INSTITUTE FOR DEFENSE ANALYSES

Analyzing Data with Ordinal Responses Observed Continuously Over Time: Data from the Shadow FOT&E

Matthew Avery
Mark Orndorff

August 2015

Approved for public release;
distribution is unlimited

IDA Document NS D-5581
Log: H 15-000823

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

About This Publication

Recent operational tests for the Shadow and Gray Eagle (two unmanned air vehicle (UAV) systems employed by the U.S. Army) have used ordinal response metrics to assess each system's ability to provide a continuous video stream to the One System Remote Video Terminal (OSRVT), the system employed by the end user to view data from these UAVs. Data collectors assign levels from an ordinal scale continuously over the duration of each mission while covariates such as the antennae used by the receiving unit and the distance of the broadcasting air vehicle to the ground receiver are also recorded. Thus, we are faced with a problem of an ordinal response variable observed continuously over time. Since the covariates are also observed continuously over time, we are faced with a problem best suited to the tool of functional data analysis. However, the problem of ordinal responses observed continuously over time is not well discussed in the literature. Here, we present a variety of approaches for analyzing this unique but relevant type of data, using the recent FOT&E for the RQ-7B Shadow as an example. The research and article were funded by the cross-divisional statistics and data science working group (C9082), but were an outgrowth of work conducted by OED for DOT&E.

INSTITUTE FOR DEFENSE ANALYSES

IDA Document NS D-5581

Analyzing Data with Ordinal Responses Observed Continuously Over Time: Data from the Shadow FOT&E

Matthew Avery
Mark Orndorff

Analyzing Data with Ordinal Responses Observed Continuously Over Time: Data From the Shadow FOT&E

Matthew Avery

Mark Orndorff

Capabilities and Mission of the Shadow

- **System Attributes**
  – 20.4 foot wingspan
  – 460 lbs
  – 7 hour endurance
  – Pneumatic launcher
  – Lands on paved runway

- **Tactical unmanned air vehicle used primarily for Intelligence, Surveillance, and Reconnaissance**

- **Provides continuous video to supported units**

- **Used extensively in Iraq and Afghanistan (over 750,000 flight hours and over 173,000 missions)**

Follow-On Operational Test for Shadow

- **Operational Testing**
  - Simulate actual environment in which warfighters will employ systems
  - More accurate characterization of system performance than developmental/contractor tests
- **Network Integrated Exercise at Ft. Bliss, TX**
  - Large exercise with many units cooperating to accomplish realistic mission
  - “Opposition force” present and active
- **Evaluate upgrades to Shadow**
  - Ground station
  - Wing span (endurance)
  - *Data link*
  - *Interface with ground units*
- **Three weeks of testing**
  - 260 hours of flight operations
  - 41 flights

Data Summary

- **Data collected from air vehicle**
  - One record per second
  - Mission lengths highly variable (<2 hours to nearly 8 hours)
- **Response Variable**
  - Video quality (Three-level ordinal, change points recorded)
- **Factors (observed continuously over time)**
  - AV altitude
  - AV distance to supported unit receiving video
- **Factors (constant within each mission)**
  - Supported Unit (Brigade HQ, various infantry battalions)
  - Uplink Frequency
  - Downlink Frequency

Data Characteristics of Interest

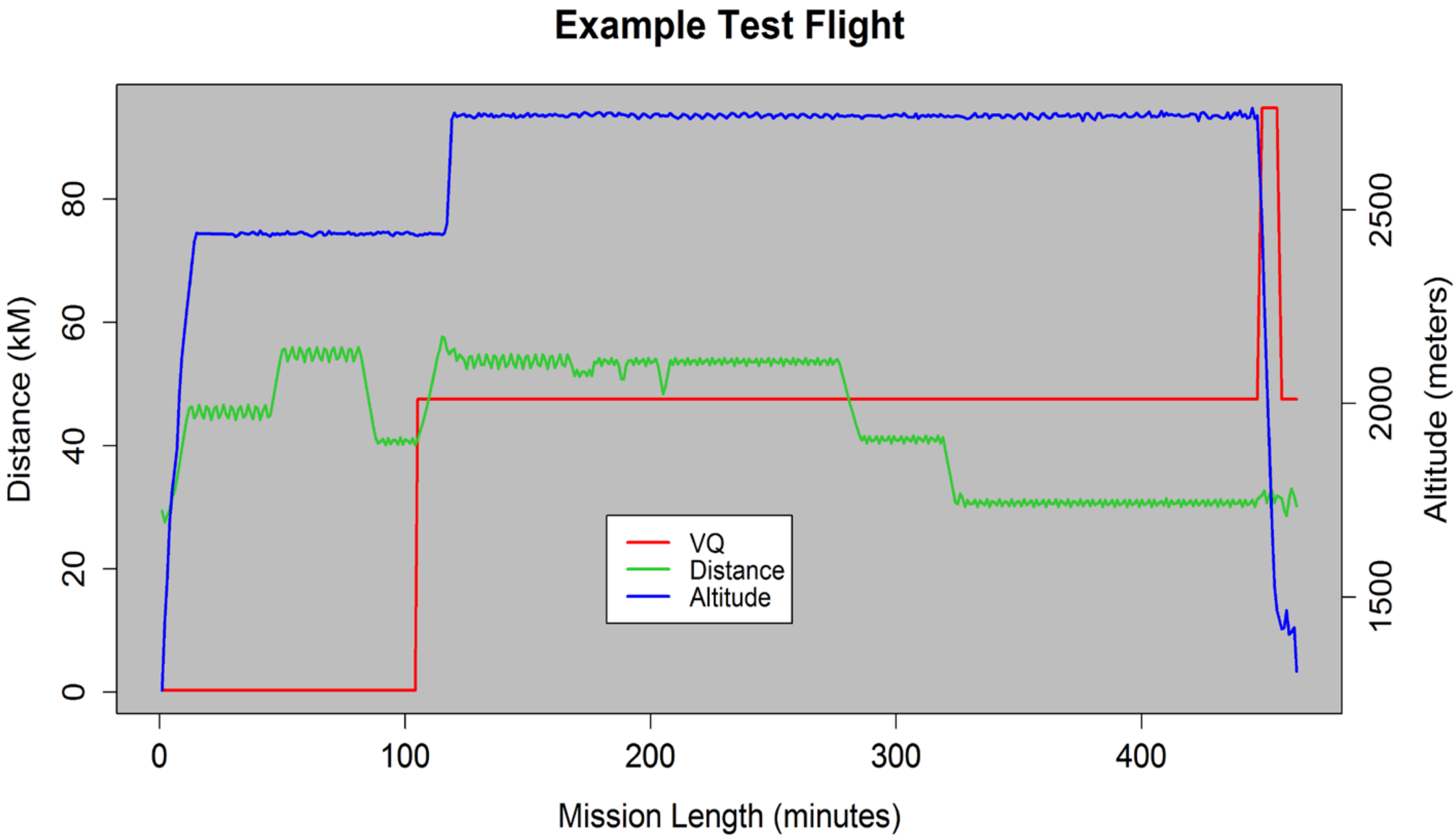
- **Ordinal response variable observed frequently over time (~1Hz)**

