Design and Analysis of Safety Critical Systems

Peter Seiler and Bin Hu
Department of Aerospace Engineering & Mechanics
University of Minnesota

February 21, 2013

Outline

- Fly-by-wire overview and design challenges
  - Analytical redundancy is rarely used
  - Certification issues
- Analysis of analytical fault detection systems
  - Motivation for model-based fault detection and isolation (FDI)
  - Probabilistic systems analysis
  - Time-correlated residuals: Operator Power Iteration
- Conclusions and future work

Commercial Fly-by-Wire

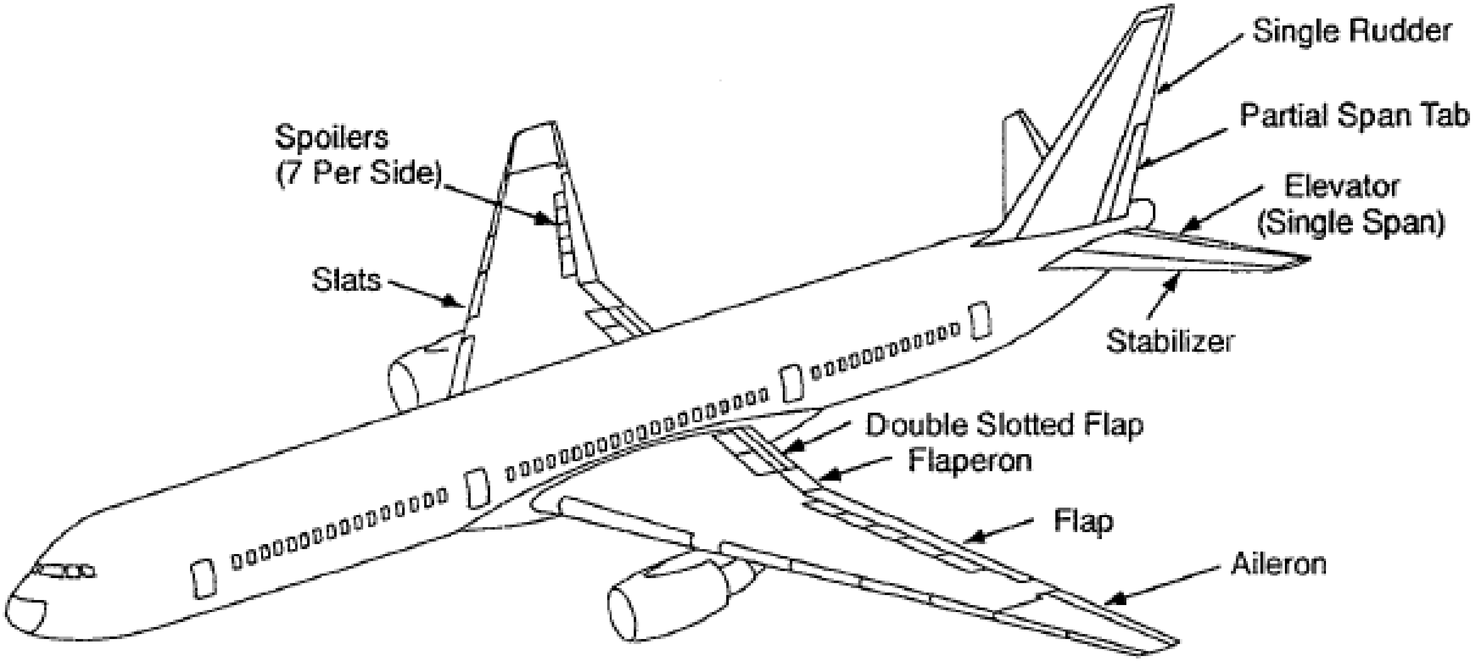
Boeing 787-8 Dreamliner

- 210-250 seats
- Length=56.7m, Wingspan=60.0m
- Range < 15200km, Speed< M0.89
- First Composite Airliner
- Honeywell Flight Control Electronics

Boeing 777-200

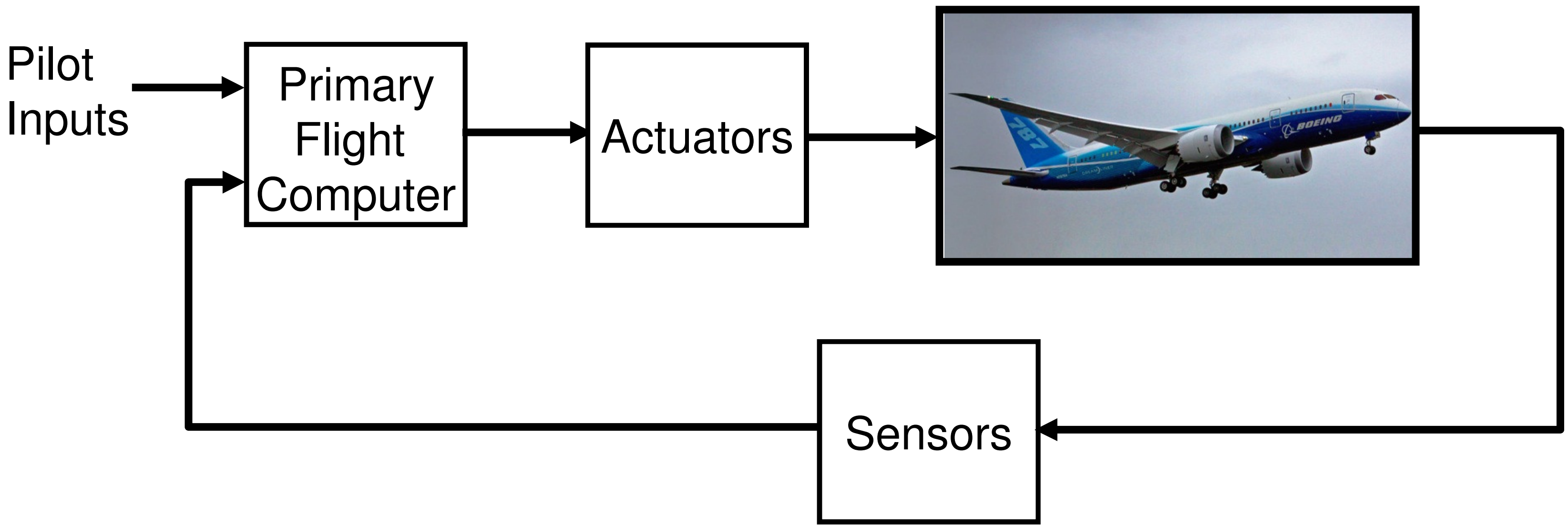
- 301-440 seats
- Length=63.7m, Wingspan=60.9m
- Range < 17370km, Speed< M0.89
- Boeing's 1st Fly-by-Wire Aircraft
- Ref: Y.C. Yeh, "Triple-triple redundant 777 primary flight computer," 1996.

