Semi-Markov Switching Vector Autoregressive Model-Based Anomaly Detection in Aviation Systems

Igor Melnyk[1], Arindam Banerjee[1], Bryan Matthews[2] and Nikunj Oza[2]

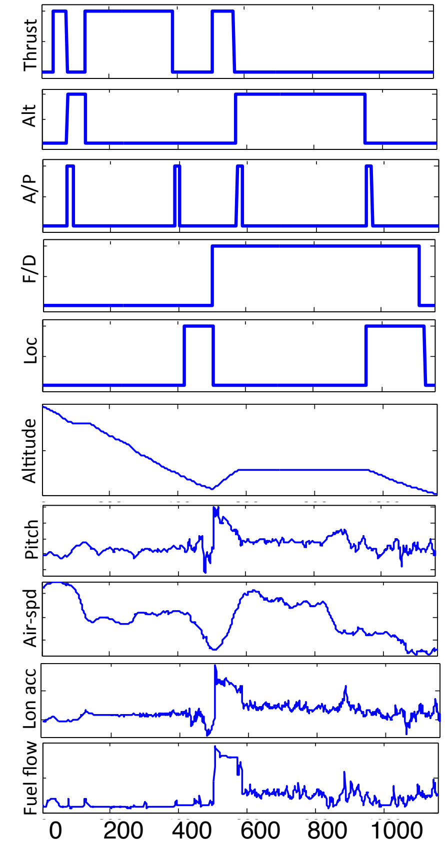
[1]Department of Computer Science & Engineering, University of Minnesota, Twin Cities, MN

[2]NASA Ames Research Center, Moffett Field, CA

Knowledge Discovery and Data Mining Conference
San Francisco, CA

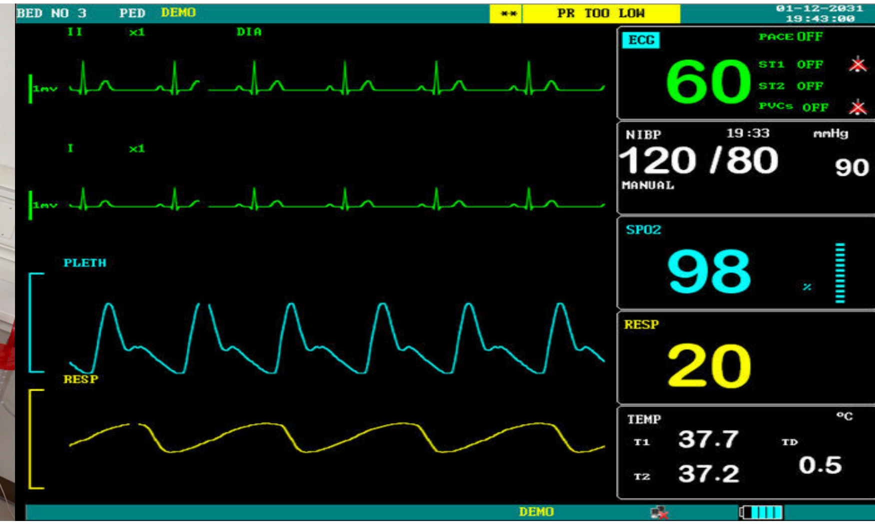
August 16, 2016

Aviation Systems



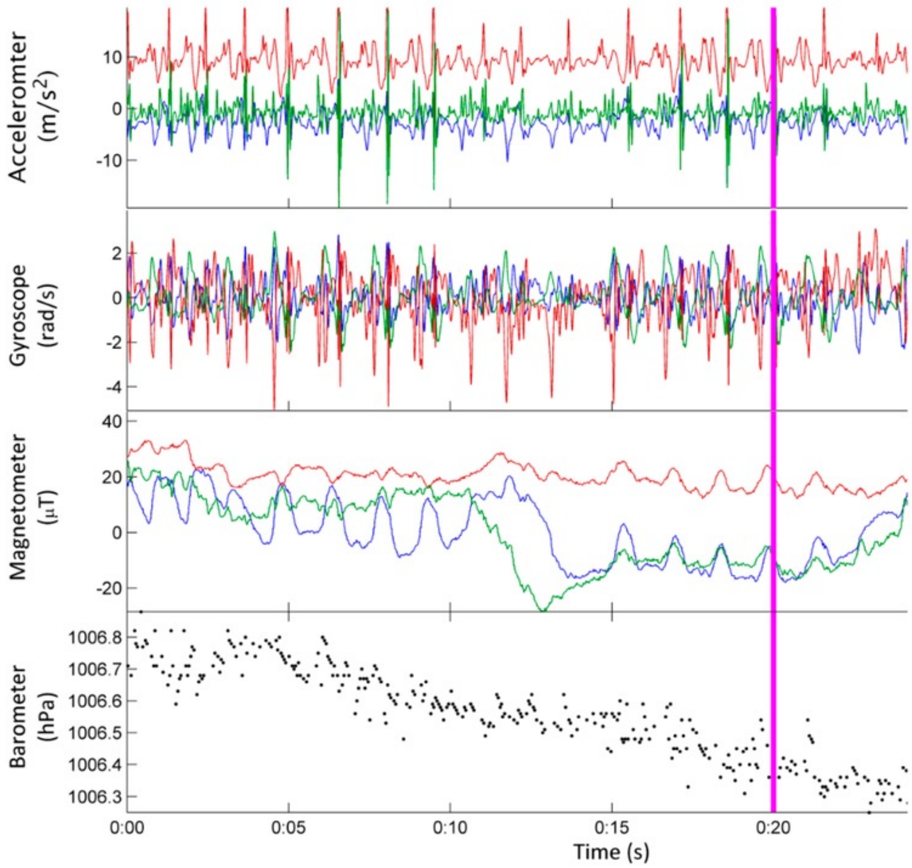
- Data
  - Flight sensors, pilot commands, weather information
- Objective
  - Monitor flight, detect anomalous activity

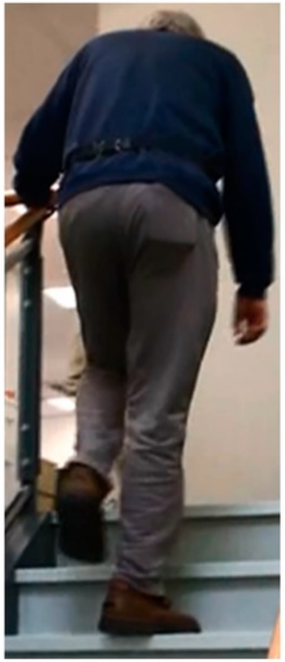
Healthcare



- Data
  - Multiple records of vital signs: blood pressure, temperature, pulse
  - Interventions: injections, pills, drop counter
- Objective
  - Monitor patient’s health

Activity Recognition





- Data
  - Multiple wearable sensors: accelerometer, gyroscope, barometer
- Objective
  - Recognize activity: running, walking, standing

Discussion

- Data
  - Dynamic
  - Multivariate
  - Heterogeneous
  - Variable length
  - Noisy
  - Partially unobservable

Discussion

- Data
  - Dynamic
  - Multivariate
  - Heterogeneous
  - Variable length
  - Noisy
  - Partially unobservable

- Challenges
  - Characterize interdependencies between multiple data streams
    - Continuous, discrete data type
  - Design of detection algorithms
    - Find patterns, trends, anomalies in unsupervised settings