- **Continuous predictors observed frequently over time (~1Hz)**
  - Operations dictate that factors like range to target and altitude (AGL) are generally piece-wise constant less variations in distance during orbit
  - Changes tended to be sharp (shift in altitude from 2,000m AGL to 3,000m AGL over the course of ~10 minutes)
    - » Closely resemble step functions

- **Within-mission variability vice between-mission variability**
  - Intramission correlation high for video quality

IDA

Video Quality Data



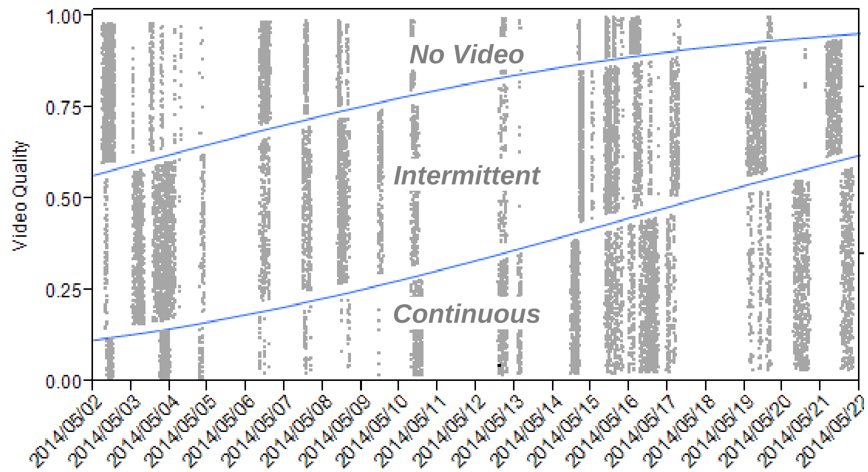
Original Analysis Approach

- **Ordinal Regression**
  - Proportional odds cumulative logit model is a standard model for ordinal data
  - Estimate factor effects for Distance and Altitude
  - Fails to account for correlation of data
- **Determining the Effective Sample Size**
  - Serial nature of the data must be accounted for
  - Estimate correlation between factor effects
  - Including all data points overstates sample size by factor of about 20
- **Analysis**
  - Randomly thinned data by a factor of 20 and estimated significance using ordinal regression
  - Monte Carlo approach for estimating p-values

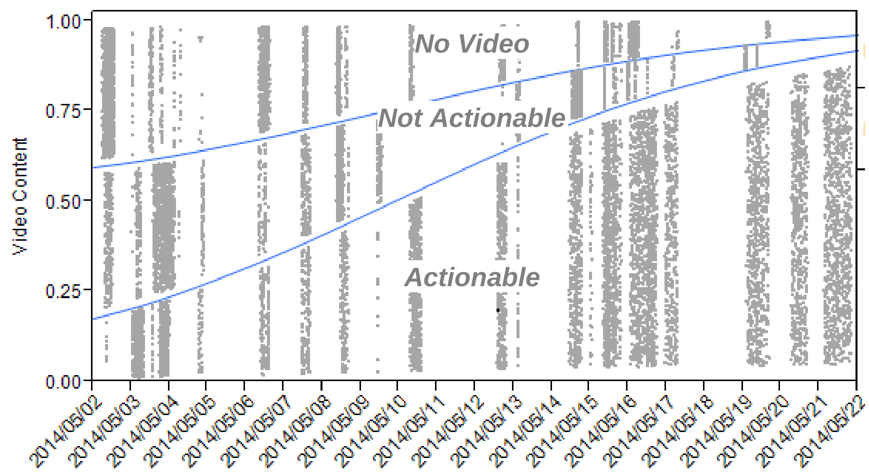
One System Remote Video Terminal
Video Quality