777 Primary Flight Control Surfaces [Yeh, 96]



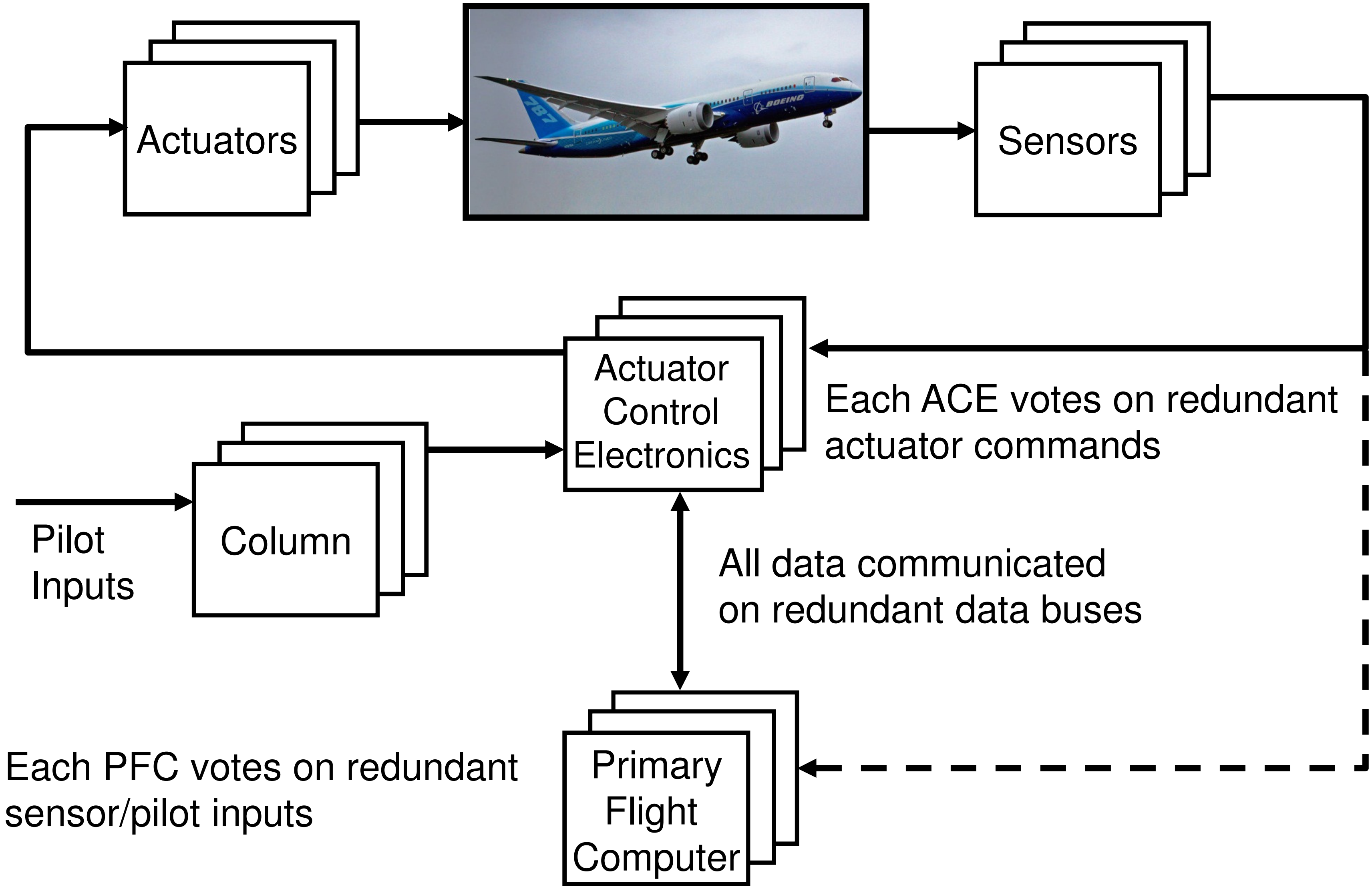
- Advantages of fly-by-wire:
  - Increased performance (e.g. reduced drag with smaller rudder), increased functionality (e.g. "soft" envelope protection), reduced weight, lower recurring costs, and possibility of sidesticks.
- Issues: Strict reliability requirements
  - $<10^{-9}$ catastrophic failures/hr
  - No single point of failure

Classical Feedback Diagram

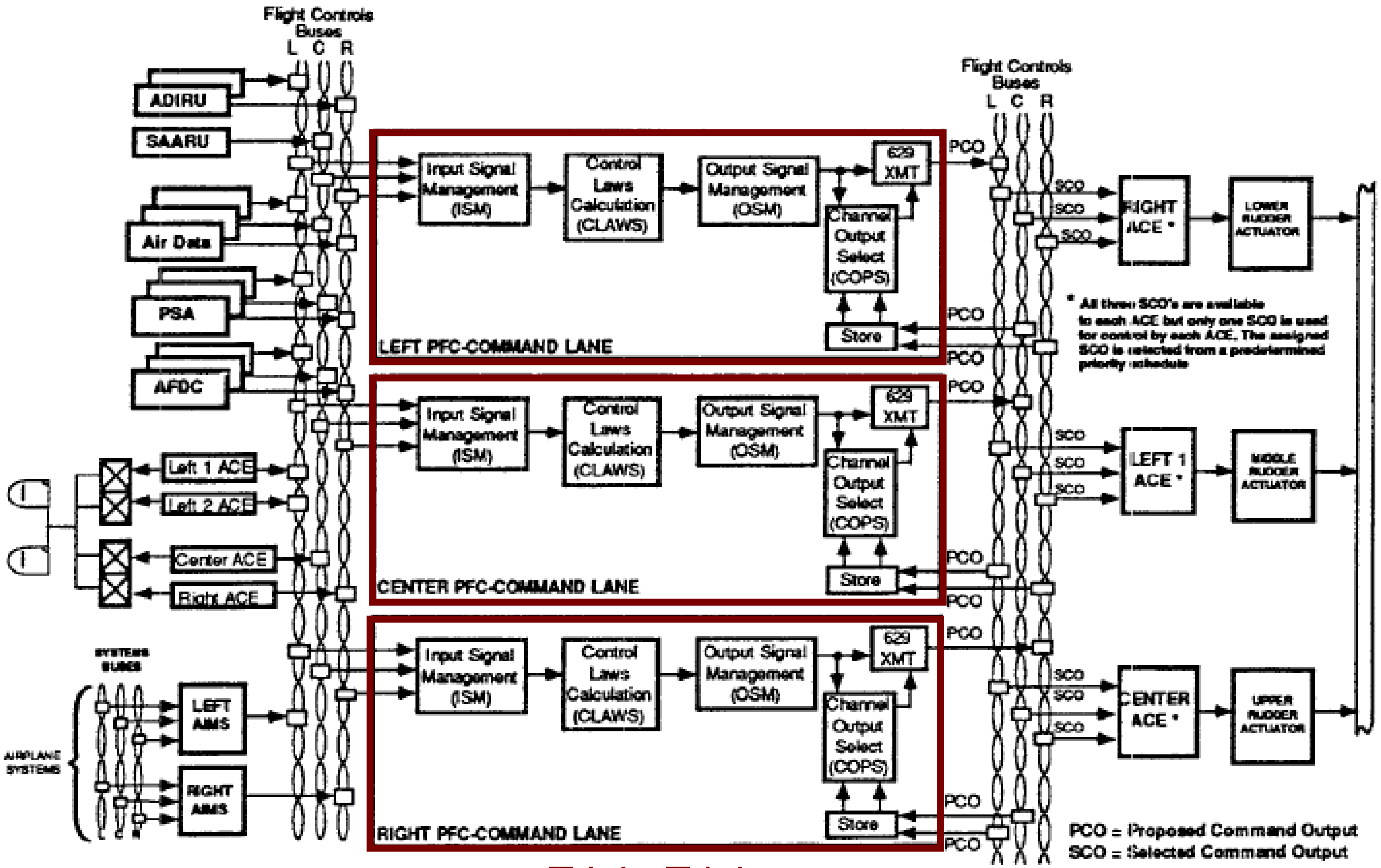


Reliable implementation of this classical feedback loop adds many layers of complexity.

