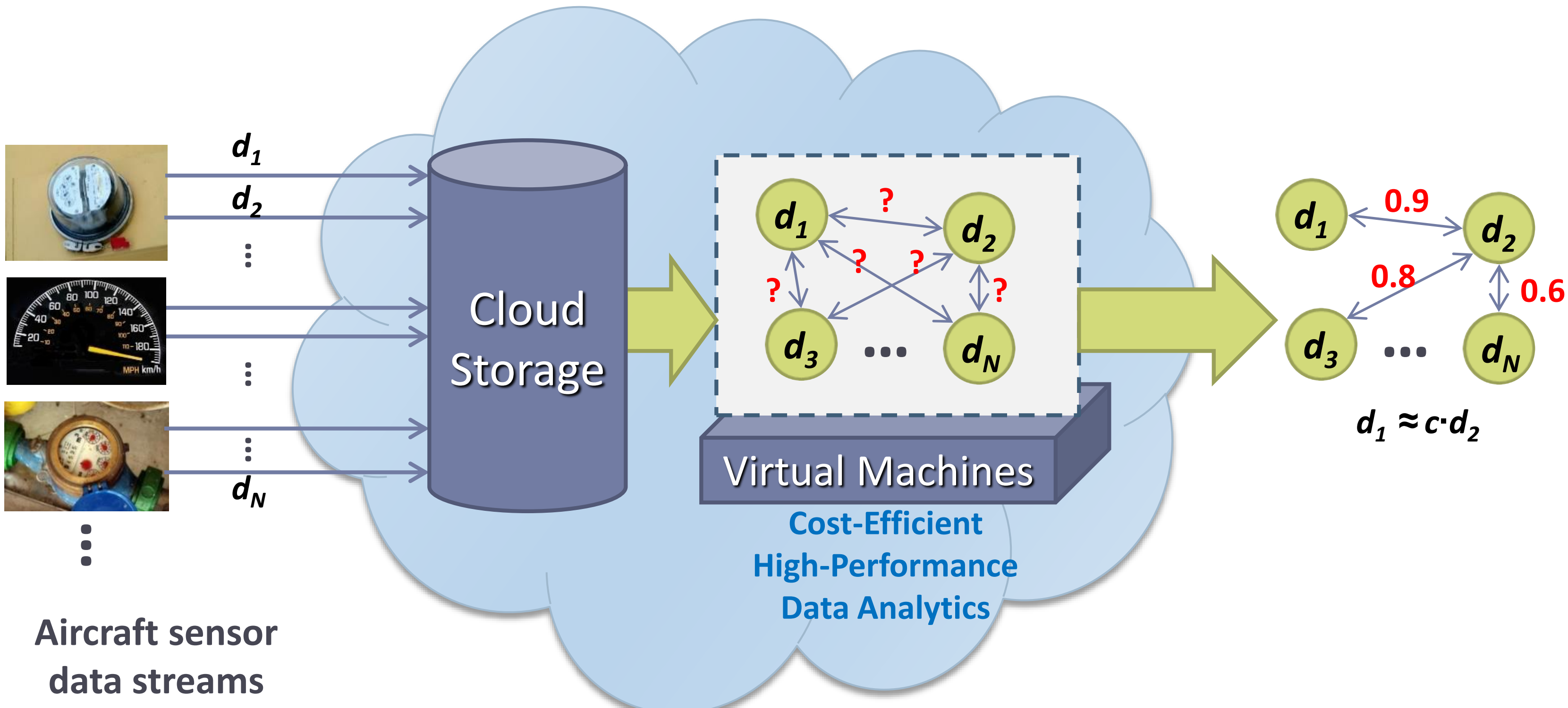




Cloud-based Offline Data Analytics

- **Scalable correlation analysis from hundreds of independently-measured sensor data streams**

  → Automating anomaly detection/correction model creation process



Research Challenges (1/4)