This Work

- Anomaly detection in aviation systems
  - Model flights using Dynamic Bayesian Network (DBN) representation
  - Detect anomalous activities
    - Mechanical problems
    - Weather causes
    - Human factors

This Work

- Anomaly detection in aviation systems
  - Model flights using Dynamic Bayesian Network (DBN) representation
  - Detect anomalous activities
    - Mechanical problems
    - Weather causes
    - Human factors

- Data
  - NASA flight dataset
  - 10^6 flights, 35 aircrafts, 300 parameters, sampled at 1Hz, duration 1-3 hours
  - Multivariate, variable length, heterogeneous
  - No labeling information available

Flight Data

Environment

continuous

- Air temperature
- Wind direction
- Wind speed
- Pressure

Controls

discrete

- Altitude mode
- Thrust mode
- Vertical speed mode
- Autopilot mode

- Selected heading
- Selected altitude
- Selected vertical speed

continuous

- Aileron position
- Elevator position
- Rudder position
- Lever position
- N1 command

Aircraft

Sensors

continuous

- Altitude
- Angle of attack
- Drift angle
- Roll, Pitch
- Heading
- Lateral acceleration
- Longitudinal acceleration
- Normal acceleration
- Ground speed
- Vertical speed
- Airspeed
- Glide slope deviation
- Localizer deviation
- Engine temperature
- Low rotor speed
- High rotor speed
- Oil pressure
- Oil quantity
- Oil temperature
- Fuel mass flow

Related Work

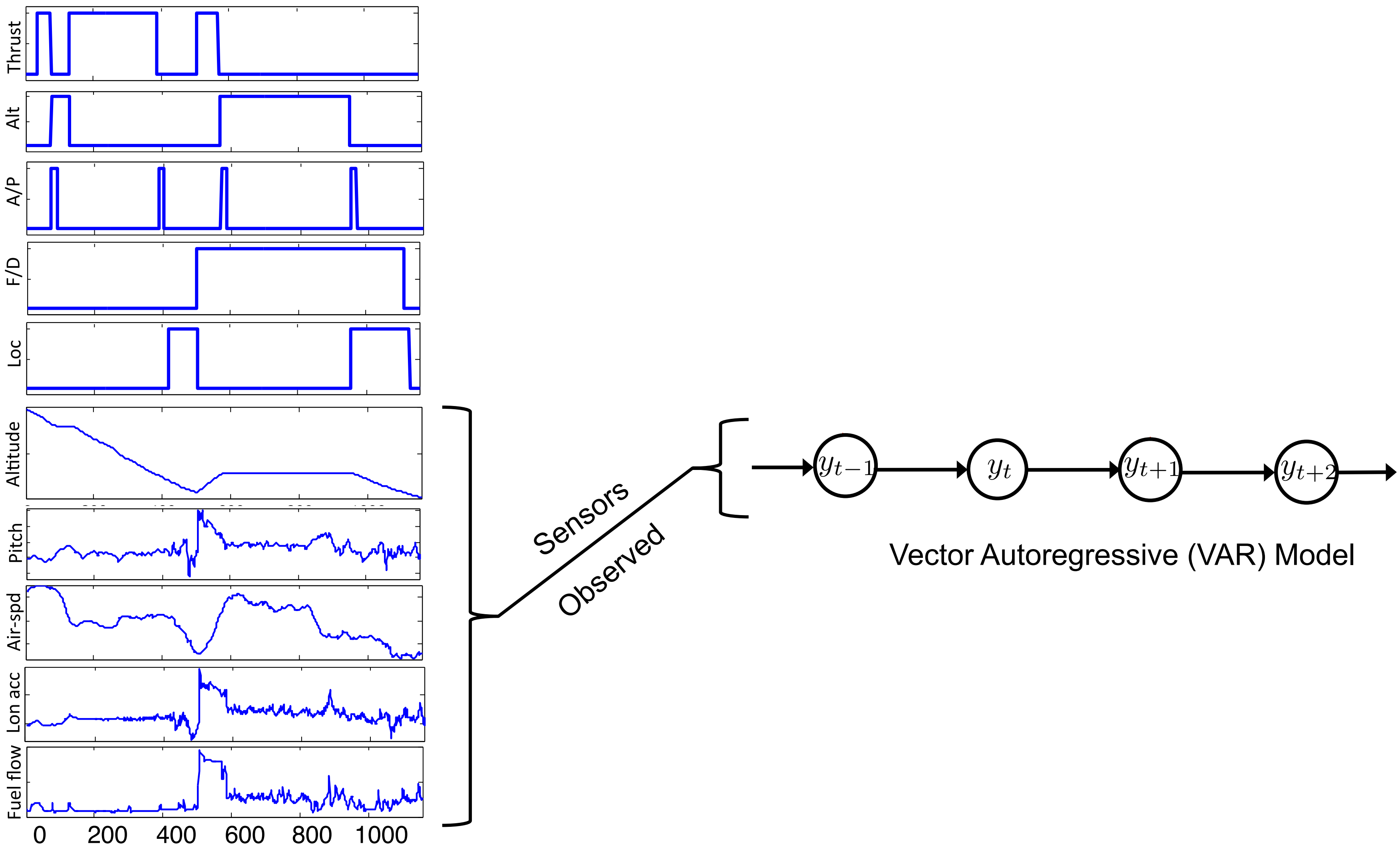
- Linear regression *[Srivastava, ’12]*
  - Detect abnormal fuel consumption in jet engines
  - Supervised approach, requires labeled data
- Clustering *[Budalakoti et al., ’09]*
  - Detect anomalies in pilot switches
  - Restricted to discrete data
- Intent inference *[Lee et al., ’14]*
  - Detect human-automation issues using pilot and sensor measurements
  - Assumes noise-free data
- Multiple kernel learning (MKAD) *[Das et al., ’10]*
  - Detect anomalies in heterogeneous flight data
  - Limited scalability due to kernel matrix updates