Triplex Control System Architecture



777 Triple-Triple Architecture [Yeh, 96]



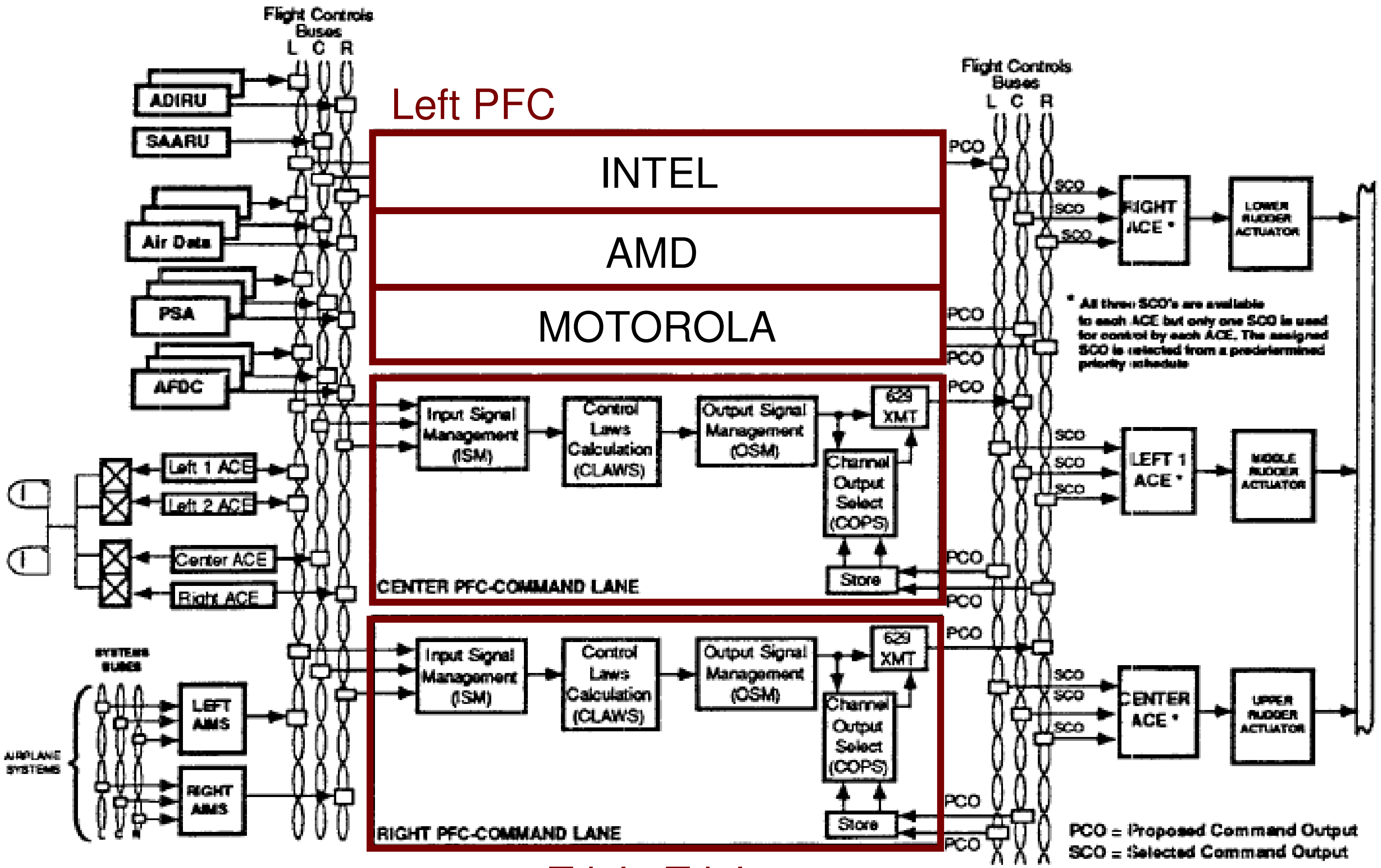
Sensors x3

Databus x3

Triple-Triple Primary Flight Computers

Actuator Electronics x4

777 Triple-Triple Architecture [Yeh, 96]



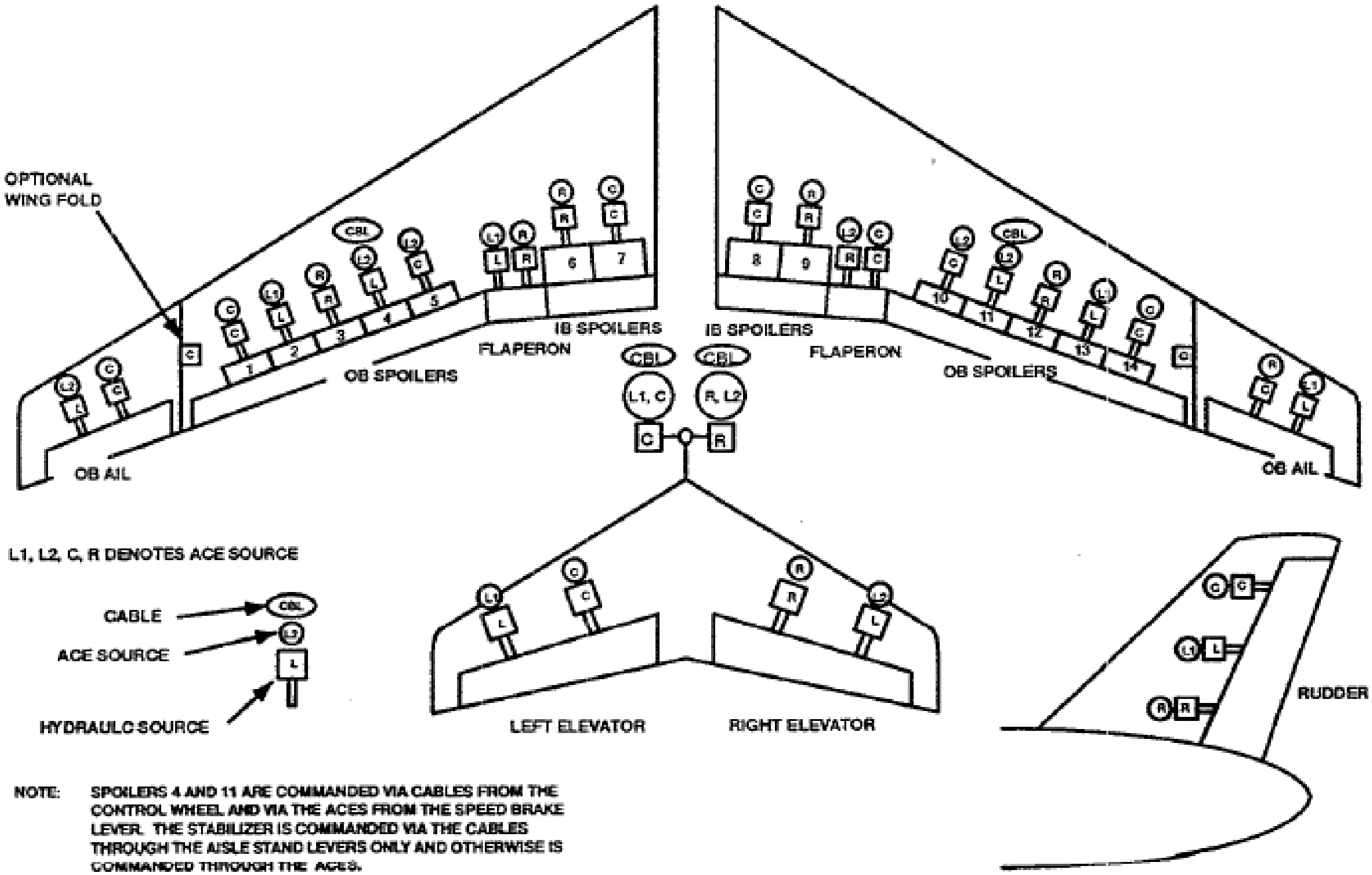
Sensors x3

Databus x3

Triple-Triple Primary Flight Computers

Actuator Electronics x4

Distribution of 777 Primary Actuators [Yeh, 96]



Degraded Modes [Yeh, 96]

Table 1 777 Primary Flight Control Modes

CONTROL MODE	PITCH	ROLL	YAW
NORMAL CONTROL	CONTROL C* Maneuver Cmd with Speed Feedback Manual Trim for Speed Variable Feel ENVELOPE PROTECTION Stall Overspeed AUTOPILOT Backdrive	CONTROL Surface Cmds Manual Trim Fixed Feel ENVELOPE PROTECTION Bank Angle AUTOPILOT Backdrive	CONTROL Surface Cmd Ratio Changer Wheel/Rudder Cross Tie Manual Trim Yaw Damping Fixed Feel Gust Suppression ENVELOPE PROTECTION Thrust Asymmetry Compensation AUTOPILOT Backdrive
SECONDARY CONTROL	CONTROL Surface Cmd (Augmented) Flaps Up/Down Gain Direct Stabilizer Trim Flaps Up/Down Feel	CONTROL Surface Cmd Manual Trim Fixed Feel	CONTROL Surface Cmds, Flaps Up/Down Gain PCU Pressure Reducer Manual Trim Fixed Feel Yaw Rate Damper (If Available)
DIRECT CONTROL	CONTROL Surface Cmd (Augmented) Flaps Up/Down Gain Direct Stabilizer Trim Flaps Up/Down Feel	CONTROL Surface Cmd Manual Trim Fixed Feel	CONTROL Surface Cmds, Flaps Up/Down Gain PCU Pressure Reducer Manual Trim Fixed Feel

Degraded functionality as system failures occur

Ram Air Turbine

Ram air turbine: F-105 (Left) and Boeing 757 (Right)
http://en.wikipedia.org/wiki/Ram_air_turbine