- **OSRVT Video Quality and Content improved over time**

Quality

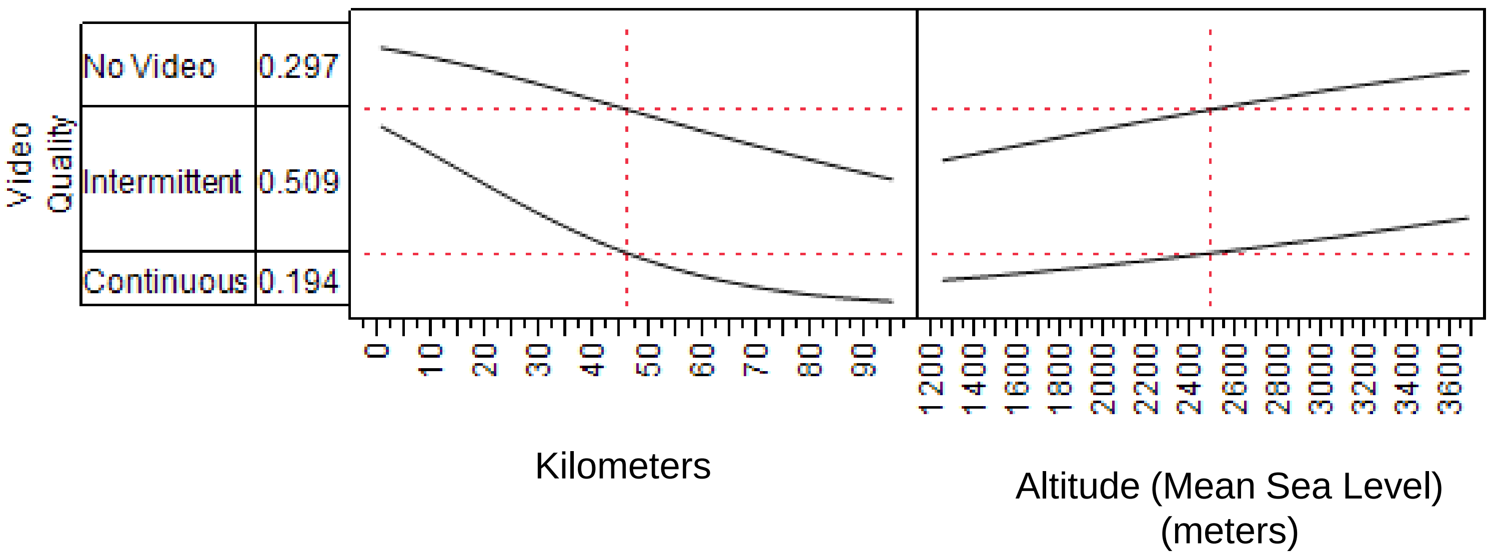


Content



One System Remote Video Terminal
Video Quality (cont)

- **OSRVT Video Quality was degraded by increased range from the Aircraft (p-value < .0001)**
- **OSRVT Video Quality improved with increased Aircraft altitude (p-value < .0368)**



Effect Likelihood Ratio Tests (using 1/20 random selection of the data)

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq	
Distance	2	2	2564.31765	<.0001*	(Average of 100
Alt	2	2	209.061571	<.0368	random samples)

Modeling the Data

- **Data are piece-wise constant**
  - Structure is fundamental to many types of UAV data, not just Shadow
  - Helps to answer, “How much data do we actually have?”