Related Work

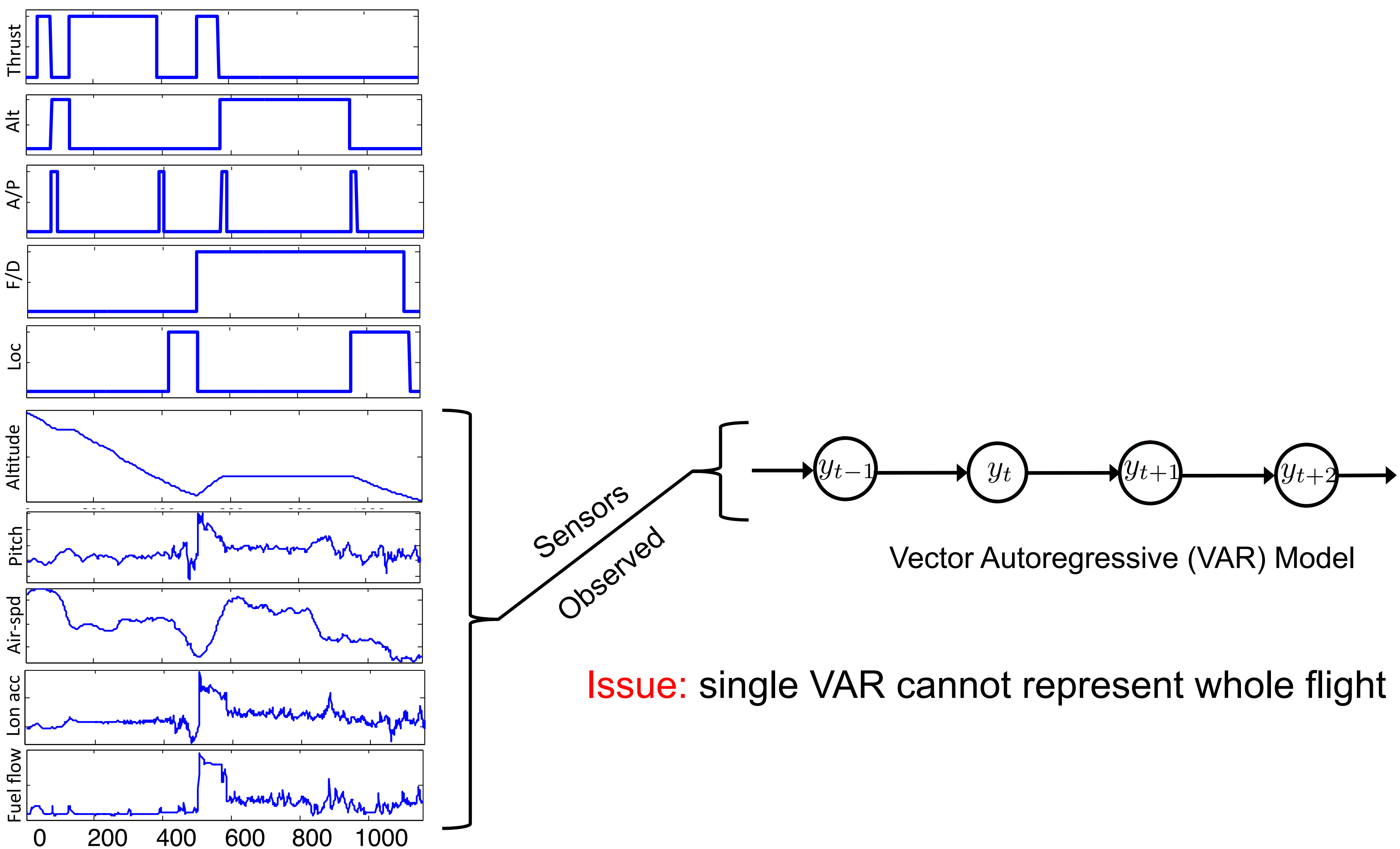
- Linear regression *[Srivastava, '12]*
  - Detect abnormal fuel consumption in jet engines
  - Supervised approach, requires labeled data
- Clustering *[Budalakoti et al., '09]*
  - Detect anomalies in pilot switches
  - Restricted to discrete data
- Intent inference *[Lee et al., '14]*
  - Detect human-automation issues using pilot and sensor measurements
  - Assumes noise-free data
- Multiple kernel learning (MKAD) *[Das et al., '10]*
  - Detect anomalies in heterogeneous flight data
  - Limited scalability due to kernel matrix updates
- Our approach
  - Unsupervised, model-based, computationally efficient
  - Works with multivariate heterogeneous time series data