Redundancy Management

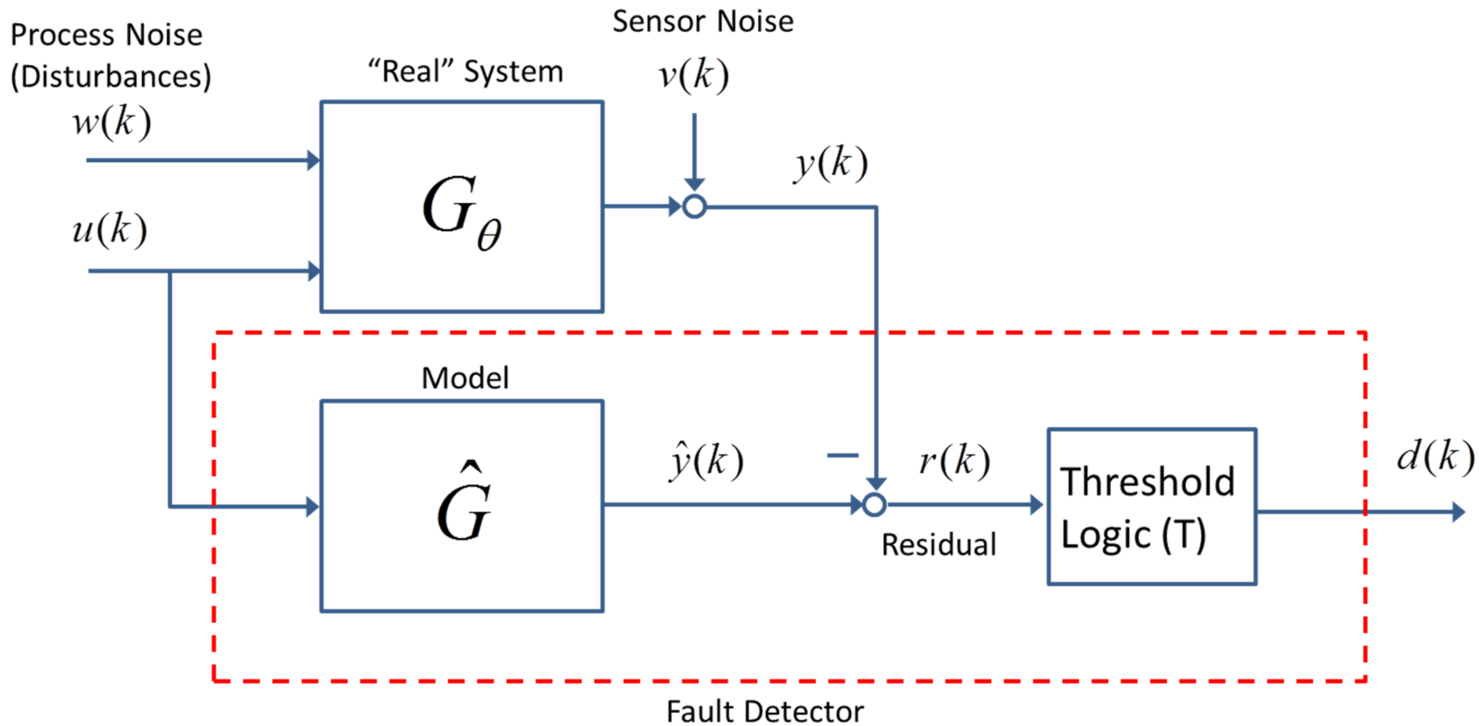
- Main Design Requirements:
  - < 10^{-9} catastrophic failures per hour
  - No single point of failure
  - Must protect against random and common-mode failures
- Basic Design Techniques
  - Hardware redundancy to protect against random failures
  - Dissimilar hardware / software to protect against common-mode failures
  - Voting: To choose between redundant sensor/actuator signals
  - Encryption: To prevent data corruption by failed components
  - Monitoring: Software/Hardware monitoring testing to detect latent faults
  - Operating Modes: Degraded modes to deal with failures
  - Equalization to handle unstable / marginally unstable control laws
  - Model-based design and implementation for software
- **Analytical redundancy is rarely used in commercial aircraft**

Outline

- Fly-by-wire overview and design challenges
  - Analytical redundancy is rarely used
  - Certification issues
- **Analysis of analytical fault detection systems**
  - Motivation for model-based fault detection and isolation (FDI)
  - Probabilistic systems analysis
  - Time-correlated residuals: Operator Power Iteration
- Conclusions and future work

Analytical Redundancy

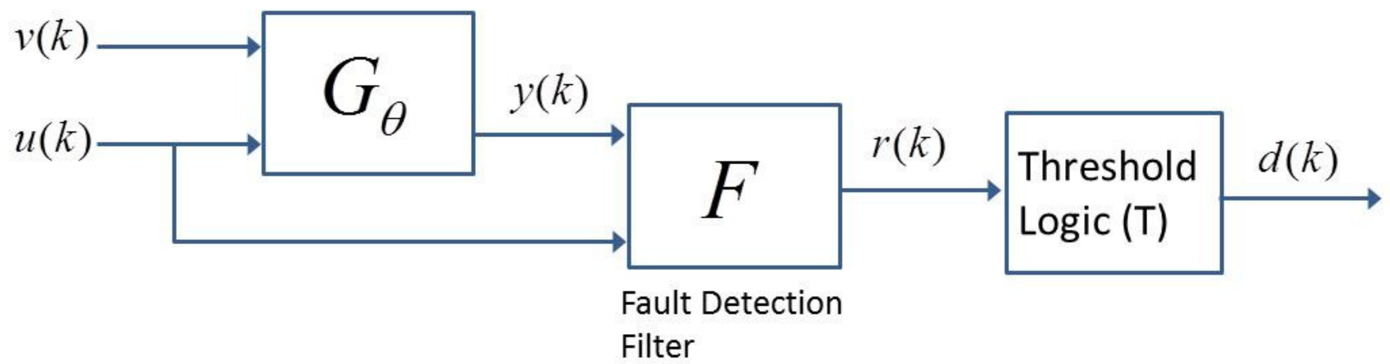
- Analytical Redundancy / Model-based Fault Detection
  - Use relations between disparate measurements to detect faults
  - Willsky, Ding, Chen, Patton, Isermann, others



Example: Parity-equation architecture

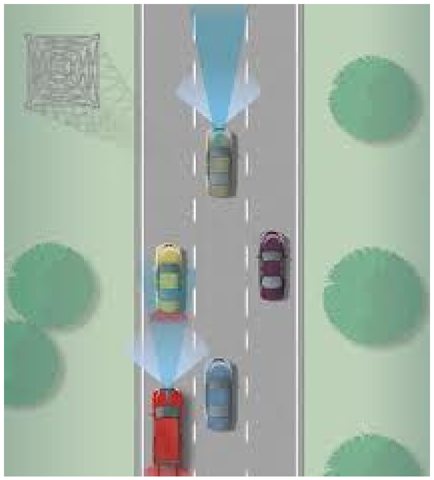
Analytical Redundancy

- Analytical Redundancy / Model-based Fault Detection
  - Use relations between disparate measurements to detect faults
  - Willsky, Ding, Chen, Patton, Isermann, others



Generic filter / threshold architecture

Motivation: Reduce Size, Weight, and Power



Automotive
Active Safety

NASA Crew
Exploration Vehicle

Unmanned Aerial
Vehicles

Many safety-critical applications can not support the high size, weight, power, and monetary costs associated with physical redundancy.

Model-based FDI for Safety Critical Applications

- FAA reauthorization requires a plan to certify UAVs for integration in the airspace by Sept. 30, 2015.
  - **Design**: Can high levels of reliability be achieved using analytical redundancy?
  - **Analysis**: How can analytically redundant systems be certified?

- Research
  - Design: Data-driven vs. model-based (Freeman, Balas)
  - Design: Robust fault detection (Vanek, Bokor, Balas)
  - Analysis: Probabilistic performance (Hu, Wheeler, Packard)

Model-based FDI for Safety Critical Applications

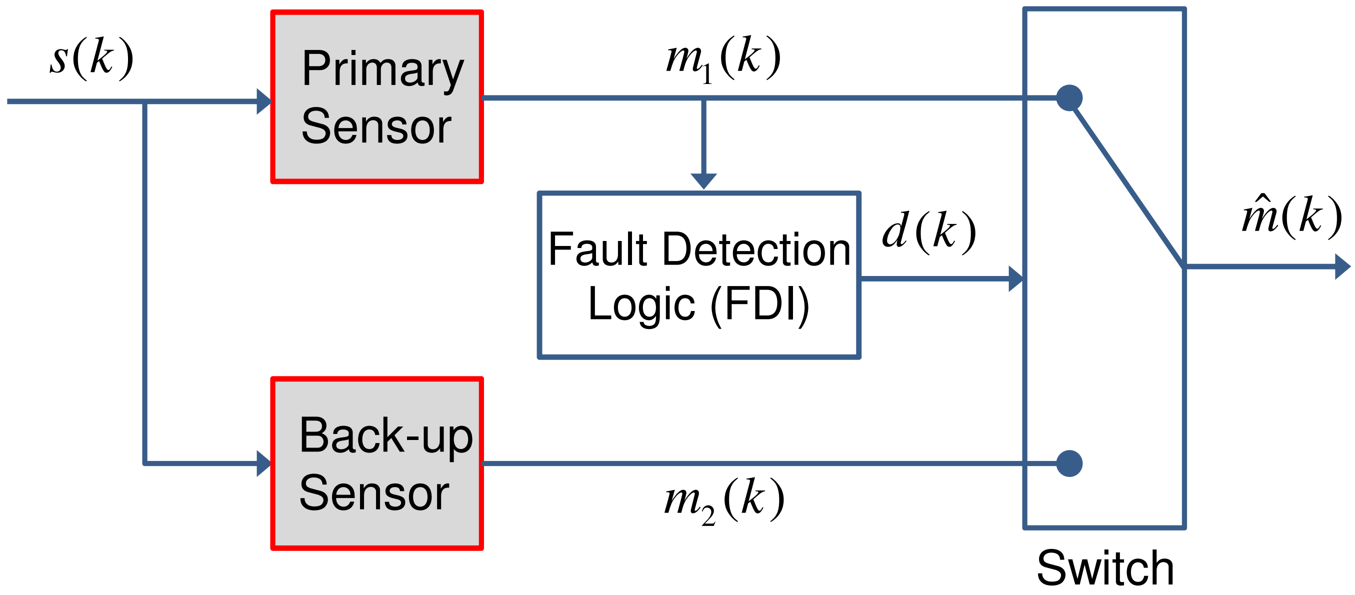
- FAA reauthorization requires a plan to certify UAVs for integration in the airspace by Sept. 30, 2015
  - **Design**: Can high levels of reliability be achieved using analytical redundancy?
  - **Analysis**: How can analytically redundant systems be certified?