- **A quantitative spatial and temporal logic as a formalism:**
  - To enable reasoning about data streams that associate values to specific points or intervals of space and time.
  - To enable geometric reasoning capabilities, in particular, trigonometric formulae to calculate with aircraft speeds, headings, range, and endurance.

v	Speed (horizontal)
α	Direction
a	Aircraft
w,x	Wind, crosswind
r	Runway

Ground speed and crosswind as functions of airspeed, wind, and runway heading

$$v_g = v_a + \sin(\alpha_w - \alpha_a) \times v_w$$

$$v_x = \cos(\alpha_w - \alpha_r) \times v_w$$

Research Challenges (2/4)

- **Extensions to logic programming to support *stochastic reasoning.***
  - **Language extensions to standard Horn clause-based knowledge bases to incorporate probabilities.**
  - **Special language support for spatial and temporal data streams.**
  - **Incremental reasoning algorithms to dynamically re-compute logical queries efficiently as new data gets injected into the application.**

If…	then…
New pilot report: icing en route	New route
New winds aloft	New altitude
New surface winds at destination	New airport
Imminent engine failure	Nearest airport

Dynamic Data-Driven Flight Plan Adaptation Examples

Research Challenges (3/4)

- **Data streaming analytics in real-time using cloud computing**
  - **More data are expected to be available through the Internet and in-flight through Next Generation Transportation system (ADS-B by 2020).**
  - **Reason about spatial and temporal data in real-time**
    - Give pilots better information to make more accurate judgments during crucial emergency moments
  - **Offline and online components**
    - Analyzing key historical data and relatively static data (e.g., terrain, aircraft models) offline
    - Combining it with dynamic data (e.g., failure conditions, weather) for real-time decision making

Research Challenges (4/4)

- **Domain-specific programming languages are needed for data scientists**
  - Easier data analyses, information generation, decision support.
  - Separation of concerns
  - Enables compiler (static) and middleware (dynamic) optimizations

Related Work

Airspeed Estimation

- S. Hansen. and M. Blanke: Diagnosis of Airspeed Measurement Faults for Unmanned Aerial Vehicles. IEEE Transactions of Aerospace and Electronic Systems. Vol 50 (1), Jan. 2014, pp 224-239.

Wind Speed Estimation

- A.Cho, J.Kim, S.Lee, and C.Kee, Wind estimation and airspeed calibration using a UAV with a single-antenna GPS receiver and pitot tube, IEEE Transactions on Aerospace and Electronic Systems, vol.47, pp. 109--117, 2011.

Fault Detection and Isolation (FDI) for Aircraft

- J. Gertler: Designing dynamic consistency relations for fault detection and isolation. International Journal of Control. Vol. 73, Issue 8, 2000, pp 720-732
- L. Trav´e-Massuy`es, T. Escobet, X. Olive: Diagnosability analysis based on component supported analytical redundancy relations. IEEE Trans. on Systems, Man and Cybernetics, Part A : Systems and Humans 36(6), 1146–1160 (2006)
- M. L. Fravolini, V.Brunori, G.Campa, M.R. Napolitano, and M.La Cava: Structured analysis approach for the generation of structured residuals for aircraft FDI, IEEE Transactions on Aerospace and Electronic Systems, vol. 45 (4), pp. 1466--1482, 2009.
- H. Khorasgani, D. Jung, G. Biswas , E. Frisk, M. Krysander:Robust residual selection for fault detection. Proc. IEEE Conference on Decision and Control, 2015, pp 5764-5769.

Fault Detection and Isolation (FDI) for Aerospace Systems

- A.Zolghadri, Advanced model-based fdir techniques for aerospace systems: Today challenges and opportunities, Progress in Aerospace Sciences, vol.53, pp. 18--29, 2012.
- J.Marzat, H.Piet-Lahanier, F.Damongeot, and E.Walter, Model-based fault diagnosis for aerospace systems: a survey, Journal of Aerospace Engineering, vol. 226, no.10, pp. 1329--1360, 2012.