University of Minnesota
Driven to Discover℠

Flight Data Modeling

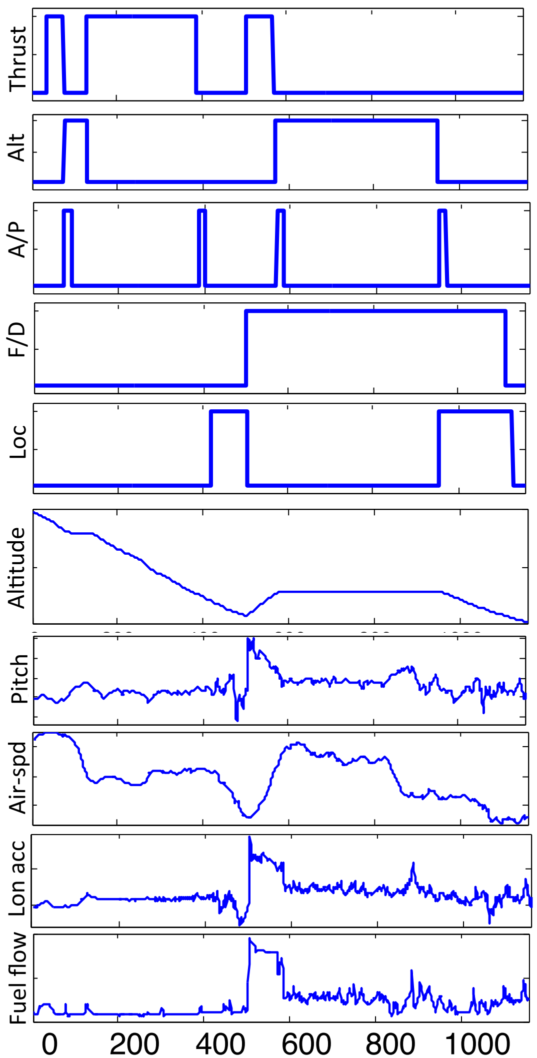


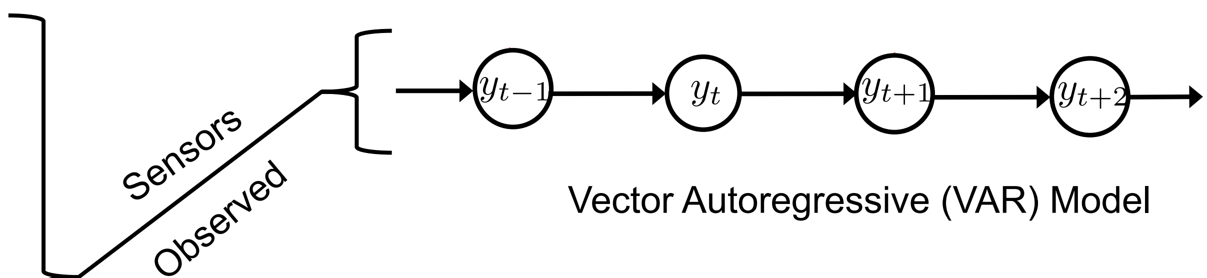
Flight Data Modeling



Issue: single VAR cannot represent whole flight

Flight Data Modeling



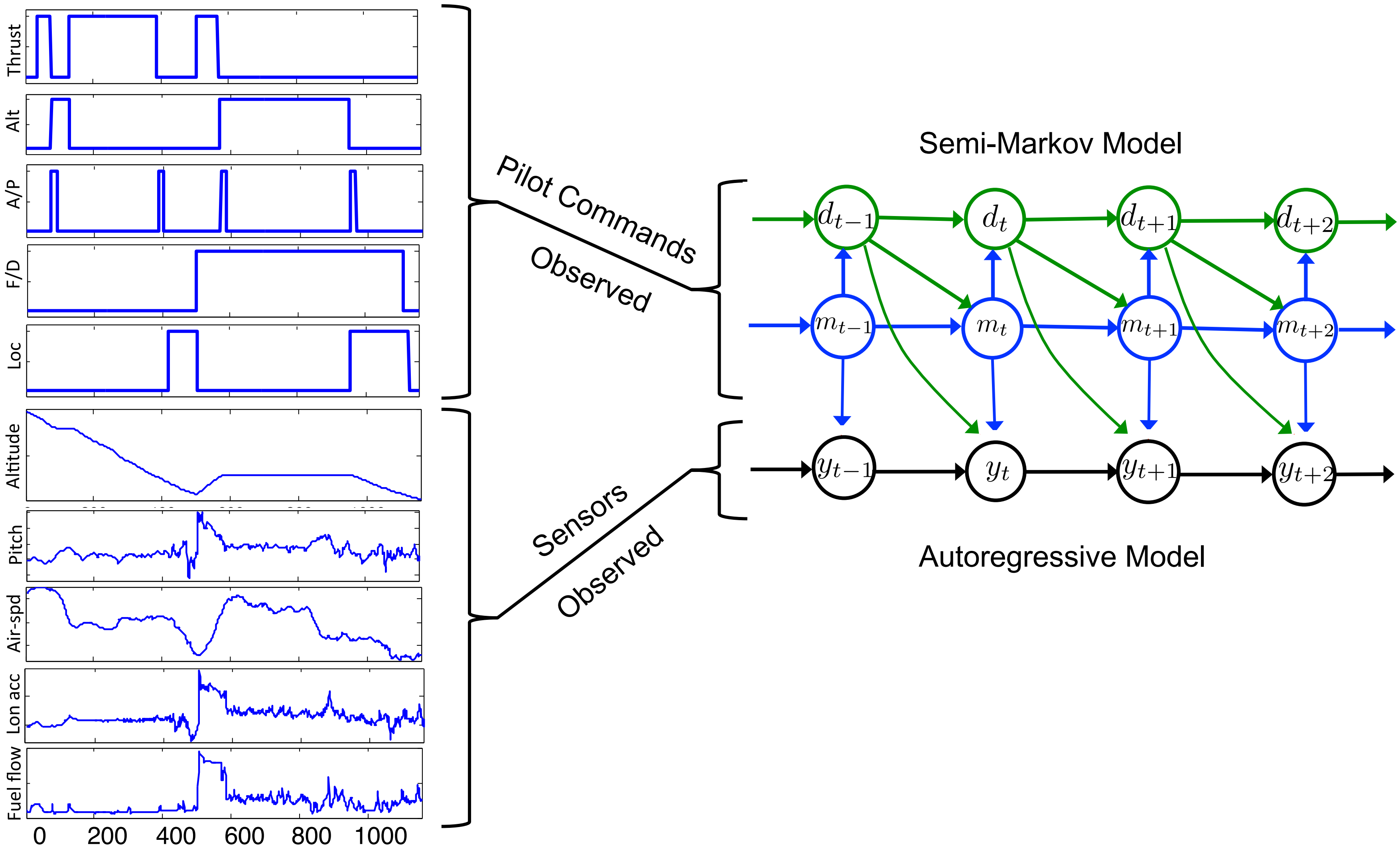


Vector Autoregressive (VAR) Model

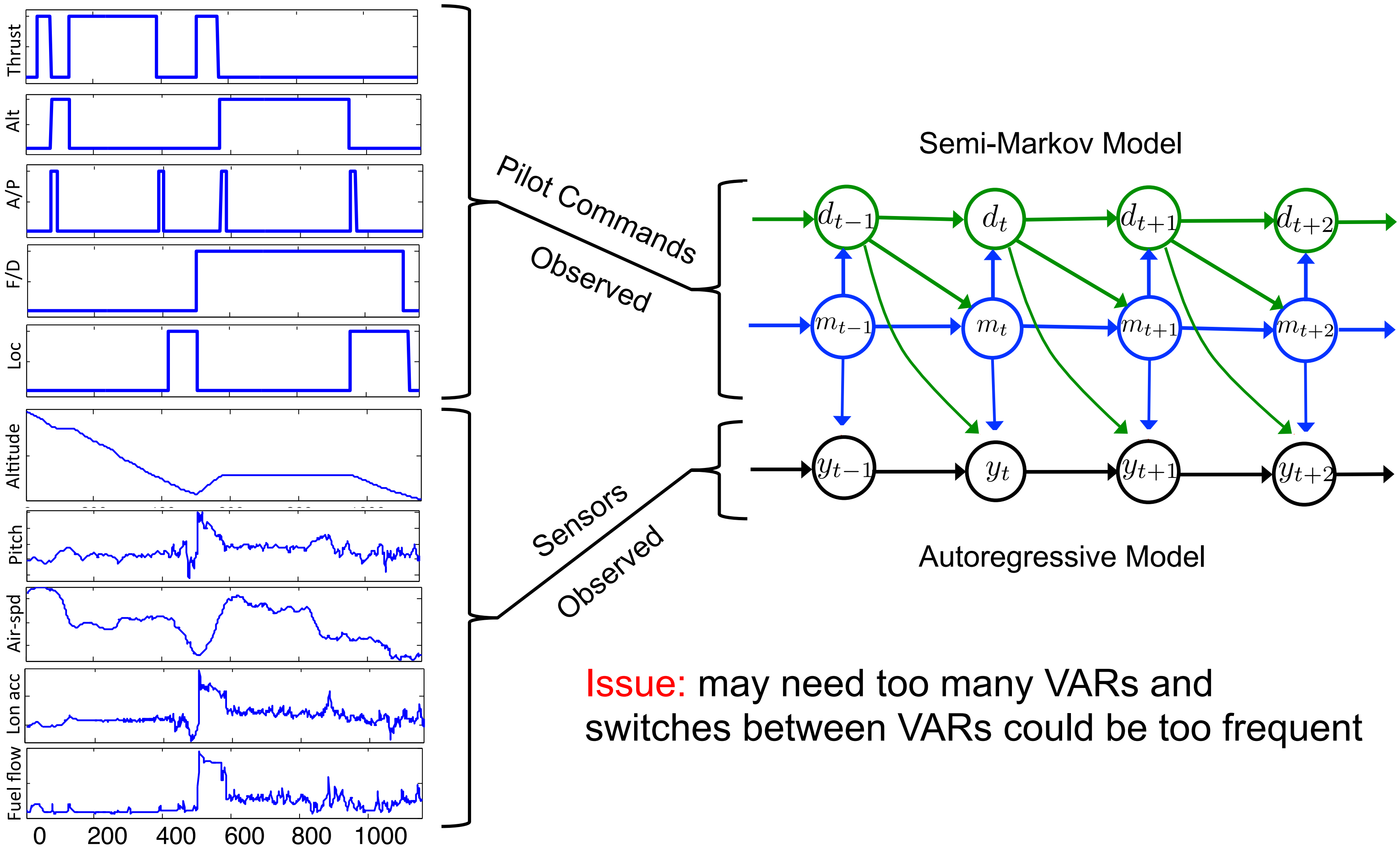
Issue: single VAR cannot represent whole flight