- Research
  - Design: Data-driven vs. model-based (Freeman, Balas)
  - Design: Robust fault detection (Vanek, Bokor, Balas)
  - **Analysis: Probabilistic performance (Hu, Wheeler, Packard)**

Certification of Analytically Redundant Systems

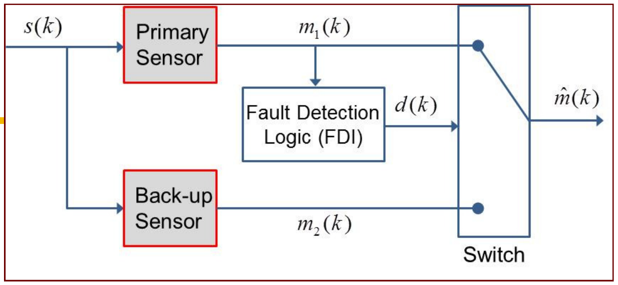
- Certification for physically redundant systems
  - Failure Modes and Effects Analysis
  - Fault Trees Analysis: Analyze system failure modes in terms of probabilities of lower-level events.
- Many issues for analytically redundant systems
  - Mixture of component and algorithm (HW+SW) failures
  - Nonlinear dynamics, model uncertainty, variation with flight condition
  - Correlated residuals
  - Strict reliability requirements
- **Proposed Approach**: Rigorous linear analysis at many flight conditions + nonlinear Monte Carlo simulations
  - Analogous procedure used to certify flight control laws

Dual-Redundant Architecture



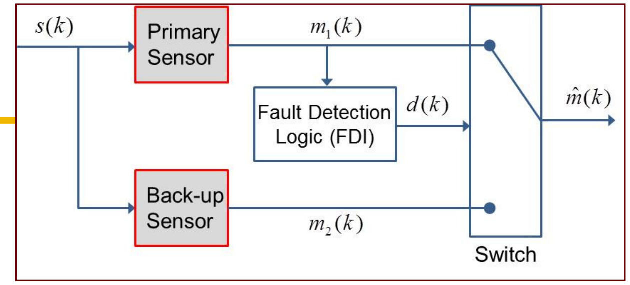
Objective: Efficiently compute the probability $P_{S,N}$ that the system generates "bad" data for N_0 consecutive steps in an N-step window.

Assumptions

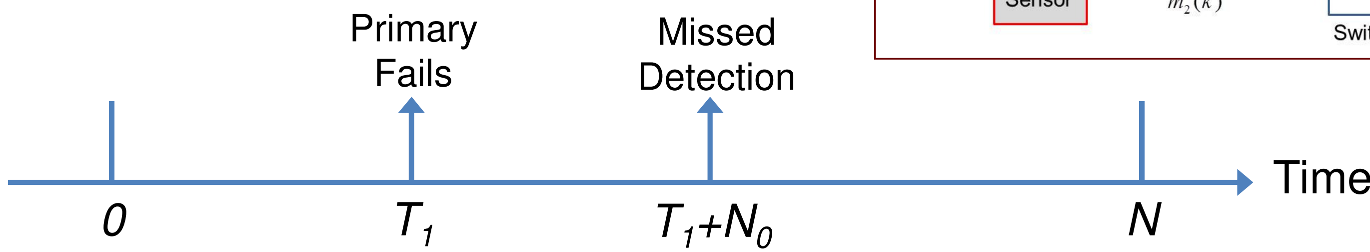


1. Knowledge of probabilistic performance
   a. Sensor failures: $P[\ T_i\!=\!k\ ]$ where T_i := failure time of sensor i
   b. FDI False Alarm: $P[\ T_S\!\leq\!N\ |\ T_1\!=\!N\!+\!1\ ]$
   c. FDI Missed Detection: $P[\ T_S\!\geq\!k\!+\!N_0\ |\ T_1\!=\!k\ ]$
2. Neglect intermittent failures
3. Neglect intermittent switching logic
4. Sensor failures and FDI logic decision are independent
   - Sensors have no common failure modes.