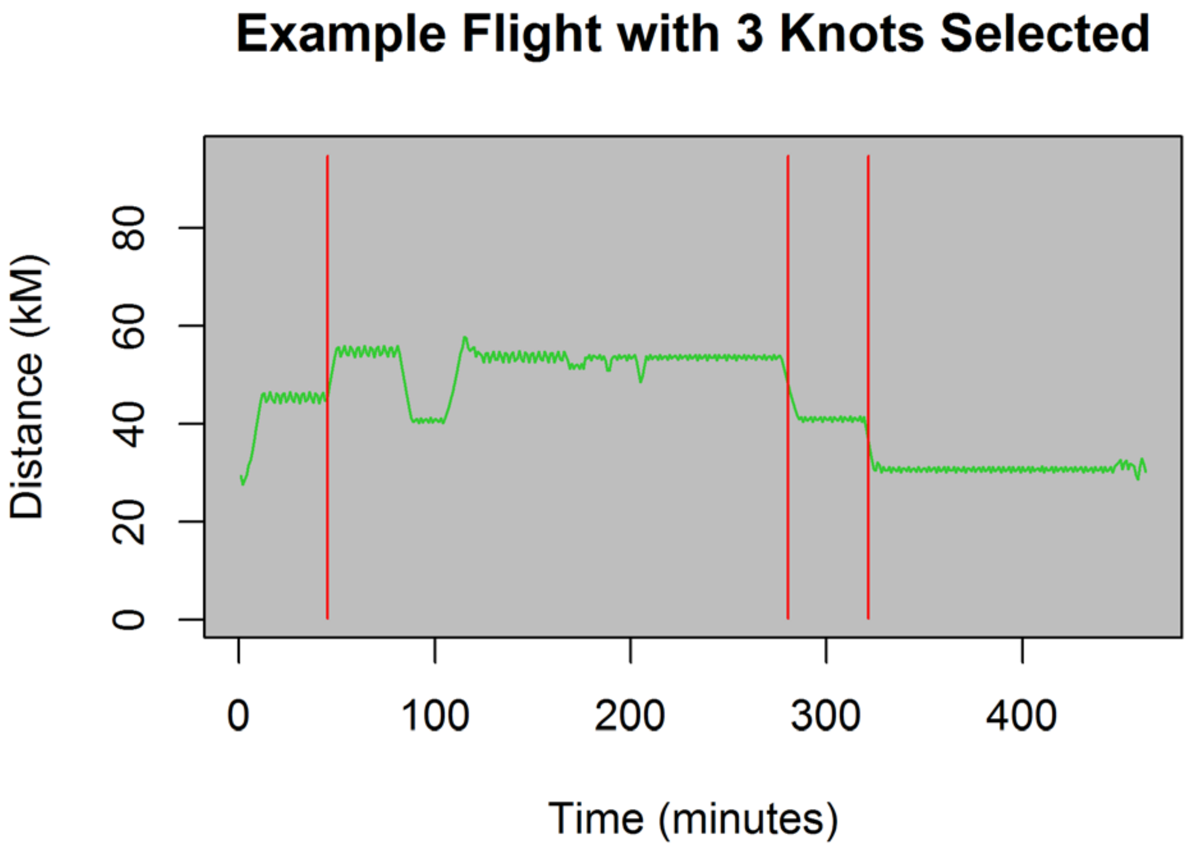
- **Within-mission correlation**
  - Uncontrollable factors that drive performance are often constant within mission
    - » Supported unit’s posture
    - » Weather
    - » Air vehicle’s state on a given day
  - Mixed model approach allows us to take this into account

Break curves into piece-wise constant segment and analyze using mixed model approach

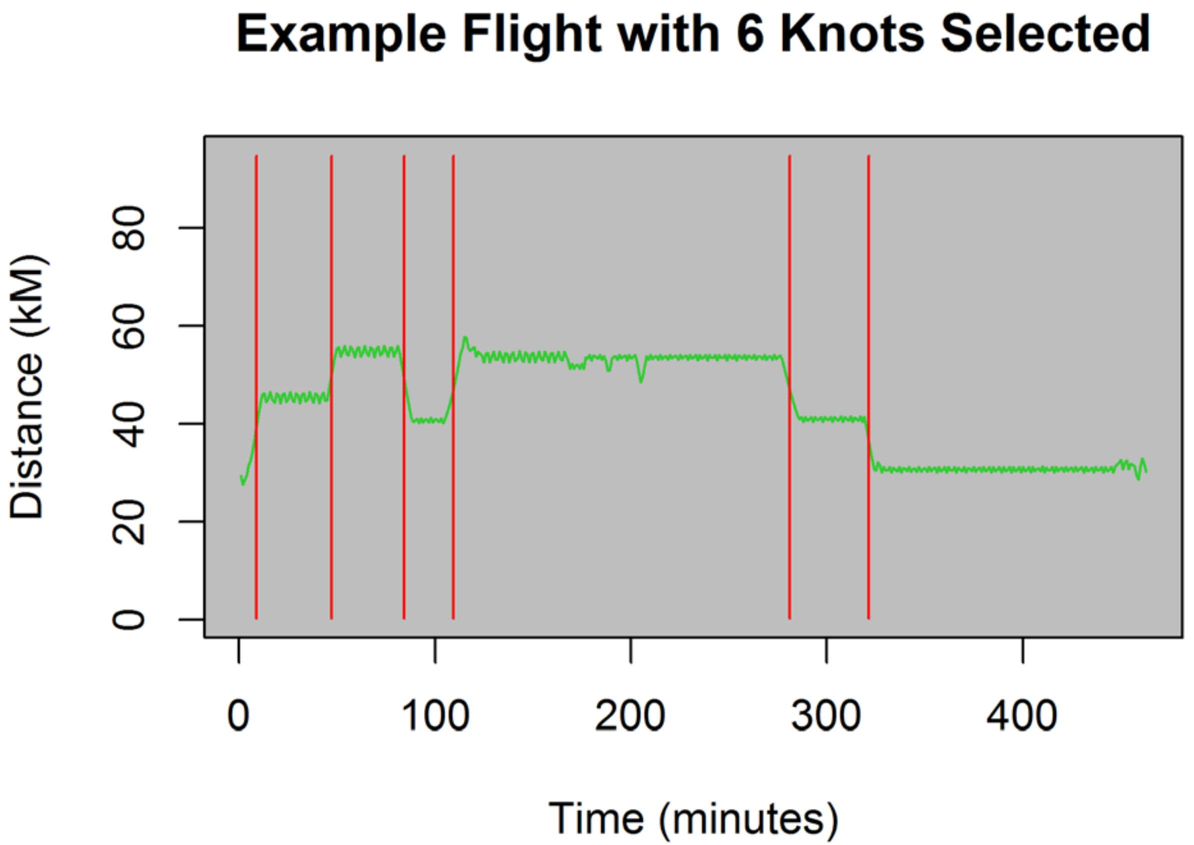
Knot Selection

- **Fit 0-degree splines using automatic knot selection algorithm**
  - Piece-wise constant
  - Place knots to minimize least-squares error over each curve
  - Number of knots determined via cross-validation
  - Penalization (based on variability of each curve) added to achieve simpler fits
    - Empirical adjustment observed to improve quality of knot selection
  - Spiriti, et al. (2013)
  - freeknotsplines package in R