Solution: use many VARs, determined by pilot commands

Flight Data Modeling

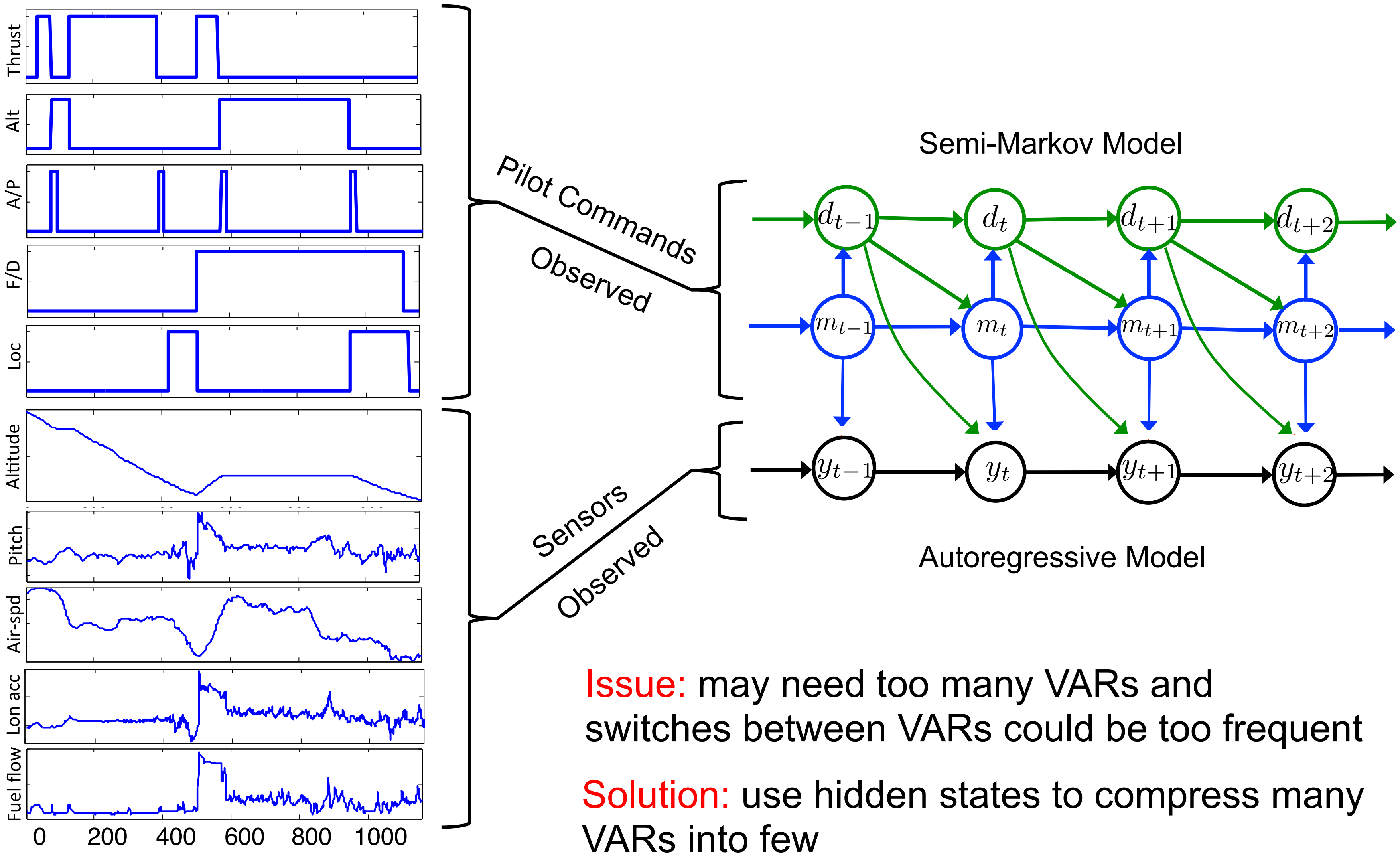


Flight Data Modeling



Issue: may need too many VARs and switches between VARs could be too frequent

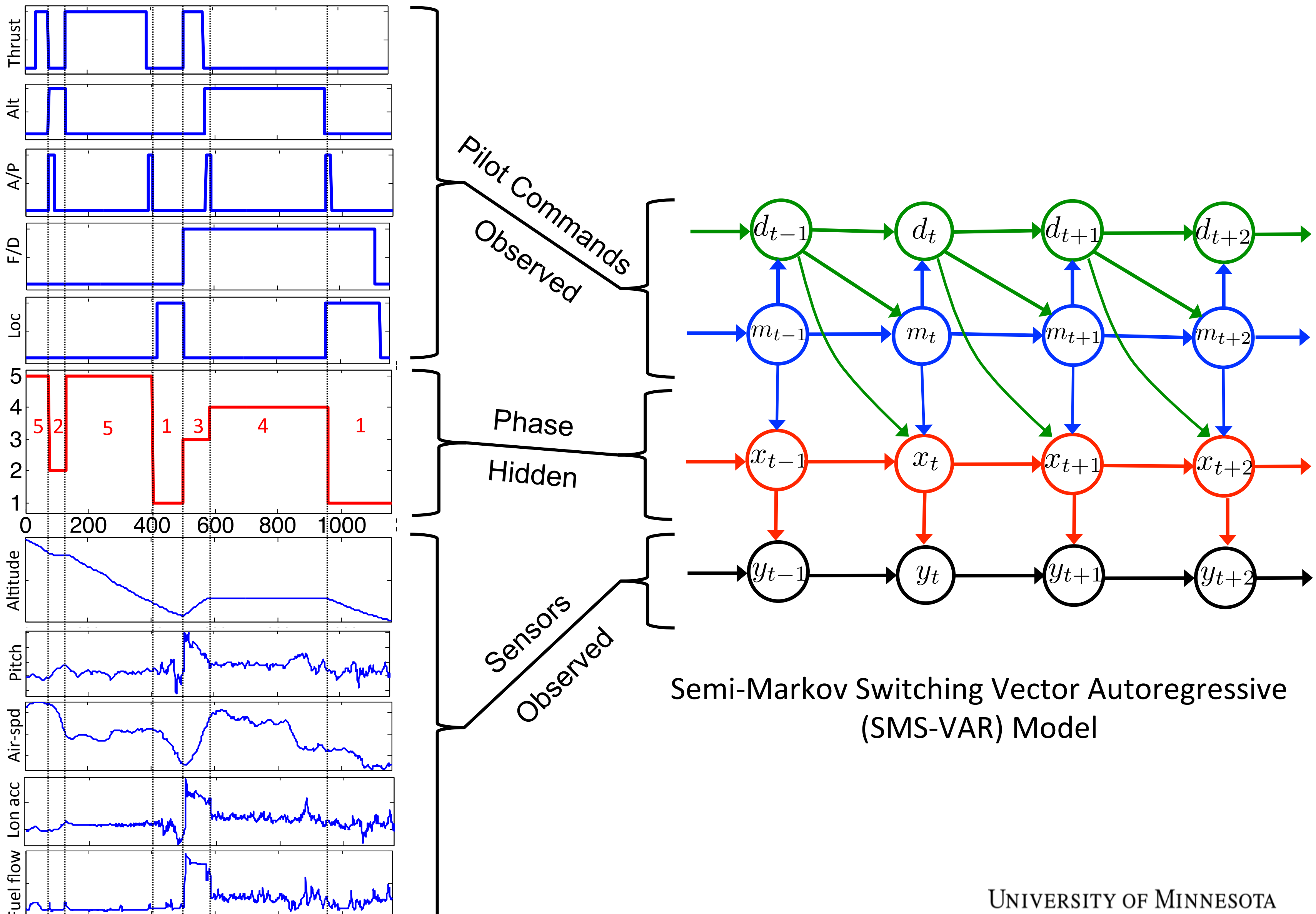
Flight Data Modeling



Issue: may need too many VARs and switches between VARs could be too frequent

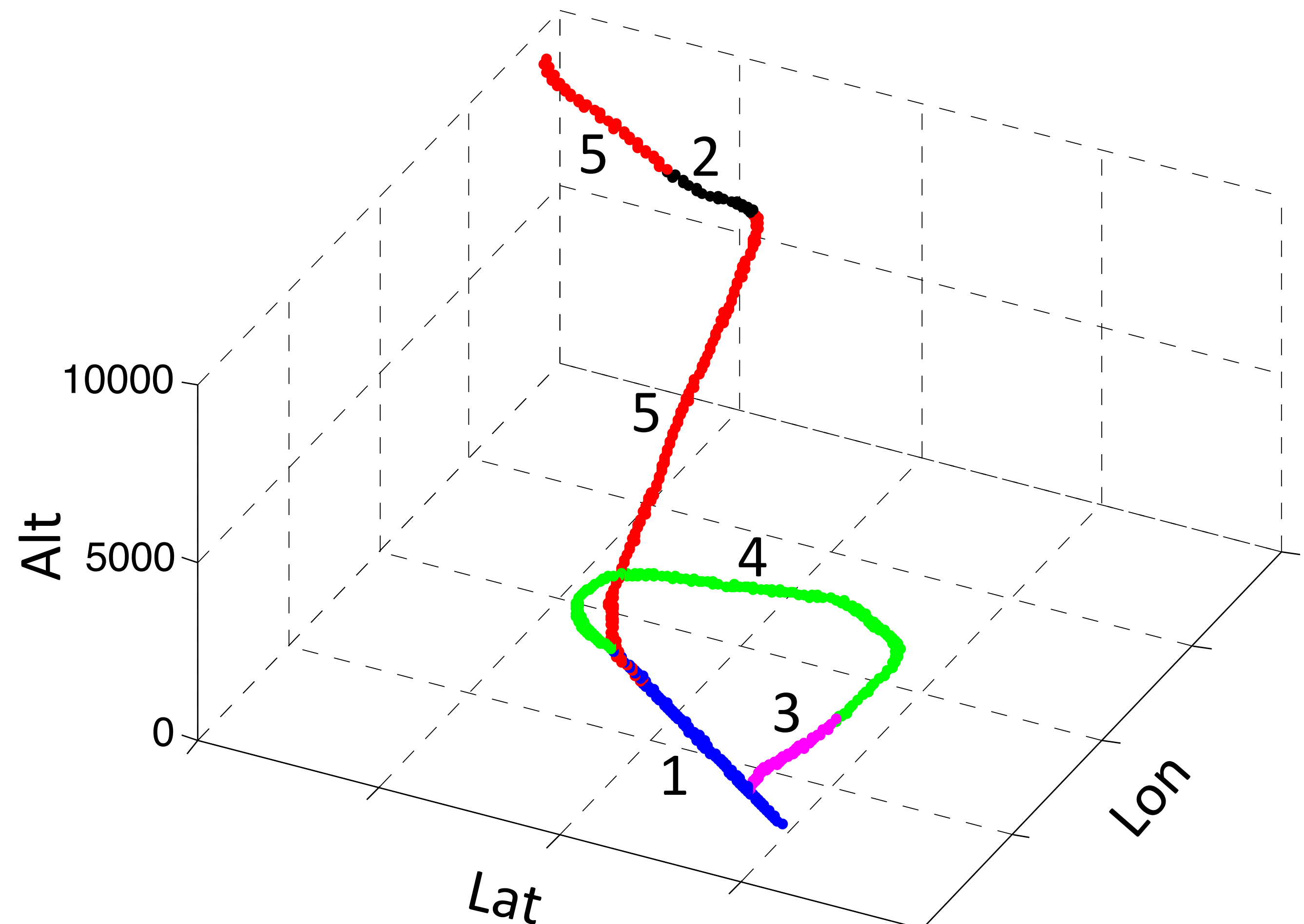
Solution: use hidden states to compress many VARs into few

Flight Data Modeling



Semi-Markov Switching Vector Autoregressive (SMS-VAR) Model

Example: Flight Trajectory Partitioning



UNIVERSITY OF MINNESOTA
Driven to Discover℠

Anomaly Detection Algorithm

- Objective
  - Given dataset of unlabeled flights, detect anomalous flights

- Step 1
  - Construct a single SMS-VAR model
    - Using all the flights (mixture of normal and abnormal)
    - Assumption: anomalous flights are rare