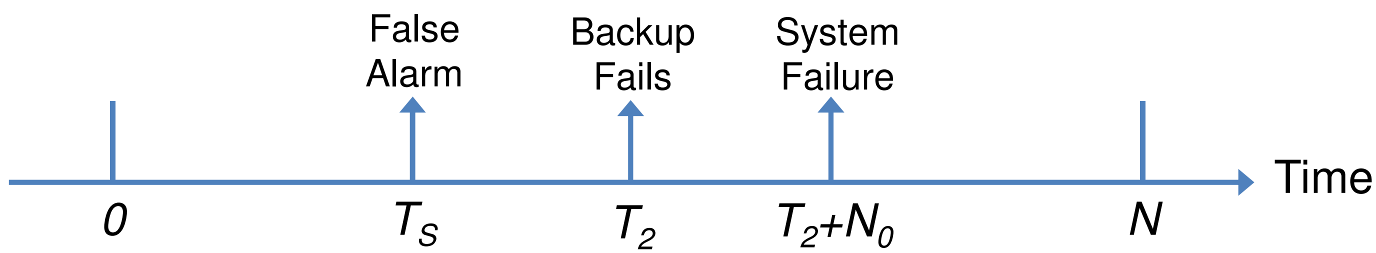
Failure Modes



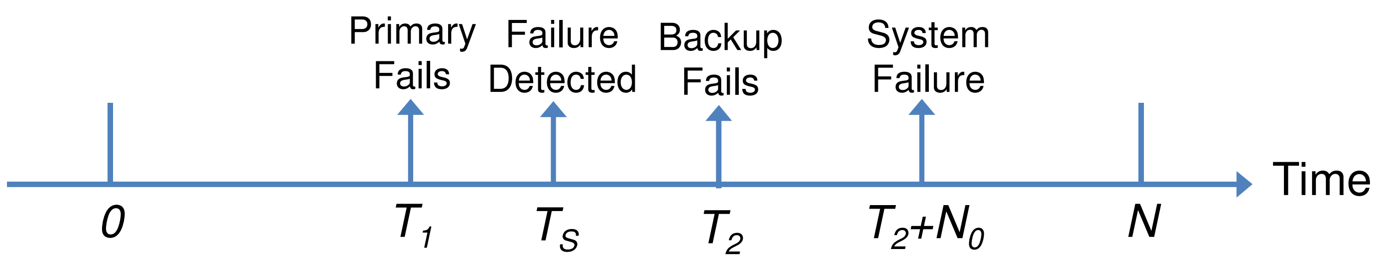
Missed Detection, M_N



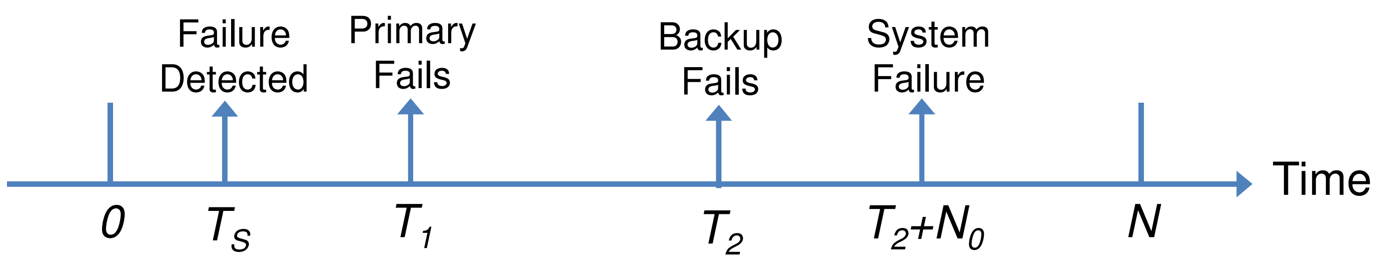
False Alarm, F_N



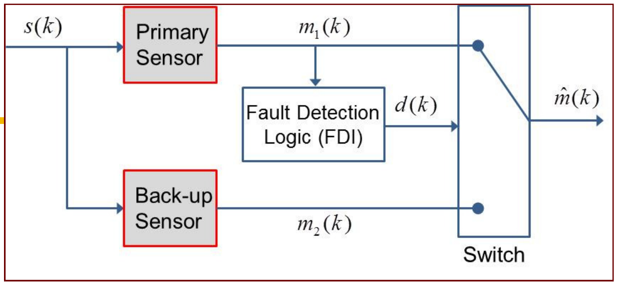
Proper Detection, D_N



Early False Alarm, E_N



System Failure Probability



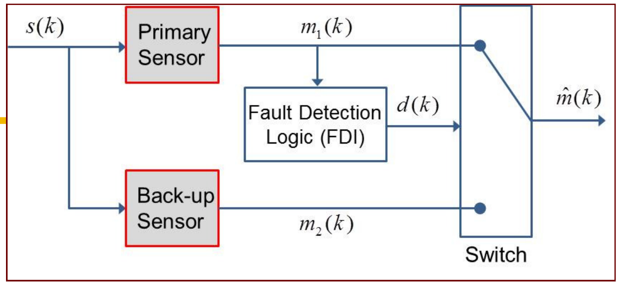
- Apply basic probability theory:

$$P_{S,N} = \Sigma_{k=1}^{N} Pr[T_S \geq k + N_0 \mid T_1 = k] Pr[T_1 = k]$$

$$+ Pr[T_S \leq N \mid T_1 = N+1] Pr[T_1 = N+1] Pr[T_2 \leq N]$$

$$+ \Sigma_{k=1}^{N} Pr[T_S < k + N_0 \mid T_1 = k] Pr[T_1 = k] Pr[T_2 \leq N]$$

System Failure Probability



- Apply basic probability theory:

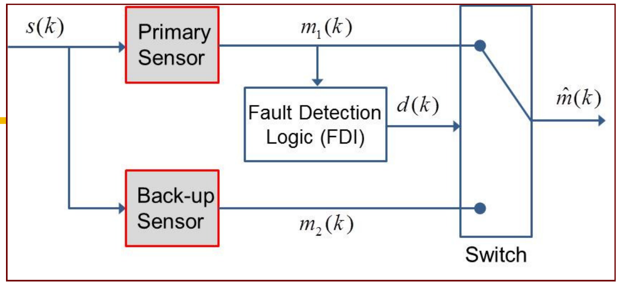
$$P_{S,N} = \Sigma_{k=1}^{N} Pr[T_S \geq k + N_0 \mid T_1 = k] \boxed{Pr[T_1 = k]}$$

$$+ Pr[T_S \leq N \mid T_1 = N+1] \boxed{Pr[T_1 = N+1] Pr[T_2 \leq N]}$$

$$+ \Sigma_{k=1}^{N} Pr[T_S < k + N_0 \mid T_1 = k] \boxed{Pr[T_1 = k] Pr[T_2 \leq N]}$$

- Knowledge of probabilistic performance
  a. Sensor failures: *P[T_i=k]* where T_i := failure time of sensor *i*

System Failure Probability

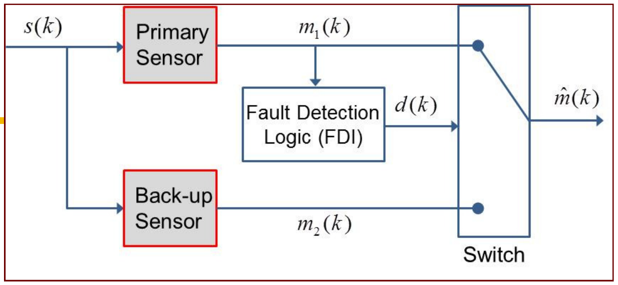


- Apply basic probability theory:

$$P_{S,N} = \Sigma_{k=1}^{N} Pr[T_S \geq k + N_0 \mid T_1 = k] \boxed{Pr[T_1 = k]}$$
$$+ \boxed{Pr[T_S \leq N \mid T_1 = N+1]} \boxed{Pr[T_1 = N+1] Pr[T_2 \leq N]}$$
$$+ \Sigma_{k=1}^{N} Pr[T_S < k + N_0 \mid T_1 = k] \boxed{Pr[T_1 = k] Pr[T_2 \leq N]}$$

- Knowledge of probabilistic performance
  a. Sensor failures: $P[\,T_i = k\,]$ where T_i := failure time of sensor i
  b. FDI False Alarm: $P[\,T_S \leq N \mid T_1 = N+1\,]$

System Failure Probability

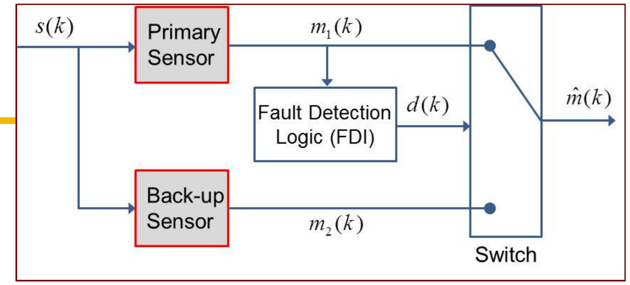


- Apply basic probability theory:

$$P_{S,N} = \Sigma_{k=1}^{N} \boxed{Pr[T_S \geq k + N_0 \mid T_1 = k]} \boxed{Pr[T_1 = k]}$$
$$+ \boxed{Pr[T_S \leq N \mid T_1 = N+1]} \boxed{Pr[T_1 = N+1] Pr[T_2 \leq N]}$$
$$+ \Sigma_{k=1}^{N} \boxed{Pr[T_S < k + N_0 \mid T_1 = k]} \boxed{Pr[T_1 = k] Pr[T_2 \leq N]}$$

- Knowledge of probabilistic performance
  a. Sensor failures: $P[\ T_i{=}k\ ]$ where T_i := failure time of sensor i
  b. FDI False Alarm: $P[\ T_S \leq N \mid T_1{=}N{+}1\ ]$
  c. FDI Missed Detection: $P[\ T_S \geq k{+}N_0 \mid T_1{=}k\ ]$

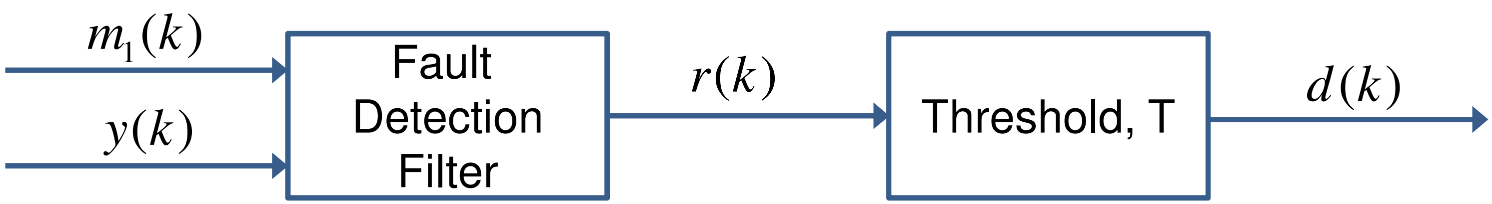
Example



- Sensor Failures: Geometric distribution with parameter *q*

$$q = 1 - e^{\frac{\Delta t}{MTBF}}$$

- Residual-based threshold logic



<u>Residual</u>

$$r(k+1) = n(k) + f(k)$$

f is an additive fault

n is IID Gaussian noise, variance=σ

<u>Decision Logic</u>

$$d(k) = \begin{cases} 0 & \text{if } |r(k)| \le T \\ 1 & \text{else} \end{cases}$$