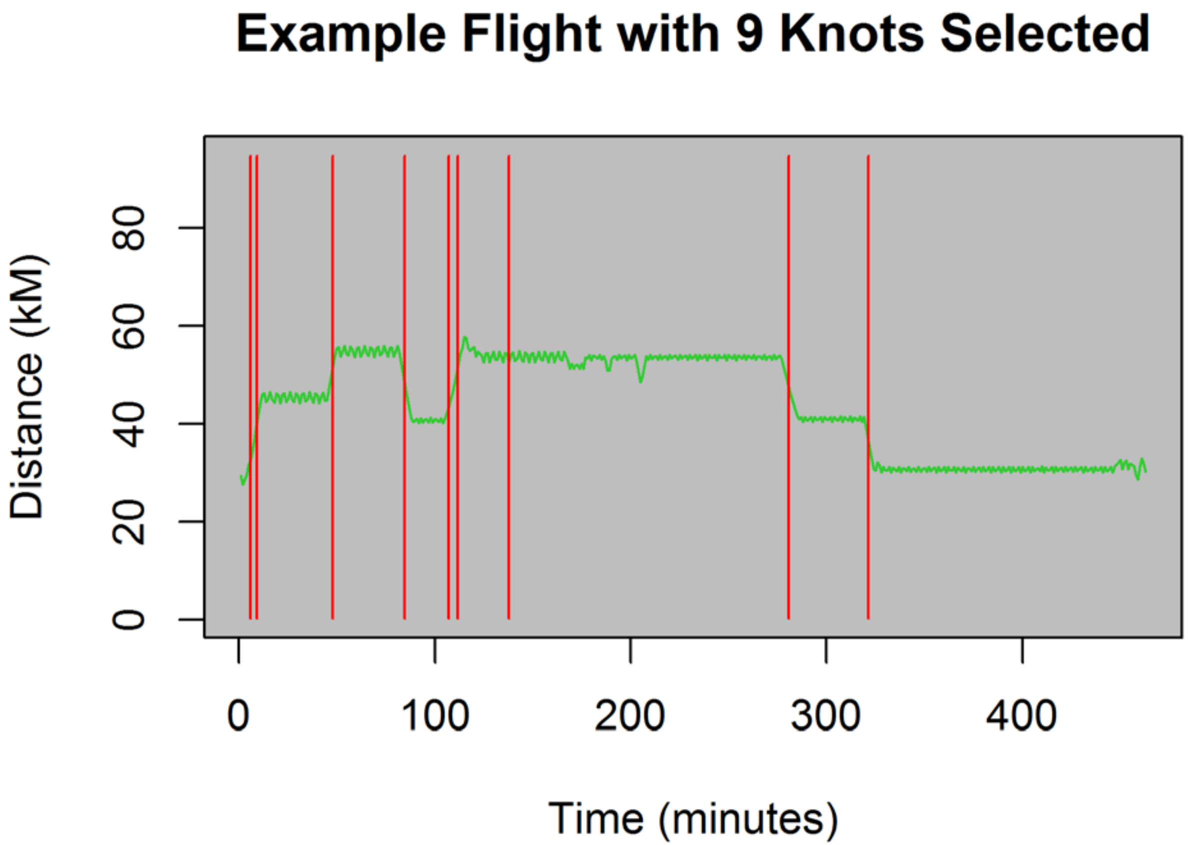
Identifying Breakpoints In Curves



Identifying Breakpoints In Curves



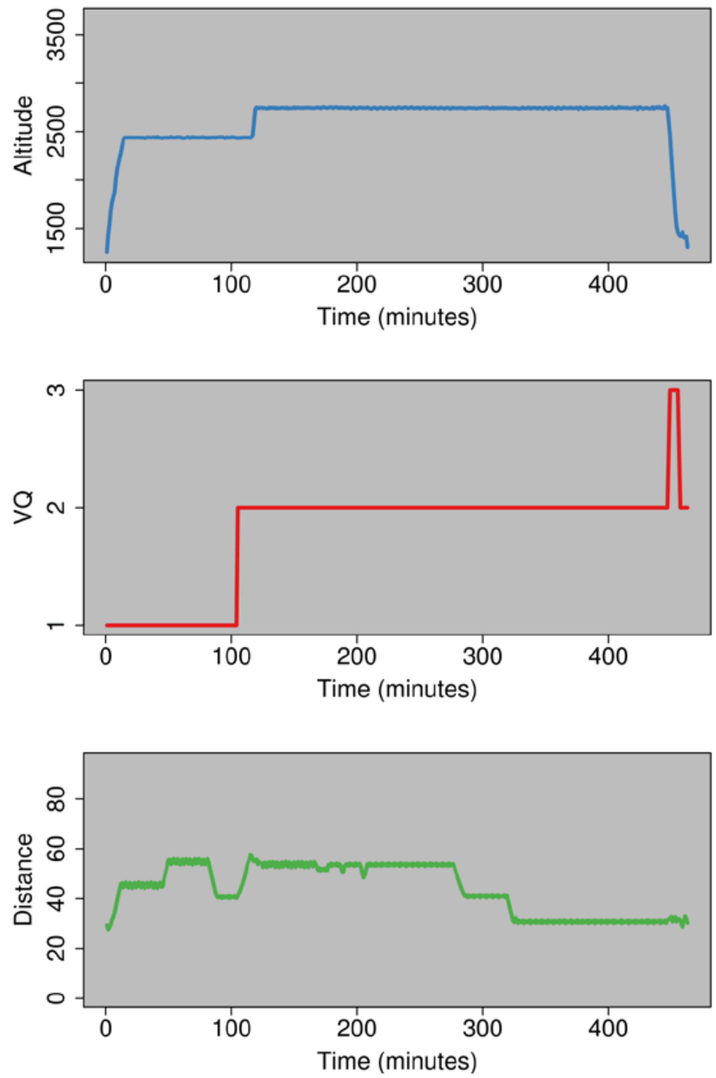
Identifying Breakpoints In Curves



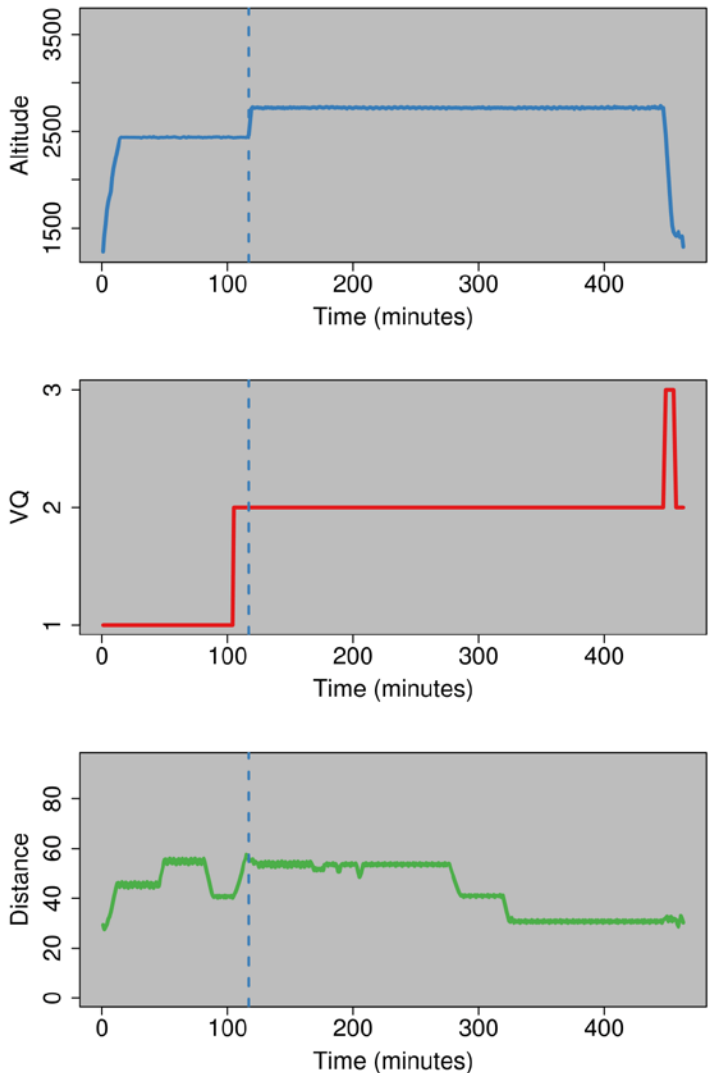
Generating Data Across a Flight

- **Each “unit of data” will consist of a time period where all continuously observed factors are constant**
  - Must identify knots for both Distance and Altitude as well as Video Quality
  - Minimum distance between knots of 5 minutes
    - » Factors may change near-simultaneously
- **Must combine results from each curve to generate data table**