    - Parameter learning is done using Expectation Maximization

Anomaly Detection Algorithm

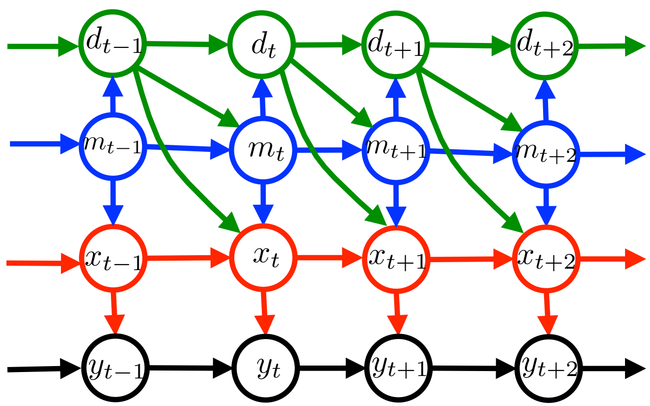
- Objective
  - Given dataset of unlabeled flights, detect anomalous flights
- Step 1
  - Construct a single SMS-VAR model
    - Using all the flights (mixture of normal and abnormal)
    - Assumption: anomalous flights are rare
    - Parameter learning is done using Expectation Maximization
- Step 2
  - Evaluate constructed model on all the flights
    - Compute anomaly scores
    - Declare anomalies when exceeding certain threshold

Anomaly Detector

- Standard way
  - Compute likelihood of whole flight $p(F) = p(\bar{d}_{1:T}, \bar{m}_{1:T}, \bar{y}_{1:T})$

Anomaly Detector

- Standard way
  - Compute likelihood of whole flight $p(F) = p(\bar{d}_{1:T}, \bar{m}_{1:T}, \bar{y}_{1:T})$
- Proposed
  - Compute dissimilarities between phase distributions
  - Main idea