Example

- Per-frame false alarm probability can be easily computed

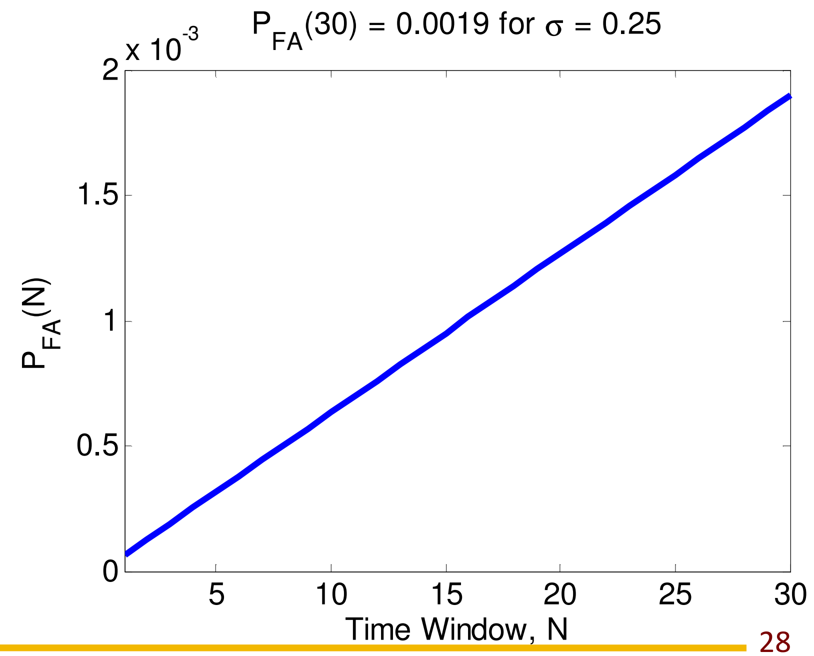
For each *k*, *r(k)* is N(0,σ²) : $P_F = \Pr\left[d(\mathrm{k}) = 1 \,\middle|\, \text{No Fault} \right] = 1 - \int_{-T}^{T} p(r)dr$

$$\boxed{P_F = 1 - erf\left(T / \sqrt{2\sigma^2} \right)}$$

- Approximate per-hour false alarm probability

$$\boxed{P[T_S \le N \mid T_1 = N+1] = 1 - (1 - P_F)^N \approx NP_F}$$

Per-frame detection probability P_D can be similarly computed.



System Failure Rate

- Notation:

$$\hat{q} := Nq$$ Sensor failure per hour

$$\hat{P}_F := NP_F$$ False alarm per hour

$$\hat{P}_D := 1 - (1 - P_D)^{N_0}$$ Detection per failure

- Approximate system failure probability:

$$P_{S,N} \approx \hat{q}(1 - \hat{P}_D) + \hat{P}_D\hat{q}^2 + \hat{P}_F\hat{q}(1 - \hat{q})$$

System Failure Rate

- Notation:

$$\hat{q} := Nq$$ Sensor failure per hour

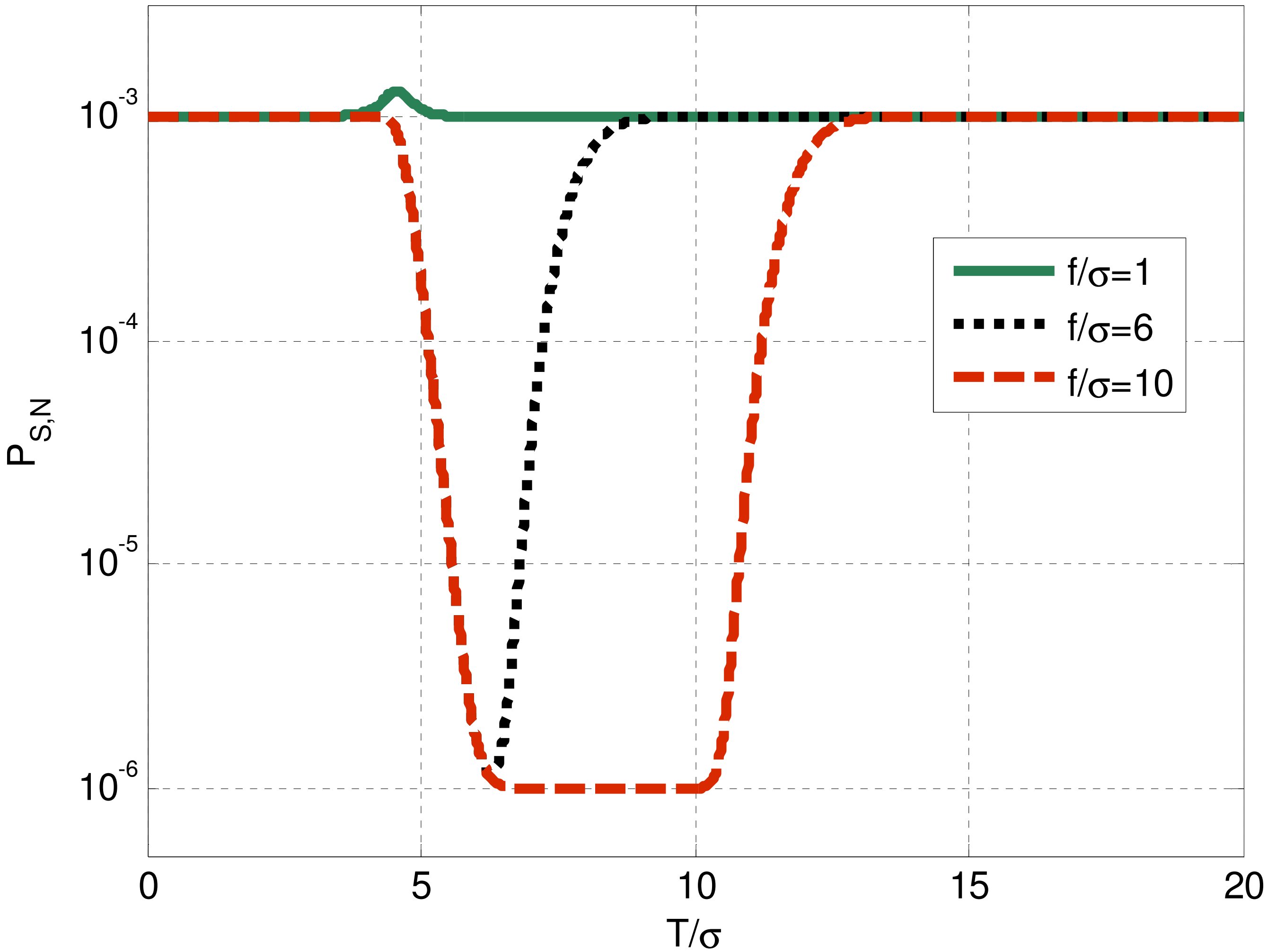
$$\hat{P}_F := NP_F$$ False alarm per hour

$$\hat{P}_D := 1 - (1 - P_D)^{N_0}$$ Detection per failure

- Approximate system failure probability:

$$P_{S,N} \approx \underbrace{\boxed{\hat{q}(1-\hat{P}_D)}}_{\text{Primary sensor fails + missed detection}} + \underbrace{\boxed{\hat{P}_D\hat{q}^2}}_{\text{Failure detected + Backup sensor fails}} + \underbrace{\boxed{\hat{P}_F\hat{q}(1-\hat{q})}}_{\text{False alarm + Backup sensor fails}}$$

System Failure Rate



Sensor mean time between failure = 1000hr
and N=360000 (= 1 hour at 100Hz rate)

Correlated Residuals

- Example analysis assumed IID fault detection logic.
- Many fault-detection algorithms use dynamical models and filters that introduce correlations in the residuals.

- **Question:** How can we compute the FDI performance metrics when the residuals are correlated in time?
  - FDI False Alarm: $P[\ T_S \leq N \mid T_1 = N+1\ ]$
  - FDI Missed Detection: $P[\ T_S \geq k+N_0 \mid T_1 = k\ ]$

False Alarm Analysis with Correlated Residuals

- Problem: Analyze the per-hour false alarm probability for a simple first-order fault detection system:

Residual Generation (*0<a<1*)

$$r_{k+1} = ar_k + n_k + f_k$$

f is an additive fault

n is IID Gaussian noise, variance=1

Decision Logic

$$d_k = \begin{cases} 0 & \text{if } |r_k| \le T \\ 1 & \text{else} \end{cases}$$

Residuals are correlated in time due to filtering

- The *N-step false alarm probability* P_N is the conditional probability that $d_k=1$ for some $1 \le k \le N$ given the absence of a fault.

$$P_N = 1 - \int_{-T}^{T} \cdots \int_{-T}^{T} p_R(r_1, \ldots, r_N) dr_1 \cdots dr_N$$

There are N=360000 samples per hour for a 100Hz system

False Alarm Analysis

- Residuals satisfy the Markov property:

$$r_{k+1} = ar_k + n_k + f_k \quad \Rightarrow \quad p\left(r_{k+1} \middle| r_1, \cdots, r_k\right) = p\left(r_{k+1} \middle| r_k\right)$$

$$\Rightarrow \quad p_R\left(r_1, \cdots, r_k\right) = p\left(r_k \middle| r_{k-1}\right) \cdots p\left(r_2 \middle| r_1\right) \cdot p_1\left(r_1\right)$$