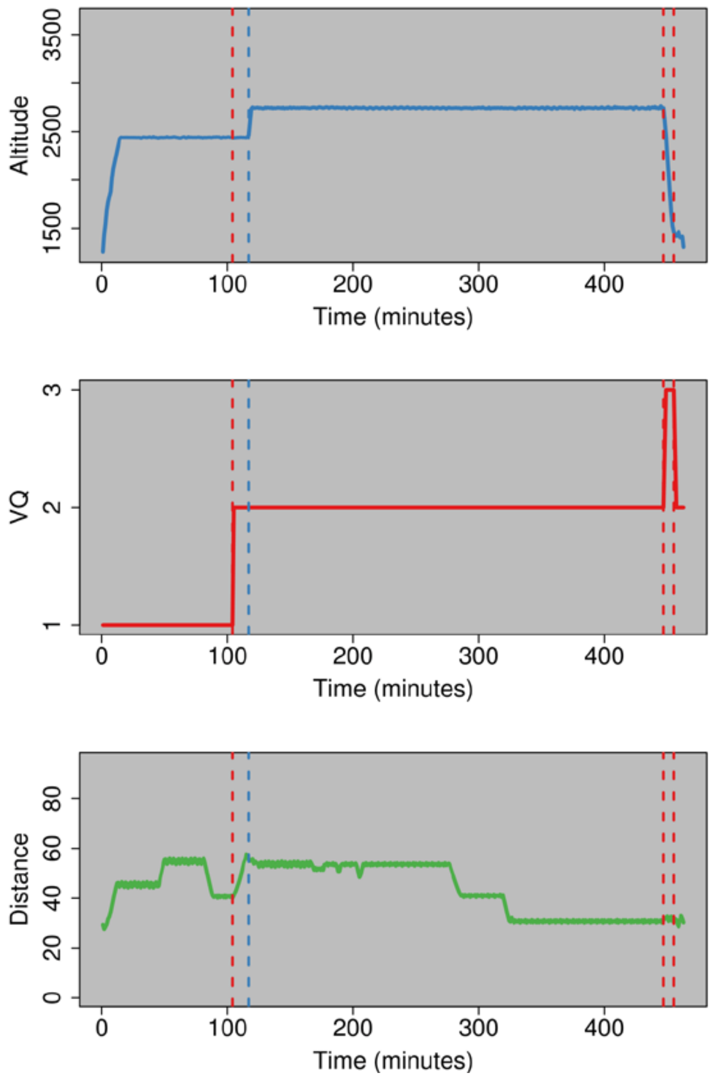
Example Flight



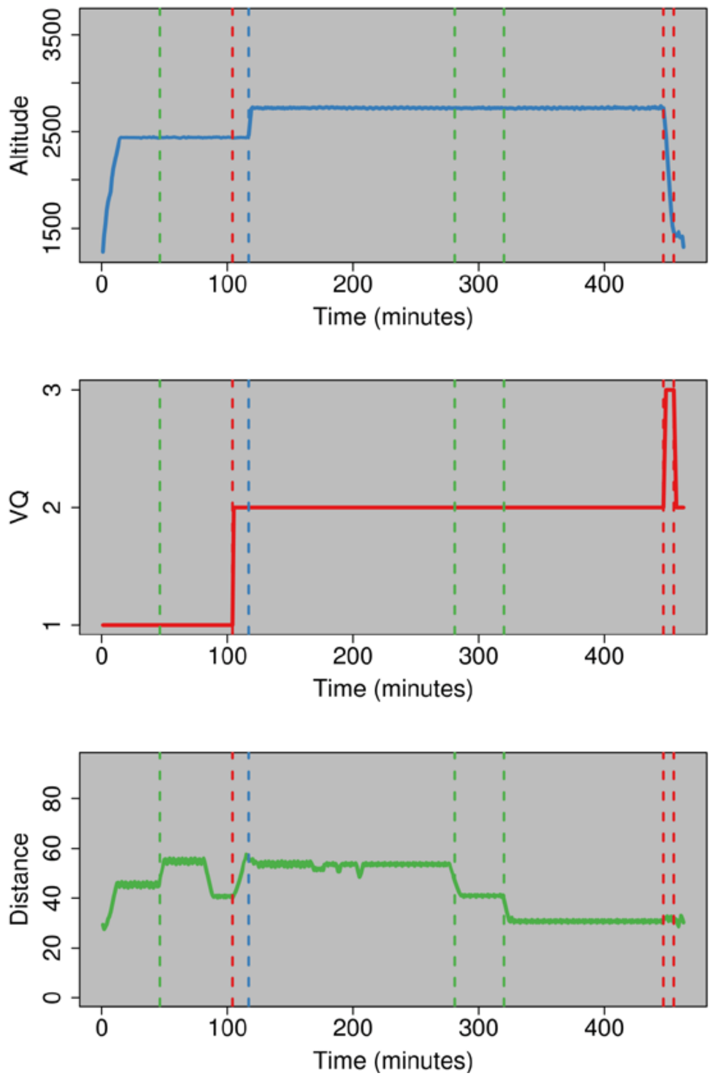
Knot for Altitude



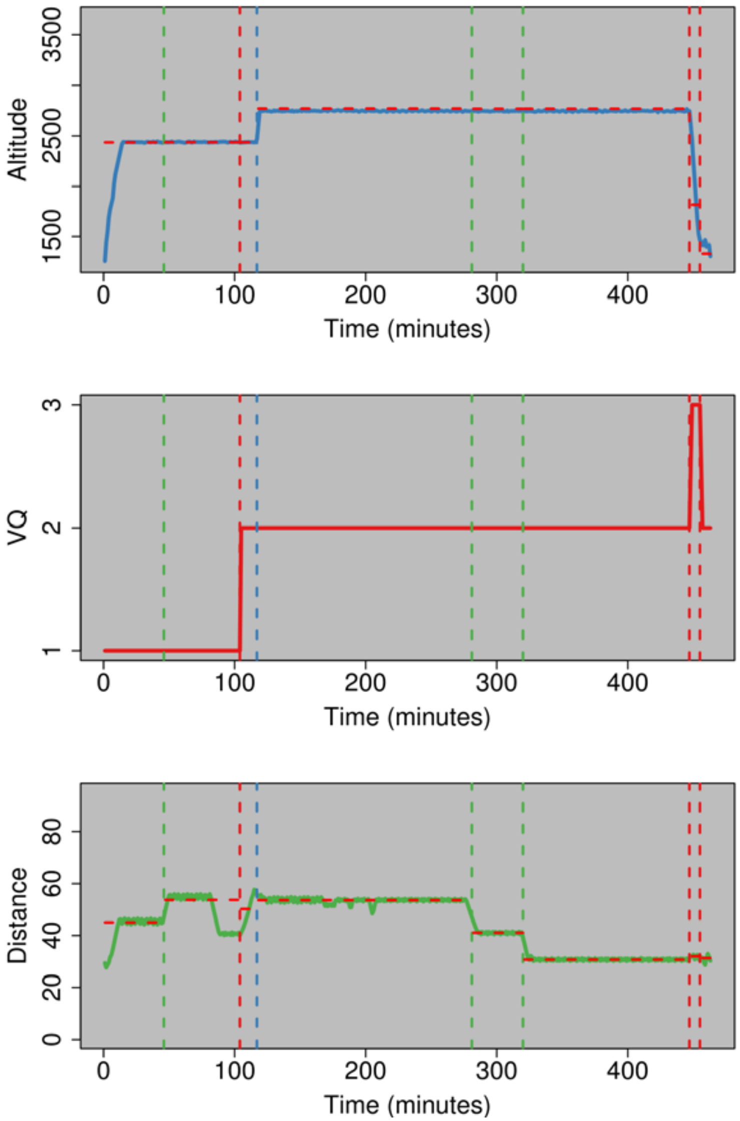
Add Knots for Video Quality



Add knots for Distance



Data Table Entries



Distance (kM)	Altitude (m)	Video Quality	Mission
45.1	2437	1	7
53.7	2438	1	7
50.2	2437	2	7
53.6	2743	2	7
41.1	2742	2	7
30.7	2742	2	7
32	1864	3	7
31.3	1417	2	7

Fit Model

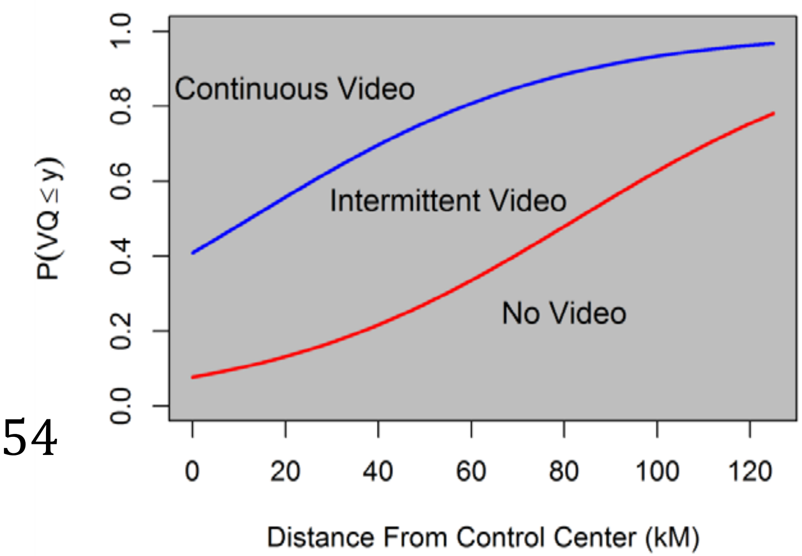
- **Generalized Linear Mixed Model fit in SAS using PROC GLMMIX**
  - Random Effect: Mission
  - Assumes compound symmetric correlation structure
  - Fixed Effects:

    Distance, Altitude, Day of Test
  - Cumulative logit model:

$$\eta_{ij} = \boldsymbol{x}_{ij}^{T}\boldsymbol{\beta} + \boldsymbol{z}_{j}\boldsymbol{b};\ i = 1, \dots, n_j; j - 1, \dots, 54$$

$$P\left(VQ_{ij} = \text{No Video}\right) = logit(\theta_1 - \eta_{ij})$$

$$P\left(VQ_{ij} = \text{No Video or Intermittent } Video\right)$$
$$= logit(\theta_2 - \eta_{ij})$$



Model Results

- **Fixed Effects**
  - Distance remains significant
  - Altitude not a significant predictor of Video Quality
  - Day of Exercise significant
    - » Crews showed superior performance at providing continuous video as they gained experience throughout the test
- **Random Effects**
  - Large mission-to-mission variation
  - Impact on Probability of Continuous video equivalent to difference in probability of continuous video between minimum observed distance and maximum observed distance

Random Effects		
Effect	Estimate	SE
Mission	1.2034	.04365

Fixed Effects			
Effect	Estimate	SE	p-value
Day of Test	-0.0995	0.03318	0.003
Distance	0.01624	0.00809	0.0459
Altitude	3.9e-05	0.00024	0.8726

Conclusions

- **Data Processing**
  - Better identification of unit of test
  - More consistent with experimental design
    - » Test plan describes tasks the air vehicle will execute throughout a flight
  - Partially accounts for serial correlation within a mission

- **Data Analysis**
  - Mixed model approach accounts for mission-to-mission variation & assess impact on system performance
  - More accurate estimation of fixed effects
    - » Altitude not statistically significant
  - Quantify uncertainty more accurately

Future Work

- **Alternative knot selection approaches**
  - Haar wavelets
  - “maximum error histograms”
    - » Data stream models for identifying changepoints
- **Alternative correlation structures within mission**
  - Autoregressive makes sense due to temporal nature of missions/flights
- **Better approaches for model selection and assessment**
  - Cross-validation approaches to assess prediction accuracy
    - » Cross-validation at mission-level
  - Properties of statistical tests for significance of random effects (Myers, et al, 2010)
  - Explore optimal weighting scheme for length of observations