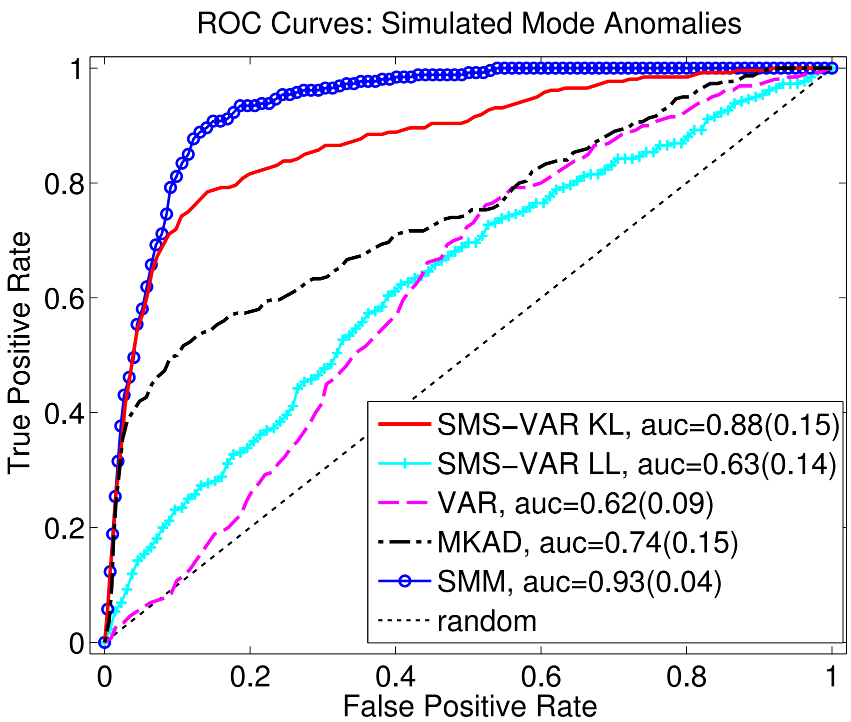
    - Compute one-step ahead prediction: $p(x_{t+1}|\bar{d}_{1:t}, \bar{m}_{1:t}, \bar{y}_{1:t})$
    - Compute filtered (after data observation): $p(x_{t+1}|\bar{d}_{1:t+1}, \bar{m}_{1:t+1}, \bar{y}_{1:t+1})$
    - Dissimilarity (KL-divergence): $D_{t+1}\left[p(x_{t+1}|F_{1:t})\middle|\middle|p(x_{t+1}|F_{1:t+1})\right]$
    - Anomaly score: std of all D_t's

Compared Algorithms

- SMS-VAR KL
  - Model: SMS-VAR (discrete + continuous data)
  - Anomaly detector: std of phase dissimilarities based on KL divergence
- SMS-VAR LL
  - Model: SMS-VAR (discrete + continuous data)
  - Anomaly detector: log-likelihood value
- VAR
  - Model: single VAR fitted to whole flight (continuous data only)
  - Anomaly detector: std of one-step-ahead prediction errors
- SMM
  - Model: semi-Markov chain (discrete data only)
  - Anomaly detector: std of one-step-ahead prediction errors
- MKAD
  - Multiple kernel learning approach (discrete + continuous data)
  - Anomaly detector: one-class SVM applied to computed kernel