- P_N can be expressed as an N-step iteration of 1-dimensional integrals:

$$P_N = 1 - \int_{-T}^{T} \cdots \int_{-T}^{T} p_R(r_1, \ldots, r_N) dr_1 \cdots dr_N \quad \Rightarrow$$

$$\boxed{\begin{aligned} & f_N(r_N) = 1 \\ & f_{N-1}(r_{N-1}) = \int_{-T}^{T} f_N(r_N) p(r_N | r_{N-1}) dr_N \\ & \vdots \\ & f_1(r_1) = \int_{-T}^{T} f_2(r_2) p(r_2 | r_1) dr_2 \\ & P_N = 1 - \int_{-T}^{T} f_1(r_1) p_1(r_1) dr_1 \end{aligned}}$$

This has the appearance of a power iteration $A^N x$

False Alarm Probability

- **Theorem:** Let λ_1 be the maximum eigenvalue and ψ_1 the corresponding eigenfunction of

$$\lambda_1 \psi_1(x) = \int_{-T}^{T} \psi_1(y) p(y \mid x) dy$$

Then $P_N \approx c \lambda_1^{N-1}$ where $c = \langle 1, \psi_1 \rangle$

- **Proof**
  - This is a generalization of the matrix power iteration
  - The convergence proof relies on the Krein-Rutman theorem which is a generalization of the Perron-Frobenius theorem.
  - For a=0.999 and N=360000, the approximation error is 10^{-156}

Ref: B. Hu and P. Seiler. False Alarm Analysis of Fault Detection Systems with Correlated Residuals, Submitted to IEEE TAC, 2012.

Results: Effects of Correlation

False Alarm Probabilities and Bounds for N=360,000

Neglecting correlations is accurate for small *a*

a	T	P_N	$1 - L_N^{(2)}$	$1 - L_N^{(1)}$
0	6.807	3.600×10^{-6}	3.600×10^{-6}	3.600×10^{-6}
0.7	9.531	3.587×10^{-6}	3.587×10^{-6}	3.598×10^{-6}
0.8	11.34	3.524×10^{-6}	3.524×10^{-6}	3.526×10^{-6}
0.9	15.62	3.167×10^{-6}	3.173×10^{-6}	3.200×10^{-6}
0.99	48.25	9.641×10^{-7}	1.177×10^{-6}	1.360×10^{-6}
0.999	152.2	1.395×10^{-7}	3.401×10^{-7}	4.446×10^{-7}

…but not for *a* near 1.

For each (a,T), $P_1 = 10^{-11}$ which gives $NP_1 = 3.6 \times 10^{-6}$

Residual Generation

$$r_{k+1} = a r_k + n_k + f_k$$

Decision Logic

$$d_k = \begin{cases} 0 & \text{if } |r_k| \le T \\ 1 & \text{else} \end{cases}$$

Conclusions

- Commercial aircraft achieve high levels of reliability.
  - Analytical redundancy is rarely used (Certification Issues)
  - Model-based fault detection methods are an alternative that enables size, weight, power, and cost to be reduced.

- Certification Approach:
  - Use linear analysis to prove performance at many flight conditions (Initial result on effect of correlated residuals)
  - Use high fidelity Monte Carlo simulations to confirm (or reject) linear results.
  - Future Work: Need to consider model uncertainty and worst-case trajectories.

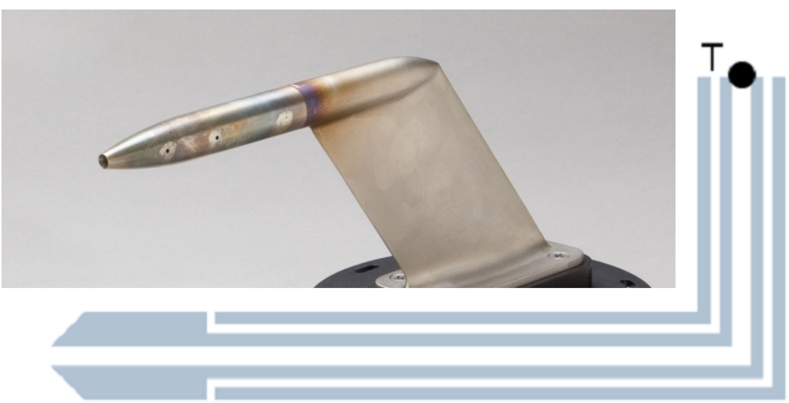
Acknowledgments

- NASA Langley NRA NNX12AM55A: “Analytical Validation Tools for Safety Critical Systems Under Loss-of-Control Conditions,” Technical Monitor: Dr. Christine Belcastro

- Air Force Office of Scientific Research: Grant No. FA9550-12-0339, "A Merged IQC/SOS Theory for Analysis of Nonlinear Control Systems,” Technical Monitor: Dr. Fariba Fahroo.

- NSF Cyber-Physical Systems: Grant No. 0931931, “Embedded Fault Detection for Low-Cost, Safety-Critical Systems,” Program Manager: Theodore Baker.

Motivation: Increased Reliability



- Air data measurements used to estimate critical flight data (airspeed / angle of attack)
- Air data failures are the suspected root cause in several accidents.

Year	Flight	Suspected Cause
1974	NW6231	Iced Pitot
1996	Birgenair301	Blocked Pitot (Insects)
1996	AeroPeru603	Blocked Static (Tape)
2008	B-2	Moisture
2009	AirFrance447	Pitot Malfunction

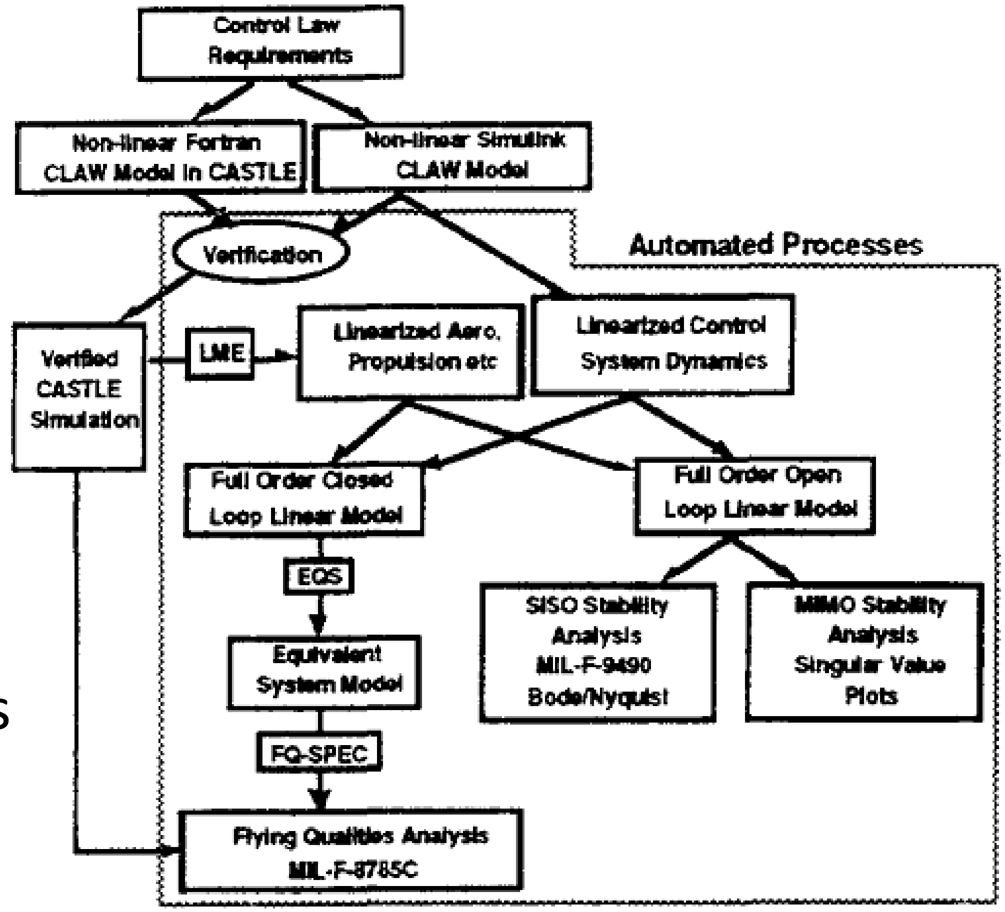
Analytical air data estimates can protect against common failure modes.

Certification of Analytically Redundant Systems

Analogy to V&V of Flight CLAWs:

- Use linear analysis to prove performance at many flight conditions
- Use high fidelity Monte Carlo simulations to confirm (or reject) linear results.
- Research: Extend linear analysis tools to polynomial systems

http://www.aem.umn.edu/~AerospaceControl/



Ref: J. Renfrow, S. Liebler, and J. Denham. "F-14 Flight Control Law Design, Verification, and Validation Using Computer Aided Engineering Tools," 1996.