Experiments: Synthetic Data

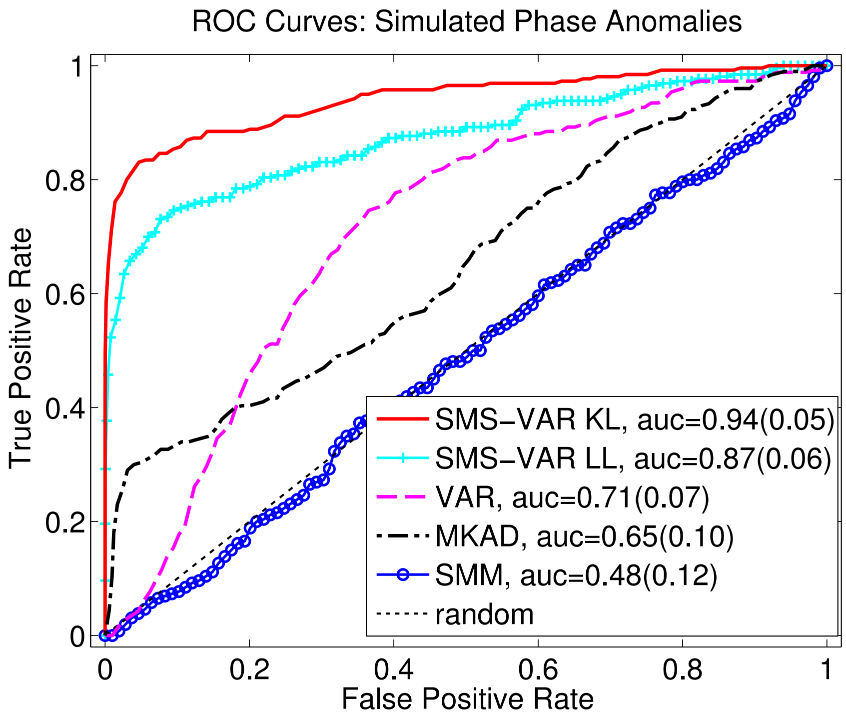
- Detecting pilot switches anomalies



  - SMM and SMS-VAR-KL have better performance

Experiments: Synthetic Data

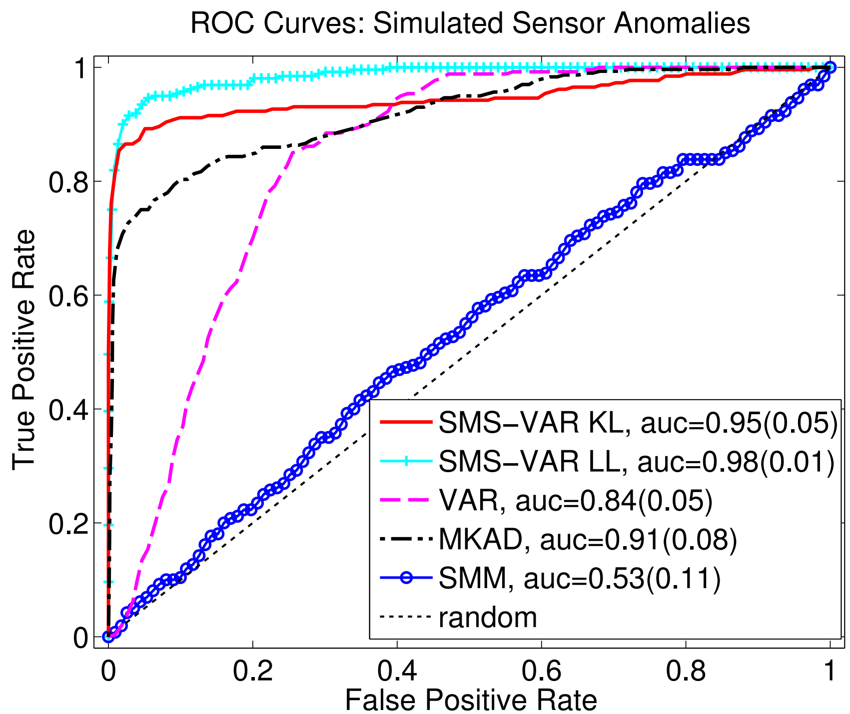
- Detecting phase anomalies



- SMS-VAR performs best (KL version is more accurate)
- SMM has lowest accuracy

Experiments: Synthetic Data

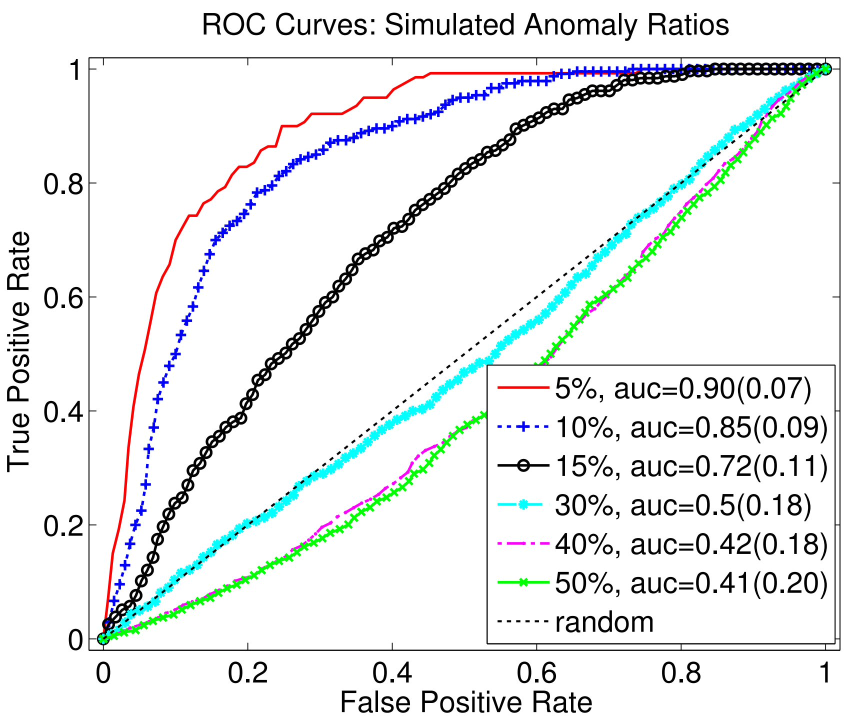
- Detecting sensor anomalies



  - SMS-VAR and MKAD perform well
  - VAR and SMM perform poorly

Experiments: Synthetic Data

- Effect of Anomaly Proportion on Accuracy

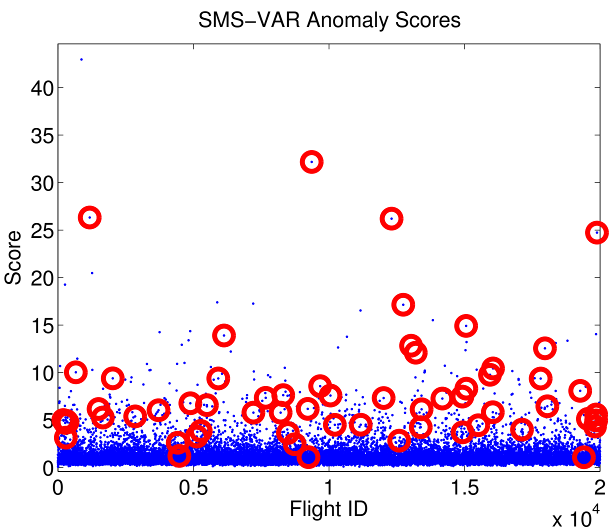


  - High fraction of anomalies leads to low detection accuracy
  - Abnormal flights start looking as normal when the fraction of anomalies is high

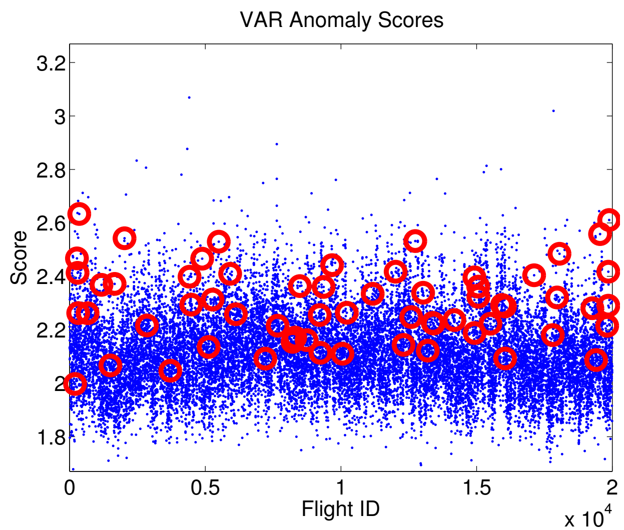
University of Minnesota
Driven to Discover℠

Experiments: Aviation Data

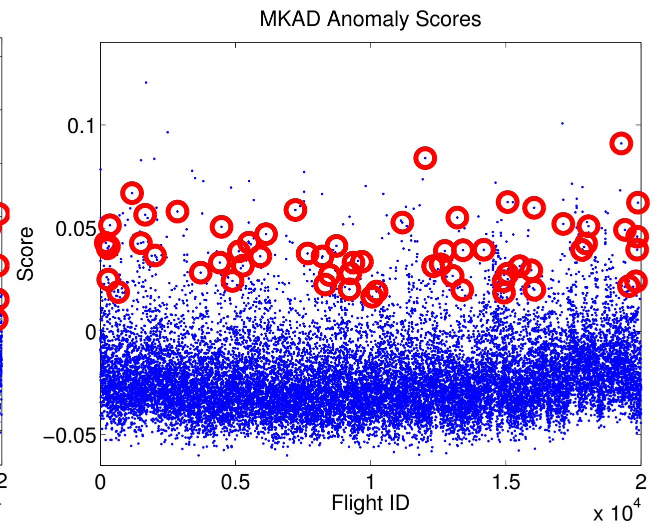
- Unlabeled data
  - 20000 flights, no labeling information, landing part of flight
  - Red circles represent Go-Around flights (total 61 of them)



a)



b)



c)

Experiments: Aviation Data

- Unlabeled data
  - 20000 flights, no labeling information, landing part of flight
  - Detected anomalous flights in the top-100 lists of each method

SMS-VAR KL	VAR	MKAD
go-around (19)	go-around (3)	go-around (17)
high speed in approach (5)	fast approach (2)	high pitch at touch down (1)
high rate of descent in approach (4)	high speed in approach (5)	high speed in approach (2)
bounced landing (2)	high rate of descent in approach (4)	low speed at touch down (1)
delayed braking at landing(2)	bank cycling in approach(2)	low path in approach (1)
late retraction of landing gear (4)	high pitch at touch down (1)	flaps retracted in approach (1)
deviation from glide-slope (2)		unusual flight switch changes (15)
unusual flight switch changes (11)		

- SMS-VAR-KL and MKAD performed better than the simple VAR
- They detected many significant discrete- and continuous-type anomalies

Experiments: Aviation Data

- Unlabeled data
  - 20000 flights, no labeling information, landing part of flight
  - Detected anomalous flights in the top-100 lists of each method

SMS-VAR KL	VAR	MKAD
go-around (19) high speed in approach (5) high rate of descent in approach (4) bounced landing (2) delayed braking at landing(2) late retraction of landing gear (4) deviation from glide-slope (2) unusual flight switch changes (11)	go-around (3) fast approach (2) high speed in approach (5) high rate of descent in approach (4) bank cycling in approach(2) high pitch at touch down (1)	go-around (17) high pitch at touch down (1) high speed in approach (2) low speed at touch down (1) low path in approach (1) flaps retracted in approach (1) unusual flight switch changes (15)

  - SMS-VAR-KL and MKAD performed better than the simple VAR
  - They detected many significant discrete- and continuous-type anomalies

Thank you!

University of Minnesota
Driven to Discover℠